DEPARTMENT
OF POLITICAL AND
SOCIAL SCIENCES

From Diversity to Mixing?

Socioeconomic Homophily in French Desegregated Middle Schools

Timothée Chabot

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Political and Social Sciences
of the European University Institute

Florence, 20 December 2021

European University Institute
Department of Political and Social Sciences

From Diversity to Mixing?

Socioeconomic Homophily in Desegregated French Middle Schools

Timothée Chabot

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Political and Social Sciences
of the European University Institute

Examining Board

Prof. Fabrizio Bernardi, European University Institute (EUI Supervisor)
Prof. Agnès van Zanten, CNRS / Sciences Po Paris (External Supervisor)
Prof. Arnout van De Rijt, European University Institute
Prof. Pierre Mercklé, ENS Lyon

Researcher declaration to accompany the submission of written work
Department of Political and Social Sciences – Doctoral Programme

I, Timothée Chabot, certify that I am the author of the work “From Diversity to Mixing? Socioeconomic Homophily in Desegregated French Middle Schools” I have presented for examination for the Ph.D. at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the Code of Ethics in Academic Research issued by the European University Institute (IUE 332/2/10 (CA 297).

The copyright of this work rests with its author. Quotation from it is permitted, provided that full acknowledgment is made. This work may not be reproduced without my prior written consent. This authorization does not, to the best of my knowledge, infringe the rights of any third party.

I declare that this work consists of 160 673 words (including tables, footnotes and bibliography, excluding the appendices).

Statement of inclusion of previous work:
I confirm that chapter 8 was jointly co-authored with Ms Marion Hoffman and I contributed 50% of the work.

Statement of language correction:
This thesis has been corrected for linguistic and stylistic errors. I certify that I have checked and approved all language corrections, and that these have not affected the content of this work.

Signature and date:

Noisy-le-Sec, the 01/12/2021

Abstract

Socioeconomic mixing at school is often considered to be a desirable objective, as it would reduce academic inequalities and help pacify inter-group relations. For this reason, socioeconomic segregation, understood as the spatial distribution of students across different schools, has been extensively studied in educational sociology. However, the interactions among students in cases where schools are effectively diverse have been notably under-explored. Does spatial diversity imply relational mixing, or do students keep interacting with socioeconomically similar peers even in formally desegregated contexts? This raises the question of *homophily*, the principle by which relationships occur at a higher rate among similar individuals.

In this dissertation, I study socioeconomic homophily among a cohort of 860 middle school students in four schools in France, followed during three years. Based on the statistical analysis of students' friendship networks and on qualitative interviews, I examine the magnitude of this homophily, and try to disentangle the relational processes through which it emerges. In particular, I consider the respective roles of dispositional and contextual factors: to which extent are friendship pairings driven by internalized traits – preferences, tastes, ways of feeling and behaving –, and, on the contrary, by the constraints and incentives that the immediate environment exerts?

Results suggest that there generally is socioeconomic homophily among adolescents, but that its magnitude drastically varies across schools. These differences are explained by a multiplicity of factors, as socioeconomic homophily emerges through the joint effect of several relational processes. Among these, the psychological tendency to search for socioeconomically similar friends only accounts for a minority of the total homophily; spatial constraints – such as classrooms and places of residence – and, most importantly, network and group mechanisms – notably transitivity –, play a key role as well. Finally, results hint at a strong variation in the local salience of socioeconomic attributes: depending on the context, youth's socioeconomic origin can be made particularly visible, or on the contrary be "flattened" and minimized, which may in turn explain students' propensity to base their friendship behaviors on this particular criteria.

Acknowledgments

In the course of the past four years, many people have contributed, directly or indirectly, to this research. The present dissertation would not be what it is if it was not for them – in fact, it would simply not exist. To all of them: thank you!

First of all, I would like to express my thanks to my two supervisors. At the EUI, Fabrizio Bernardi has followed my work for four years, and I have benefited from his guidance and support. Even as I progressively drifted away from his own research areas, he kept encouraging me to follow my interests and curiosity, and he helped me find resources and people that could push me further in this direction. Among other things, he strongly advised me to go to the university of Groningen for an exchange and recommended me there, which proved to be a key step in the course of the PhD. I was lucky to find a supervisor that put up with my desire for autonomy (should I say stubbornness?) the way he did. At Sciences po, Agnès van Zanten accepted to carry on as a PhD supervisor after she had already supervised my Master thesis in 2016/2017. I have benefited greatly from her extensive knowledge of the literature in the sociology of education, and from her deep understanding of the French education system. She also had the difficult task of keeping me in touch with qualitative methods and perspectives, as I happily ran toward more and more quantitative razzle-dazzle. Finally, it was thanks to her that I met Sylvie Da Costa, who opened up two of my five fieldworks.

Next, my thanks go to the external jury members, Pierre Mercklé and Arnout van de Rijt. Their comments on the first version of the manuscript proved extremely valuable, both for the dissertation itself and for giving me ideas about future research and publications. I must also thank Arnout for the time he took to talk with me about a year ago, while I was still maturing important ideas; and Pierre for the opportunity he gave me to come talk in his classes about my work.

Several other academics have provided decisive help at different times during the PhD. René Veenstra welcomed me at the University of Groningen and at the ICS, where I met a formidable group of experts in Social Network Analysis. I particularly benefited from the advice of Marijtje van Duijn, Robert Krause, and of course Tom Snijders, who – in addition to inventing half of the methods used in the dissertation – took some time to listen to the confuse nonsense that would eventually give birth to contribution scores. I must also thank Andreas

Flache for his insights on agent-based modeling and how it could intersect with network analysis. I am also indebted to the members of the Statnet and Socnet mailing lists, which answered my questions on several occasions and guided me through my discovery of SNA methods; in particular Robert Klause, Tom Snijders (again!), Per Block and Philippe Leifeld. Paulo Serodio also helped me deepen my understanding of ERGMs. In France, Sylvie Da Costa gave me access to the two schools of Paris 2 and Savoie 2, thanks to her contacts at the general secretariat of catholic schooling. She has also shown great interest and enthusiasm for my work over the years (which is honestly the one thing that PhD researchers crave the most). Pauline Proboeuf opened the doors of Paris 1 for me, as well as she gave me precious advice on post-doctoral positions in France. Though not directly related to the thesis, I also want to express my gratitude to Philippe Bézès, for bringing me on board with SOG PRO and for his cheerfulness, as well as to Pierre Blavier, for the Méricourt adventure and for the postdoc call to candidacy that he told me about.

Special thanks go to Marion Hoffman, whom I also met at Groningen and who co-authored chapter 8 with me. The brilliant innovation that is the ERPM gave some unexpected value to the summer camp data I had collected, and it now opens the way for exciting new research. Marion was also kind enough to share her experience and deep understanding of network models with me, as well as she recommended me important journal articles, both of which proved important to the statistical analyses developed in chapters 5 and 7. As a bonus, we got to complain to each other about our respective issues with our PhDs (which is the second thing that PhD researchers crave the most).

The persons most essential to this research are the ones that must unfortunately remain anonymous: the members of the four schools and of the summer camp in which the data collection took place. I want to express my gratitude to the headmasters of Paris 1, Paris 2, Savoie 1 and Savoie 2 for welcoming me to their schools, as well as to the teachers, the administrative staff and the members of the *Vie Scolaire* that helped me to organize the survey and that shared part of their professional experience with me. Special thanks to Thierry, Chantal and Emmanuel for their warm welcome and their help. Regarding the summer camp study, Carine and Sandrine made my presence to the camp possible. During the three weeks of the study, the entire staff has been incredibly helpful and supportive, and I had tons of fun working, playing, dancing and (on occasions) drinking with them: Marie, Rémi, Claire, Livio, Hugo, Milou, Ludo, Solène and all the others. Thanks as well to Léa, for collecting data in the pilot study of 2017 (and for being a great friend in general). And finally, my biggest thanks must go to all the adolescents that took part in the study, both in the schools and in the camp. Not only are their answers the raw material out of which the entire dissertation is made, their energy and humor has also inspired me all along the thesis.

On a more personal note, I want to thank the friends that have supported me during this long journey: Quentin, for his brilliant mind and for all the discussions that fueled my theoretical thinking, both about the PhD and social science in general; Paul and Pef, for their presence over the years and for their interest for my work; Raoul, for his enthusiasm and kindness when I needed it the most. Thanks as well to Louis, for welcoming me out of the blue, helping me save a relationship I had nearly broken, and turning me into a star of the internet; and to Fanny, which made it all possible in the first place. Thanks to Thérèse and Jean for their kindness during the final push.

I must of course thank my parents for their help and support. As a sociologist working on educational inequalities, I know all too well how much early socialization and the transmission of cultural capital matters in one's academic success. I also know that this transmission is not automatic nor easy, but that it demands constant educational work on the part of parents (even for the positively adorable kid that I am sure I have been). I would not be writing a PhD if this work had not been done with great care and dedication throughout the years. Thanks as well to my brother, who had the brilliant idea of becoming a mathematician (something very useful to have in one's relationships when doing statistics!). In particular, it was thanks to him that I first understood the inner-working of the ERGM and what it substantially implied.

It is not entirely clear to me whether I should thank Masahiro Sakurai or curse him. Either way, Inkling, R.O.B, Lucina and Banjo have been faithful companions all along the PhD (and Greninja to an extent, though I feel like we did not get the story we deserved).

Finally, I want to thank the *chapougnes*, Trois-Pattes and Acha-Pacha (bonjour!), for their headbutts, trampling and grumbling. They are the ones that spent the most time with me as I was writing the dissertation – on top of being the most adorable creatures ever seen on the face of the Earth (objectively speaking). And of course, the most special thank goes to Louisa. It is hard to know where to start, as the list of everything she has done for me seems endless. No one has red, commented and heard about my thesis as much as she has, and her feedback has proven incredibly valuable on countless occasions. The most important, however, is elsewhere – I think she will understand without me spelling it out.

Table of Contents

List of Tables

List of Figures

Introduction

Friendship relations, though they may be perceived by individuals as intimate, personal relationships, are a key dimension of social stratification. They are both a vector of socialization, that shape one's tastes, values, habits, or worldviews, and a valuable capital granting access to important resources. For these reasons, the socioeconomic segregation of friendship networks contributes to the maintenance and reproduction of class inequalities. One of the strongest regularities found by empirical studies about friendship in our societies is *homophily*: friendship ties are disproportionately likely to be observed among individuals that share important characteristics, such as sex, age, ethnicity, or social class (Bidart, Degenne, and Grossetti 2011; Lazarseld and Merton 1954; Verbrugge 1977). Among these, the homophily of social class, or *socioeconomic homophily*[1], appears as a major determinant of friendship formation among adults (McPherson, Smith-Lovin, and Cook 2001).

How precocious is this particular form of homophily? In other words, is it also to be found among children and adolescents? There are at least two reasons for shifting the focus away from adults and toward youth specifically. First, the social life of children is an object of interest in and of itself, and socioeconomic homophily among children may have different implications than among adults. For instance, it may impact schooling inequalities and, more generally, facilitate the inter-generational reproduction of inequalities; or it may affect the well-being of youths and their opportunities to develop fulfilling relationships, particularly in the school context. Second, from the perspective of sociological theory, the study of youth friendships offers crucial insights into the underlying processes of homophily at large. This is because they are at an early stage of their lives, characterized by intensive friendship formation as well as a relative indeterminacy of certain personal traits. Moreover, many youths find themselves in much more diverse relational contexts than what adults typically experience. For these two reasons, children and adolescents provide an opportunity to look into the origins of socioeconomic homophily across the life-course, and to identify several factor contributing to its emergence that are, in the case of adults, too intertwined for such an analysis to be possible.

1 I will use the adjective "socioeconomic" to refer to the notion of social class. This is because the adjective "social" has an ambiguous meaning, and can be understood as referring to social class (e.g. "social origin"), to sociability (e.g. "social life") or to society at large (e.g. "social norms").

Socioeconomic Homophily among Youths: School Segregation and Social Inequalities

The impact of socioeconomic segregation among youths has long been a concern for the sociology of education and of inequalities. Mostly, the focus has been on spatial segregation across schools: depending on their socioeconomic and ethnic backgrounds, students attend different tracks, schools and classrooms. Taking the example of France, on which this project focuses, the literature is abundant (Delvaux and Van Zanten 2006; Merle 2012; Oberti, Préteceille, and Rivière 2012). The prime reason for this interest is that school segregation contributes to inequalities of academic achievement. Indeed, the socioeconomic composition of a school has an impact on students' results, controlling for observable individual-level characteristics (Coleman [1966] 1995), and the same is true of classrooms and working groups (Duru-Bellat and Mingat 1997). This means that separating upper and lower-class students aggravates the initial advantage, due to family socialization, of the former over the latter. Moreover, beyond academic performance, students in highly segregated working-class environments are exposed to highly homogeneous socialization contexts that reinforce their class-specific dispositions, such as ways of speaking, cultural practices and moral values (Lepoutre 1997; Willis 1977). Conversely, at the other end of the social spectrum, the segregation of upper-class students contributes to the acquisition of valuable cultural and social resources from an early age, and helps parents to maintain control over their child's education (van Zanten 2015). Educational segregation thus contributes to social reproduction.

The concerns about school segregation extend straightforwardly to the question of socioeconomic homophily in students' friendships. Assuming that certain schools are diverse in terms of the composition of their population – which, as we shall see, is indeed the case for some of them –, there is no guarantee that this spatial diversity will entail relational mixing. Therefore, if the alleged virtues of a desegregated school system lie in the interactions between students from different backgrounds, then studying the composition of schools is not enough: it is necessary to go down at the level at which these interactions occur. Strongly segregated friendship networks within diverse schools might not offer much of an advantage in terms of inequality reduction compared to outright spatial segregation[2].

2 There are likely certain benefits of spatial mixing that play out at the level of schools' or classrooms' aggregate features, such as the quality of teachers, the frequency of classroom heckling, or the access to information regarding orientation and career paths. In that regard, the effect of school diversity on academic equality does not necessarily depend on peer-to-peer relations exclusively. Still, horizontal socialization among students counts for a large part of the learning processes that occur at school, so it seems quite implausible that

Furthermore, one thing that is rarely discussed in the literature is whether socioeconomic diversity may sometimes have negative consequences. The extent to which students from different backgrounds get along with one another may impact their well-being at school, as well as it may affect their perception of other social groups and leave lasting implications on their prejudices or political opinions in later life. In this regard, it would be difficult to argue in favor of socioeconomic diversity if it turned out to heighten conflicts among students, to elicit antagonism rather than friendship, and to deteriorate their image of socially different peers. Incidentally, the fear of conflict or violence among students is one of the reasons put forward by upper-class parents to explain their avoidance of diverse schools (van Zanten 2015), though we lack solid evidence about whether such a fear is justified.

The Origins of Socioeconomic Homophily: Preciousness and Rigidity of Class Dispositions

Beyond its intrinsic interest, the study of socioeconomic homophily among children and adolescents offers valuable insights into the principles underlying friendship formation at large. In doing so, it also informs us about the precociousness of the internalization of certain class-specific dispositions by children and adolescents and in turn, about the mechanisms through which socioeconomic homophily gets reproduced in society at large.

To understand why this is the case, it is first necessary to consider how inter-individual relationships are formed. A long-standing question among friendship and education scholars is that of the respective roles of contextual and dispositional factors in shaping sociability practices. We shall come back to this distinction in detail in the body of the dissertation, but it can be summarized as follows. Friendship formation can be conceptually separated into two components, “meeting” and “mating” (Verbrugge 1977). “Meeting” refers to the fact that it is first necessary for people to meet before becoming friends. As a consequence, friends are always chosen within a pool of potential partners, which depends on the sociability contexts in which one evolves: the composition of a person’s school, neighborhood, activity club or workplace strongly determine what potential friends they have access to[3]. However, even for a given amount of daily contact, not all relationships are equally likely to emerge. The second

the extent of relational mixing in a school would make no difference in regard to these outcomes.

3 Interestingly, this may not be as true for online relations, which have become increasingly important in youths’ social lives over recent years (though many online relations are extensions of face-to-face ones, so the effect of spatial segregation should not disappear entirely). In any case, I focus here on face-to-face relations, online ones being arguably different in nature (which is not to say, though, that they are not “real” friendships, as commentators are often quick to conclude).

step of friendship formation is therefore “mating”, the process by which individuals select certain friends – consciously or not – among the persons they have met. This is dependent on a number of individual traits that make certain relationships more attractive, easier to form, or more resilient in time.

Of course, there are a number of ways in which these two components interact (e.g. choosing to attend certain sociability contexts in order to meet a certain type of person), but the distinction remains analytically useful. To what extent are friendship pairings driven by internalized traits – preferences, tastes, ways of feeling and behaving – and on the contrary, by the constraints and incentives that the immediate environment exerts? Are friends “chosen” following elective affinities, or rather “found”, through the interplay of contact opportunities in class, in the neighborhood and within peer groups? These questions are central to the study of homophily and even friendship in general, as they have critical theoretical and practical implications. A relational feature that is primarily driven by dispositional selection will be affected by changes in individuals’ socialization, perception of others or expectations about their relationships. On the contrary, one that depends on contact opportunities will be more sensitive to the spatial and temporal organization of social activities.

The study of socioeconomic homophily at school constitutes a particularly rich entry into these questions. Social class – or, in the case of children, socioeconomic origin[4] – strongly affects both dimensions, the contextual and the dispositional. Indeed, socioeconomic groups live and work in different places, attend different schools or go to different places for their leisure; as well as having different tastes, values, worldviews, ways of speaking or even gestural habits[5]. The question of these factors’ respective roles, as well as of their interactions, is thus particularly salient for this type of homophily. However, in the case of adults, disentangling them may often prove impossible. Friendships are built all along the life course, as individuals often keep friends from different periods of their life (Bidart et al. 2011). The state of a person’s network will thus depend on the composition of the many different contexts that they have lived through, but also on the type of disposition-driven selection that operated in each of these spaces, as well as on the differential survival rate of established relations over time and changing contexts.

4 I use the terms “socioeconomic origin” and “socioeconomic background” interchangeably; the choice of one over the other in different parts of the text is simply meant to limit repetitions.

5 References supporting this view will be presented in chapters 1 and 2.

By contrast, studying homophily among school students offers a number of methodological advantages. First, schools offer well-bounded spatial and temporal settings, making the collection of reasonably exhaustive data about the state of relationships in a given population possible. The fact that students generally have little control over key elements of the structure of contact opportunities – classroom, school and neighborhood – also makes it easier to distinguish their effect from that of dispositional factors. Of course, it remains impossible to capture all the relevant factors in a single study, but this is a significant improvement compared to the study of adults' personal networks. Second, school years constitute a period of intensive friendship formation: the sheer volume of relational change that occurs in a relatively short time provides greater statistical leverage to assess the role of different combinations of factors on the probability that relationships have to emerge and persist. Third, certain schools exhibit high levels of socioeconomic diversity, offering valuable case-studies to see whether spatial desegregation can entail relational mixing. By contrast, in entirely segregated spaces, the question of the dispositional factors that may affect homophilic bonding is simply irrelevant, as there is no opportunity for socioeconomic mixing whatsoever. As we shall see, diverse schools are by no mean rare in France; and middle school in particular may be one of the most diverse periods of the life course for many individuals (even though there is still segregation, in the sense that diversity is less important than what should be expected under a random allocation of students across schools)[6].

Finally, one interesting characteristic of children and adolescents is that they have yet to join the workforce, or even to clearly engage in a career path (at least before the split between general and vocational teaching). Consequently, they do not really belong to a social class, in the sense that adults do; rather, they are characterized by a certain socioeconomic origin, typically defined from the occupation and education of their parents. This allows for an examination of the specifically transmissible dispositions that may affect homophilic behaviors: if children present certain habits, beliefs or tastes that push them toward socially

6 As a side note, it is interesting that in the French case at least, many public discourses that regret the lack of socioeconomic diversity focus on middle school. For example, in 2016, the economist Thomas Piketty engaged in a debate in the columns of the newspaper *Le Monde* with the Minister of Education of the time, Najat Vallaud-Belkacem, about the lack of socioeconomic diversity in middle schools (Battaglia 2016). Yet most other social contexts are likely far more segregated than middle school, without raising as much public concern (in particular, socioeconomic segregation is higher at the university and on the labor market). Most likely, this is linked to the particular place that schooling occupies in the French political imaginary, as it concentrates many beliefs and expectations about meritocracy, social equality and national integration. On the myths associated to public education in western societies, and their particular prominence in France, see Duru-Bellat (2015), introduction, pp.1-4.

similar peers, then it means that socioeconomic homophily pertains, at least partly, to heritable traits, rooted in early socialization.

Of course, school students are not only under the influence of their parents; peer-to-peer socialization in particular, plays an important role as well. This is especially true of teenagers, who experience a shift from parents to peers as the main normative group of reference (Adler and Adler 1998; Octobre et al. 2010). Therefore, in diverse schools, a potential tension exists between the dispositions transmitted by the family, and that which may be learned from peers of different origins. As a result, the strength of socioeconomic homophily in students’ friendships will not only depend on whether they have already acquired certain class-specific dispositions that impact friendship formation from their family, but also on how rigid or resilient these dispositions happen to be when exposed to more diverse socialization contexts. Simply put then, strong homophilic behaviors on the part of students would suggest that psychological barriers to the formation of mixed friendships exist from an early age, and that these are most likely transmitted by family socialization. On the contrary, if students’ inclinations toward socioeconomically similar friends are weak, then one may be willing to conclude that such psychological barriers are either irrelevant in general, or that they are formed later in life, and that secondary socialization from peers can over-write them to an extent. This is a (very) simplified presentation, and, as we will see, students’ relational patterns ought to be interpreted with greater care; for example, considering the role of teachers, or that of residential segregation within the recruitment sector of a school. The generalization from the origin of socioeconomic homophily among teenagers, to its origin among adults, is not straightforward either. The point remains, though, that adolescence is a period of relative dispositional flexibility, which together with the spatial mixing found in certain schools, allows for an examination of the precociousness of certain homophilic dispositions. In turn, this can help us to understand the origin of these dispositions among adults, and the mechanisms through which socioeconomic homophily gets reproduced from one generation to the next.

A Gap in the Literature

Despite all this, socioeconomic homophily between school students has been remarkably under-studied so far. Although seminal works in the sociology of adolescent friendship placed it at the core of their preoccupations (Cohen 1979; Coleman 1961; Hollingshead 1949), there are very few recent studies that directly tackle this subject. Among these, Papapolydorou (2014) for the United Kingdom and Malacarne (2017) for the United States make a similar remark: little is known about how socioeconomic origin affects friendship formation among children and teenagers. This comes as a surprise, because the question of socioeconomic homophily at school stands at the intersection of extremely dynamic fields of research: the sociology of education, of social reproduction, and of sociability and friendship bonds. Yet this may be the very reason why this topic has received little attention: standing at the edges, so to speak, of these different fields, it grazes their core concerns, but mostly falls in-between their respective domains of investigation.

A more complete literature review is proposed in chapter 1, so I will not go into detail here. However, to give a rough sketch of the sociological literature that the thesis engages with, two main bibliographic clusters can be identified. On the one hand, teenagers' friendships are the focus of an abundant literature in Social Network Analysis (SNA) which has developed elaborate quantitative tools to study human relational structures. The notion of homophily is pivotal to these works, as it is acknowledged to be one of the dominant features of friendship networks in general, and of adolescent ones in particular (Block 2018; McPherson et al. 2001). However, socioeconomic origin has attracted little attention; homophily is more frequently envisioned in terms of gender, race and ethnicity, or academic achievement and attitudes. Even when it is considered, it is primarily conceptualized in terms of economic resources and consumption power, leaving aside questions of cultural differences among socioeconomic groups, and therefore of students' class-specific socialization.

On the other hand, socioeconomic origin and the inequalities that it entails among youth is a core concern of the sociology of education and more precisely, of social reproduction, notably within the French and British traditions. This time however, it is homophily, and more generally, the principles driving friendship formation among peers, that are largely left out of the focus of researchers, for at least two reasons. First, in France particularly, the sociology of education has long been marked by a Durkheimian perspective that essentially saw students as passive entities waiting to be socialized. This remained true even for later critics of Durkheim's model – for instance Bourdieu and Passeron ([1970] 1990) or Boudon ([1973] 2011) – who remained primarily concerned with the distribution of degrees and with labor-market outcomes (Duru-Bellat 2015, pp. 22-29). This contrasts with an American tradition that has often considered schools as "small societies" (Coleman 1961), and of which the SNA literature partly comes from. Second, even though some works in this literature have nevertheless studied student social life, they have often done so in segregated working-class environments. Coherent with the key concern about spatial segregation that I mentioned earlier, ethnographers have examined in great detail the processes of mutual influence among working-class students in highly segregated schools. Socioeconomically diverse schools have however received less attention. To be fair then, the constitution of homogeneous friendship groups among children and adolescents has been considered by this literature, but essentially through the prism of spatial closure and contact opportunities, more so than through the prism of internalized dispositions or the purposive selection of certain friends.

In the present work, I strongly advocate for the integration of these two approaches. Beyond the sole question of socioeconomic homophily, the lack of discussion between the SNA literature and the sociology of social reproduction proves harmful to the respective objectives of both strands. On the one hand, the study of school students' relationships should not ignore the impact of socioeconomic factors, including their cultural and psychological dimension, which we know shapes human behavior in decisive ways across most aspects of social life. It would also benefit from a more holistic understanding of the local context, at school and in the neighborhood, and of its impact on adolescent sociability. Despite recent efforts in this direction (e.g. Kruse and Kroneberg 2019; McFarland et al. 2014), most SNA studies on adolescent friendship tend to adopt a formalist or universalist stance, trying to uncover generic regularities in the organization of social networks derived from fundamental psychological or relational mechanisms. While this is a legitimate objective, the ethnographic literature on student sociability has also shown that relational patterns and cultural expectations about friendship ties strongly vary across social environments. Integrating these contextual elements would enrich SNA models and theories about youth friendship networks. On the other hand, social reproduction studies may have dedicated too little attention to horizontal relations among peers, relative to the role of parents or teachers. With that said, there are still a number of studies that have looked into peer group dynamics, with fine-grained ethnographies of youth sociability at school or in the neighborhood. Such studies would greatly benefit from integrating the methods of SNA into their toolbox, as these offer the advantages of quantitative methods in terms of systematicity and reproducibility, while being well-adapted to ethnographic approaches and to the study of meso-level contexts (Lazega 2014).

The Present Study

The thesis investigates socioeconomic homophily among middle school students in a context of socioeconomic diversity. It is primarily based on research conducted in four French middle schools, all of which welcome students originating from the working-, middle- and upper-classes in similar proportions. One cohort of students has been followed for three years, through a mixed method design that combined questionnaire surveys and semi-structured interviews. Additionally, a second study will be used to a lesser extent, which was conducted during a three-week summer camp in the French Alps (with no relation to the main in-school sample). In a nutshell, the objective is two-fold: to document the aggregate levels of socioeconomic homophily observed in these different contexts, and to identify the relational processes that explain them.

The choice of middle school specifically is motivated by the fact that it is the time at which peer-to-peer sociability becomes central to children's social life (Adler and Adler 1998; Coleman 1961), and at which they progressively develop a "critical distance" (Cousin and Felouzis 2002) toward teachers and parents. Moreover, studying homophily in primary school would raise certain methodological problems, notably the fact that the notion of elective friendships may be less relevant to the sociability of young children (affinities among certain children do exist, but they are likely to be less stable, and declarative methods may be less reliable, which is why the study of young children often relies on observational methods). High school is generally a less diverse schooling level in France, due to the separation of general and vocational tracks, which operates on a strong selection on socioeconomic, ethnic and academic traits (Palheta 2015). Yet a certain level of socioeconomic diversity is a precondition for the study of socioeconomic homophily.

At this point, a clarification is in order. The concept of homophily can have different meanings depending on the authors. It is sometimes used as a descriptive concept, which refers to an observed outcome – there is homophily inasmuch as friendships among similar individuals are effectively more frequent; and sometimes as an explanatory one, which refers to an ongoing process – individuals look for friends similar to themselves. Quite tellingly, the adage traditionally used by researchers to illustrate the notion is itself ambiguous in that regard: when saying that "birds of a feather flock together", it is not entirely clear whether said birds are actively trying to flock toward their peers, or whether they effectively happen to be flocked together, neither of which logically implies the other. Following an established practice in the SNA literature (Goodreau, Kitts, and Morris 2009), I will use the noun "homophily" in the first sense exclusively, to designate a descriptive outcome about the state of friendship ties. The various processes or factors that may contribute to this homophily will be referred to using an adjectivized form: homophilic selection, homophilic behavior, homophilic preferences, etc.

In this descriptive sense, homophily is defined as an over-representation of ties among similar individuals within a given population, compared to what would be expected under a random allocation of ties. This means that homophily is always defined relative to a population of reference; socioeconomic homophily can be considered at the level of a classroom, a school, a neighborhood, or even an entire country. This is the key difference between the notions of *homophily* and of *homogeneity* (or of *diversity*, which I essentially use as an antonym to homogeneity). Highly segregated schools are characterized by extremely homogeneous peer groups, in the sense that most students have friends socially similar to themselves; but they will usually exhibit weak homophily, because, *within the population of the school*, differences in socioeconomic backgrounds do not impact relationships much. This is why, in homogeneous populations, homophily is practically meaningless as a concept. At the same time, however, segregated schools strongly contribute to socioeconomic homophily at regional and national levels, because they make friendships among similar students more likely in this broader, more diverse population.

With this in mind, it is possible to examine the many factors that may contribute to the apparition of socioeconomic homophily in friendship networks. In order to do this, the first step is to document the levels of socioeconomic homophily observed in the studied contexts. As trivial as this may sound, there is to the best of my knowledge, no descriptive estimate in the literature that would indicate what levels of socioeconomic homophily may be expected in French diverse schools.

> *RQ1: Given a situation of socioeconomic diversity, to what extent is there socioeconomic homophily in teenagers' friendships?*

Second, socioeconomic homophily need not be static, but can evolve over time, especially in the case of teenagers who are undergoing drastic developmental changes over a short period of time. There is also no reason to expect that the observed patterns will be similar across different contexts: the four schools of the sample, as well as the summer camp investigated in the auxiliary study, may exhibit different relational structures and levels of homophily.

RQ2: Does the strength of socioeconomic homophily vary over time and/or across contexts?

Third, adolescent sociability should not be considered as unidimensional. By this, I mean that it is made-up of different types of relationships that cannot be perfectly subsumed under a single notion like “friendship”. Students’ networks are made of close friends, but also of mere “pals” or acquaintances; they sometimes have enemies, or are victims to bullying; and their relationships at school often differ from those outside of it[7]. Therefore, any assessment of socioeconomic homophily should try, where possible, to differentiate between these different aspects of youth sociability.

RQ3: Does the strength of socioeconomic homophily depend on the type of relationship that is considered?

These three first questions are descriptive, in the sense that they do not attempt to provide an explanation for the observed patterns. For this, it is necessary to consider the factors that impact friendship formation among students, and that make ties more likely to appear or to persist between peers from similar socioeconomic backgrounds. Some of these factors may be forms of purposive selections of same-background friends, because of explicit preferences, or of prejudices associated to socioeconomic positions. They may also pertain to more subtle and unconscious dispositions that facilitate friendship formation among similar peers, such as shared tastes, interests, or a greater ease with communicating. Alternatively, homophilic selection may operate on characteristics that are correlated to students’ socioeconomic background, resulting in an indirect inducement of socioeconomic homophily through a compositional effect. Likely candidates in that regard are ethnicity and academic achievement. Moreover, even in diverse schools, the structure of contact opportunities among students often remains unbalanced: depending on their background, they can live in different neighborhoods, attend different classrooms, or participate in different out-of-school activities. Finally, one key insight of social network analysis – if not of sociology altogether – is the fact that inter-individual relationships are dependent from one another, thus the very notion of a social network. This means that students’ friendships are impacted by their position in the relational structure, and that the formation of future ties depends on existing ones: group pressure, friends of friends or enlarged social circles are major drivers of friendship

7 The case of romantic mating is also important but will not be investigated here. Not only do I lack data on this point, it also raises very different issues, such as the intersection of socioeconomic background and gender norms, or the criteria of physical attractiveness among middle school students.

formation, above and beyond the affinities that may exist between two individuals. Altogether then, this brings us to the fourth research question:

> *RQ4: What are the relational processes through which socioeconomic homophily appears among youths?*

Following the above discussion, this question may be further split into sub-questions, such as: (a) is there direct homophilic selection based on socioeconomic background? (b) Are there compositional effects that induce socioeconomic homophily due to selection based on correlated characteristics? (c) Are their differences in structural opportunities for contact among students that induce socioeconomic homophily? (d) How do group and network processes – understood as the dependency across network ties – affect socioeconomic homophily? This quick classification of homophily-inducing processes may seem arbitrary to the reader, but a more complete justification will be provided later (chapter 2).

Finally, these processes may themselves change across time and contexts, which naturally bring to the last research question:

> *RQ5: Do the relational processes identified in response to RQ4 vary over time and/or across schools, and can this account for the trends observed in response to RQ2?*

To conclude this introduction, a word of warning as per the way in which homophily should be conceptualized. Despite the fact that the notion stands at the core of the thesis, it is important to keep in mind that it is far from being the only relational feature that characterizes students' sociability. Crucially, even though homophily may be conceptually separated from these other features, a network eventually forms a single inter-dependent structure, whose different aspects are partly constitutive of one another. For example, the strength of homophily will depend on the extent to which the network is centralized or clustered[8]. Regarding socioeconomic origin in particular, students from certain backgrounds may be socially more active than others, having more friends altogether; as well as they may be more popular and hold a position of power and prestige within peer groups. Such elements need to be taken into account in order to interpret homophilic patterns correctly, which means that socioeconomic homophily should be considered within the broader relational context of which it is part. This has important implications in terms of quantitative network modeling, but also of substantial interpretation, as we shall see in the empirical chapters.

8 Centralization refers to the concentration of ties around a small number of popular individuals, while clustering refers to the tendency of ties to form closed triangles or larger groups. These are network analysis concepts that will be detailed in chapter 3.

Chapter Lay-Out

The thesis is made up of eight chapters, in addition to the introduction and conclusion. **Chapter 1** proposes a literature review on the themes of adolescent peer relations and socioeconomic origin. It presents some of the established results within this literature, as well as its blind spots and ongoing debates, which will help to refine the research questions presented in the introduction. **Chapter 2** presents a general theoretical framework for the study of socioeconomic homophily at school. It mostly draws on concepts taken from Social Network Analysis but proposes to re-frame them through the meta-language of human action proposed by Lahire (2005), which centers on the notions of context and dispositions. **Chapter 3** presents the fieldwork, data and methods. It introduces the four schools in which the study took place, presents important features of the followed cohort of students, and discusses the methodological choices made in the research design.

Moving on to the empirical chapters, **chapter 4** proposes a descriptive assessment of socioeconomic homophily in the studied schools. It uses simple quantitative measures to document the strength of socioeconomic homophily across different types of relations, assess its evolution over time, and compare it to other prominent forms of homophily (ethnic, academic and gender homophilies). **Chapter 5** then attempts to disentangle the relational processes that have generated the observed features, through the use of longitudinal network models called *Stochastic Actor-Oriented Models* (SAOM). Moreover, an original approach to the use of these models in the study of homophily is proposed, which I call *contribution scores*. It relies on simulations derived from empirically calibrated models to explore hypothetical scenarios, which permits us to infer the impact of specific relational processes on the formation of homophilic network structures. **Chapter 6** is based on qualitative methods. It relies on semi-structured interviews conducted with a sample of students and examines the mental and discursive categories that these students use to classify their peers and orient themselves in the relational space of the school. **Chapter 7** tries to combine the findings from chapters 4 to 6 in order to explain the differences in homophilic patterns observed across schools. Since there are only four schools in my sample, it faces what is often referred to as the “small-n problem”, making the chapter overall more exploratory than others. Hypotheses are proposed about the origins of inter-school differences, but these are meant to be tested more systematically on larger samples. Finally, **chapter 8** shifts the focus away from school and toward a vacation context (quite fitting, then, for the end of the thesis!). It examines socioeconomic homophily in a three-week summer camp and argues that this particular context is especially suitable to isolating the effect of attendees' internalized dispositions toward socioeconomically similar friends. It also introduces a methodological innovation, in the form of a longitudinal extension of a recently developed class of statistical models, called the *Exponential Random Partition Model* (ERPM). The chapter is adapted from an article co-authored with Marion Hoffman, the inventor of these models.

In the conclusion, I summarize the main findings from the thesis and evaluate them in regard to the initial research questions. Broader implications for the study of friendship formation and of socioeconomic homophily are discussed, together with the limitations of the proposed approach and leads for future research.

Chapter 1: A Literature Review on Adolescent Peer Relations and Socioeconomic Mixing

One cannot really identify a distinct scientific field centered around the question of adolescent socioeconomic mixing. Rather, separate research traditions, each built around their own research interests and articulated to a broader disciplinary field, have treated the issue as one piece, among others, of their puzzles. As a result, few studies explicitly focus on the impact of socioeconomic background on adolescent relationships, but many of them address neighboring questions. They pertain, for instance, to ethnic, gender and academic homophily; to generic principles of friendship formation within adolescent peer groups; or to social reproduction and class-dependent socialization, at school or in the family. In most cases, socioeconomic homophily is approached from the side, so to speak, and as a secondary concern.

In the first part of this review, I go through the main bibliographic clusters I identified, so as to give the reader an overview of the literature my thesis engages with. In each case, I briefly discuss the main conclusions of the reviewed works, as well as their relations (or lack thereof) with other clusters. The second section focuses on the contradictions or disagreements within this literature, and the ways in which they will be addressed in the dissertation. It also discusses how these works can inform the research on socioeconomic homophily, even though they mostly pertain to neighboring subjects. Therefore, while the first part tries to follow the existing bibliographic boundaries as closely as possible, the second, on the contrary, aims at drawing and combining insights from sometimes unrelated fields.

1. Research Traditions Studying Adolescent Social Mix

There are two main bibliographic clusters with which the thesis engages: Social Network Analysis studies in the school context, and ethnographic works centered on the question of schooling inequalities and social reproduction. While the first primarily looks into relational structures, raising the issue of socioeconomic background as one among many other factors to structure students' networks, the second, on the contrary, is crucially interested in social class, with peer-to-peer sociability being but one aspect of social class belonging and socialization. The interests of both traditions sometimes overlap, but the methodological and theoretical distance among them is large enough that they rarely engage with one another. Both traditions are diverse and call for further internal distinctions.

1.1 Social Network Analysis in the School Context

Social Network Analysis (SNA) is a very broad and dynamic field, which studies human relations using the statistical methods of network analysis (derived from graph theory in mathematics). Although the field is primarily defined by its statistical toolbox, it also has a number of shared concepts, theories and empirical results. Those are grounded in the belief that social networks exhibit recurrent properties that go beyond the cultural content of the relationships they are made of. In other words, part of the field adopts a formalist position, in the Simmelian sense, i.e. that it distinguishes between the “form” and “content” of social relations (Simmel 1950). This formalist approach states, for example, that most social networks exhibit the fundamental properties of homophily, balance and popularity (Block 2018; McFarland et al. 2014). This is actually a controversial position, and while many SNA researchers accept the idea that there is such a thing as a “network science” or “network sociology”, which can uncover recurrent network mechanisms, others criticize this formalist approach, asserting that networks are methodological constructs that should not be mistaken for social realities, and thus do not exhibit properties beyond that of the particular social relations of which they are made (Lazega 2014).

In any case, SNA as a field is deeply related to analyses of school networks. Pioneers of the network approach, most notably Coleman, worked on students’ peer groups (Coleman 1961). Even now, many key methodological advances in network analysis were first developed and tested on school data: inferential models that can solve the issue of observations’ dependency (Goodreau et al. 2009) longitudinal models (Snijders 2017) or multi-level analysis of several networks (Lazega and Snijders 2015). The reason is primarily practical: schools are one of the most convenient settings to collect complete-network data. Students are neatly ordered by classrooms, nested in schools; it is generally much easier to make them fill hour-long surveys about intimate matters than it is with adults (sitting at a table and checking boxes on a sheet being, after all, the daily student condition); friendship is a relatively straightforward indigenous concept, which makes it easy to define network ties; clear institutional boundaries separate the inside from the outside, sparing researchers the hardships of identifying relevant social circles and sampling rules; and, last but not least, networks are of a very reasonable size – never more than a few hundred nodes – compared to the large mass of data that many social networks represent. Due to this strong methodological orientation, and coherent with the formalist approach mentioned above, part of the field at least leans toward explanations in terms of generic psychological or relational mechanisms, which is an important difference with the second tradition I will come to discuss.

Altogether, this research tradition is both large and heterogeneous. Therefore, for the sake of clarity, I further divide it into 3 groups. Note, however, that there are a number of pathways, common interests and mutual citations across these.

1.1.1 Social Psychology and the Developmental Approach

The biggest chunk of literature on youth friendship networks comes not from sociology, but from social psychology. There is a long tradition of sociometric studies in psychology that has tried to model the impact of social ties on individual psychological outcomes, such as depression, self-confidence, aggressiveness, or norm-breaking behavior. While older studies relied on classical statistical methods (e.g. Coie, Dodge, and Coppotelli 1982), SNA is a precious toolbox for studying sociometric nominations and has now been widely adopted.

Of particular interest to sociologists are the numerous studies that tackle issues of status inequalities among students. This has been approached from both sides of the status distribution: popular or “cool” students, and bullied, isolated, or rejected ones. Popularity refers to the fact that certain individuals are more appreciated, listened to, respected, looked-up to or obeyed than most of their peers; in a word, they are high-status. Social psychologists have emphasized the dual nature of youth popularity. They distinguish between “acceptance” – how well-liked an individual is – and “social impact” – their visibility, prominence, and dominance in the group (Coie et al. 1982). Usually, both aspects are only weakly correlated (Cillessen and Mayeux 2004; Rose, Swenson, and Carlson 2004). Students with high acceptance would be more oriented toward school norms, as well as more prosocial and empathetic toward peers, while students with high impact would frequently exhibit aggressive behavior, be cited as bullies, and foster anti-school norms (de Bruyn, Cillessen, and Wissink 2010). As for bullying, it is generally associated with social impact and perceived as the consequence of power asymmetry and strategic decisions. Pellegrini and Long (Pellegrini and Long 2002) or Salmivalli (2010) see it as frequently being a means oriented toward specific ends, such as maintaining one’s standing in the group. This contrasts with older psychological perspectives on bullying, which exclusively saw it as the expression of inherently aggressive or disruptive personalities (Salmivalli 2010).

Apart from peer status, another important area of research pertains to homophily. Recent research has emphasized the fact that homophily can occur through two distinct mechanisms: “selection” refers to the idea that similar peers tend to bond, and “influence” the idea that friends socialize each other into becoming more similar, regarding a number of psychological and behavioral characteristics (Brechwald and Prinstein 2011). Both selection and influence have been found to operate regarding delinquent behavior (Snijders and Baerveldt 2003) or substance use (Dijkstra et al. 2015). Homophily also appears for statutory positions, with popular students more often being friends with one another (de Bruyn et al. 2010). Furthermore, popularity is correlated to engagement in drug consumption or delinquency (Dijkstra et al. 2009; Troop-Gordon, Visconti, and Kuntz 2011), although this should probably be related to social impact more than acceptance.

Most of these studies are rooted in a developmental perspective. They share an assumption that there is a fairly “normal” or “proper” development for children and adolescents and focus their attention on potential deviations from this desirable pattern. As such, they tend to have a pathogenic view on certain youth behavior (bullying, anti-school norms, drug consumption, sexual behavior – all of those being often coined “risk” or “problem” behavior). As Galitz and Robert (2014) point out, the literature on bullying in particular, has been increasingly framed through a Public Health approach. This approach is, for example, visible in the impact/acceptance theoretical pair: one type of popularity (acceptance) is seen by researchers as positive and desirable, when the other (impact) is associated with pretty much any “bad” behavior one can think of – again: bullying, drug consumption, anti-school behavior etc (Troop-Gordon et al. 2011 makes an extremely salient example of a binary and moralist view on youth popularity). As we shall see, this contrasts with studies from other traditions, which have a different take on school norms and desirable developmental patterns.

Moreover, this literature is relatively insensitive to variation among social groups. Classical sociological variables such as ethnicity or social class are absent from most of the studies I have cited, and there is also a lack of cross-national comparisons. Again, this is an important difference with other strands of literature.

1.1.2 Peer Status, Local Orders and Learning Conditions

The themes of peer status, homophily, and selection-influence processes have also been studied by sociologists. One of the founders of this tradition is Coleman (1960), who uses the term “adolescent subculture” to refer to the relative autonomy of adolescent peer groups in establishing their own norms, values and status hierarchies. Coleman explicitly regretted that these subcultures were generally not oriented toward school success but valued romantic or athletic abilities instead. In another work, he also showed that the socioeconomic composition of schools has an impact on students’ performance, net of individual socioeconomic backgrounds (Coleman [1966] 1995; note, however, that this is a result based on multilevel regression, not network analysis).

While unrelated to network analysis, the ethnographies of Eder (1985) and Merten (1997) are also important references for the field, being two fine-grained descriptions of popularity mechanisms among adolescent girls (middle school for Eder, junior high for Merten). They both emphasize the ambiguity of popularity, in that “popular” students are not necessarily well-liked: they can resort to aggression and meanness to enforce their status but stand at a risk of becoming rejected if they cannot maintain at least some level of likability among peers. Informed by these monographies, and essentially pursuing Coleman’s policy interest in promoting pro-school norms within adolescent peer groups, later sociological studies on adolescent school networks have been very concerned with the relationship between peer status, “problem behavior” and academic success. For instance, Rodkin et al. (2006) show that sociometric popularity is associated with aggressive behavior for a sample of sixth-graders. With a longitudinal design, Moody et al. (2011) examine the link between popularity and drugs in mid-adolescence – the former predicting increased consumption of the latter.

To my knowledge, there is no similar study explicitly dedicated to popularity in French schools. However, various ethnographic studies suggest that popularity hierarchy may be less salient than in the US context. In particular, ethnographies in US schools (together with numerous TV shows) commonly identify a well-bounded group of “popular” students, or a “leading crowd”, which is known to all students and is explicitly referred to as such. By contrast, in most descriptions of French schools, there does not appear to be such a clear group of popular students, though this does not mean that there are not more subtle or hidden forms of hierarchy (e.g. Cousin and Felouzis 2002; van Zanten 2012).

Researchers are also strongly interested in the effect of friends on academic success. It is a consistent result that there is strong homophily among students based on academic achievement (grades) and engagement (attitudes toward school) (Flashman 2012; Lomi et al. 2011; Tuma and Hallinan 1979). Recent studies try to determine the extent to which this is due to selection and influence respectively. Burk, Kerr and Stattin's longitudinal study (2008) indicates that peers not only select their friends based on similar levels of school engagement and delinquent behavior, but that there is also an effect of influence among friends: they socialize one another into participating in acts of delinquency, with further effects on school engagement. Similarly, Lomi et al. (2011) find evidence of both mechanisms. However, Fortuin, Geel and Vedder (2016) only find evidence for selection, not influence; as do Smirnov and Thurner (2017). A natural experiment conducted in a boarding school for high-achieving students in Peru (Zárate, Angrist, and Pathak 2019) even found processes of negative influence: middle- or low-performing students that have high-performing friends experienced a deterioration of their grades over time, possibly because of discouragement and negative comparison with these friends.

Therefore, while selection based on grades is a well-established process, it remains uncertain whether there is a substantial influence effect as well. The contradicting results found in the literature suggest that this may depend on environmental factors impacting the salience and direction of peer influence, such as different national and school cultures, but, to the best of my knowledge, this has not yet been explored. Moreover, Lewis, Gonzalez, and Kaufman (2012), applying selection-influence model to the Facebook networks of college students, found clear evidence of selection in terms of cultural tastes (music and movies), but not of influence. Though there are are reasons to believe that peer influence is particularly strong for middle and high school students (Adler and Adler 1998), whereas Lewis et al. consider slightly older college students, this still supports the idea that selection is overall more salient than influence (assuming, of course, that the social processes underlying these phenomena is somehow comparable for academic attitudes and leisure tastes).

A different piece of literature investigates classroom dynamics to explain educational resistance. Rather than students' individual attributes (such as ethnicity and social class) or mutual influence through dyadic ties (as with peer effect on grades), educational outcomes are primarily envisioned as stemming from the organizational context: certain classroom environments are more appropriate for learning than others. Relational dynamics, more than fixed subcultures, tend to produce a higher-level context that has an impact on students' behavior and performance. McFarland (2014) argues that student defiance toward teachers, and resistance to learning, appears when some students have enough social support from their peers to be backed up in their challenging of adult authority. He postulates that most individuals have the desire to bypass school norms (or any form of external control more generally); but that only popular ones have the resources to act on it. This is a way of reversing the popularity-school achievement relationship: to McFarland, it is not that resistant behavior begets popularity, but rather that popularity allows for resistant behavior. There is a small cluster of works that focus on such meso-level interactions within the classroom and their impact on the educational climate (Geven, O. Jonsson, and van Tubergen 2017; Müller et al. 2016).

1.1.3 Social Integration and Equality of Opportunities

The final group of SNA studies I will discuss, unlike the two previous ones, frontally tackle the fact that students belong to different social groups. While these studies share with the two previous strands a strong concern for academic performance, here the focus is not so much on the mechanisms through which these performances could be fostered for any and all students (e.g. by manipulating students' relations in order to break patterns of resistance to the teachers, as McFarland 2001 proposed). Rather, they look at the unequal opportunities for school success met by different social groups. This concern is extended not just to school success, but to social integration within peer groups: the question is both whether certain students tend to be marginalized or excluded (e.g. students from ethnic minorities) and whether students manage to establish inter-group contact.

The primary characteristic for which such questions have been raised is doubtlessly ethnicity or race[9]. In the US context, with a strong public interest for racial desegregation policies, there is a long tradition of network studies on racial homophily in children and adolescent friendships. Unsurprisingly, friendships between students belonging to the same racial group are more likely than between students of different racial groups, even in formally desegregated schools (Hallinan and Williams 1987; Moody 2001; Shrum, Cheek, and Hunter 1988). Researchers have also considered the case of multiracial adolescents, who tend to either amalgamate into a single-race group (e.g. mixed Native American-White students tend to mostly befriend Whites) or act as relational bridges between racial groups (e.g. Black-Whites are somehow intermediate between Blacks and Whites in terms of their friendship choices) depending on the considered mixed-race group (Doyle and Kao 2007). In terms of

9 In general, "race" is used in US studies and "ethnicity" in European ones. "Race" is commonly operationalized through self-identification, while "ethnicity" can be measured either through self-identification, or (more frequently) migratory background (being born, or having parents born in a foreign country). This is linked to the particular situation of minority groups in both continents (Afro-Americans are a native population in the US, while an important part of non-white European residents come from relatively recent immigration), as well as to specific research traditions and national cultures. See Safi (2013) for a discussion of these concepts and their history.

time trend, racial homophily increases as students grow older, peaking in middle school and then remaining largely stable (Shrum et al. 1988).

However, this is not only due to discrimination or same-ethnic relational preferences. First, of course, towns, neighborhoods and schools tend to be racially segregated, meaning that the probability of contact is higher among children belonging to the same racial groups, fostering homophilic friendships even in the absence of any discriminatory choices on the part of individuals. Mouw and Entwisle (2006) estimate at about one third the part of racial friendship segregation among American adolescents that is due to residential segregation. Most of this effect is due to between-school segregation (students attending different schools), but there is also a smaller within-school effect (in a given school, people tend to befriend peers living close to them).

Then, levels of racial homophily also depend on the racial composition of schools, *even when homophily is a net measurement of this racial composition*. Indeed, homophily is generally quantified through relative probabilities of tie occurrence (odd-ratios of same-group vs cross-group ties), which are not sensitive to the differences in opportunities for cross-group friendships mechanically induced by the population composition. Using such indicators, Moody (2001) found that the average level of racial homophily increases as school composition becomes more racially mixed; it peaks in schools with moderate levels of mixing but decreases again when levels of mixing become very high.

This somehow counter-intuitive result – more compositional mixing can result in more relational segregation – could be explained by two concurrent phenomena. First, the overall level of mixing in the population might impact the salience of race as a distinction criterion. If but a few students are from a different group than the rest of their peers, then race as a whole might look unimportant to students. On the other hand, the presence of several racial groups of significant size may elicit tensions or rivalries among them, making race a salient characteristic within local subcultures. This is coherent with Hallinan and Williams' (1989) findings that white students exhibit less racial homophily than black ones, but only in predominantly white schools; in predominantly black schools, white individuals appear more homophilic. Thus, racial identities might be stronger when individuals feel threatened or in a minority. Second, when few cross-race ties are available, minority students might be forced to reach out to other racial groups to find friends with whom they match on other important characteristics. Say that there are only two black students in a classroom: if one of them is a football fan, but the other is not, then the first would need to befriend white students if he wants someone to talk about football with. In other words, homophily on a given characteristic is stronger when individuals have a large variety of friendship options to choose from, as Blau (1994) theorized. McFarland et al. (2014) formalized these two ideas in a model predicting greater racial and/or socioeconomic homophily in schools as both the size (greater choice available) and heterogeneity (greater uncertainty and potential tensions) of the student population increase. Testing it on American data, they found support for their hypotheses. Remember though, Moody (2001) also found that with high levels of mixing in schools, relational homophily decreases again – a curved pattern that is not easy for these theories to account for.

Finally, even in formally desegregated environments, a number of constraints and opportunities can foster same-race friendships. Tuma and Hallinan (1979) showed that racial homophily appears significantly smaller when taking into account structural features of the school context. Indeed, students from the same racial groups tend to have similar grades, as minority students usually do worse at school – and, as we saw in the previous section, students often select friends with similar grades. Furthermore, even within a given school, students are frequently sorted into different classrooms and working groups, depending on the options they take or on their academic level, all of which correlate to racial background. Similarly, Moody (2001) finds that racial homophily is lower in schools where extra-curricular activities are racially integrated.

Moving on to Europe, we also find several studies concerned with ethnic homophily. Again, the greater likelihood of same-ethnic friendships occurring is clearly established (Baerveldt et al. 2004; Geven et al. 2017; Kruse et al. 2016). The mediating effect of third factors, however, is less clear. Examining high-school friendship networks in the Netherlands, UK, Sweden and Germany, Kruse et al. (2016) finds no evidence that residential segregation explains ethnic homophily within a given school. Similarly, Smith, Maas, and van Tubergen (2014) put to test the "by-product hypothesis", that is, the idea that ethnic homophily is mediated by socioeconomic and cultural similarity. They find no evidence that any of these factors explain ethnic homophily. However, Stark and Flache (2012) did find, for Dutch middle school students, that similarity in tastes and opinions (music, school attitude and "deviant" behaviors) mediated ethnic homophily; but only in classrooms where ethnicity and opinions correlated (which seems logical). In certain classrooms, where ethnicity and opinions did *not* correlate, ethnic homophily was not fostered, but on the contrary was reduced by opinion homophily. This heterogeneity across classrooms in the opinion/ethnicity correlation may explain why Smith et al. could not detect any effect at the aggregate level. Altogether, the evidence for mediating factors of ethnic homophily is not clear, and sheer discrimination may play an important role. Comparing these results with the US studies mentioned above, where the mediating impact of third factors was much clearer, the question remains open as to whether the difference is due to actual national differences, or to the methodological designs of the studies[10].

As for peer status and social integration, Plenty and Jonsson (2017), working on Swedish data, found that students with an immigrant background were at a higher risk of being socially isolated, rejected by peers or bullied. However, this relationship is mediated by classroom composition: the higher the density of students with an immigrant background is, the lower victimization and exclusion is for these students, while conversely, the higher it is for native students. In other words, peer hierarchies would not be linked to intrinsic levels of status attached to different ethnic origins, but rather to local patterns of majority and conformity. In a similar vein, Geven et al. (2017) found that students belonging to ethnic minorities exhibit less "problem behavior" (skipping class or arguing with teachers) when they attend schools with a high proportion of co-ethnic peers. This relation is mediated by the number of in-school friends: when a significant portion of classmates are of the same ethnic group, minority students have more in-school friends, which in turn is associated with a reduction in problem behavior[11]. Finally, Kruse and Kroneberg (2019) argue that ethnic homophily depends not only on the composition of schools, but also on the level of inter-school ethnic stratification at a regional level. Based on German data for secondary school students, they found that in areas with high inter-school stratification – i.e. where ethnic minorities are strongly under-represented in the most prestigious schools and tracks – immigrants tend to strongly identify as Germans in the prestigious and selective tracks, but weakly in vocational ones. In turn, this impacts their propensity to reach out to out-group friends (immigrants who identify as Germans are more likely to nominate native peers). In areas with low stratification, however, there is hardly any difference between prestigious and relegated tracks in minority students' propensities to identify as Germans, and there is a

10 Most of the European studies mentioned here use the same data, drawn from the CILS4EU project (Kalter et al. 2016). As for the American ones, many of them rely on the Add Health sample, although some researchers also collected data of their own. The fact that a few large studies provide data for a significant part of the research being conducted means that their methodological choices and the idiosyncrasies of their design can have a disproportionate impact on the field.

11 Note that this relation may also be confounded by ethnic biases on the part of teachers, which is not discussed in the article (e.g. teachers in schools with few ethnic minorities may have a higher chance of having prejudices against these minorities, which may facilitate conflict with certain students).

weaker relationship between German identification and the propensity to nominate out-group friends (controlling in all cases for the ethnic composition of schools)[12].

There are fewer studies that look at socioeconomic homophily. In most cases, it is simply used as a control variable, researchers being primarily interested in race/ethnicity. Going through the result tables of these articles, the statistical effects associated to socioeconomic homophily generally appear to be relatively small – much smaller, at least, than those of ethnic homophily. Smith et al. (2014) found no socioeconomic homophily at all while controlling for ethnicity, gender and sociometric popularity (German, Dutch and Swedish high-school students). In Kruse et al.'s models (2016), socioeconomic homophily, although statistically significant, has a much smaller effect size than ethnic homophily (German and Dutch high-school students). Similarly, McFarland et al. (2014), using the Add Health data set, found consistent socioeconomic homophily across US secondary schools, but of a magnitude largely inferior to that of racial homophily. Altogether, this seems to suggest that socioeconomic homophily among adolescents is generally weak. However, there are reasons to question this result, to which I come back to below (section 2.1.1).

In fact, the SNA studies that directly address the question of socioeconomic origin are more interested in social integration and exclusion than in homophily. Olsson (2007) finds that youth poverty increases the risk of social exclusion at school. Surveying Swedish adolescents, he shows that economic deprivation reduces the number of friends, both looking at parental income and at adolescents' own economic resources (e.g. pocket money or spending power, which are correlated to parental income but also depend on the relative share of resources that the family is willing to dedicate to the child)[13]. Hjalmarsson and Mood (2015) and Plenty and Mood (2016) reach similar conclusions. Similarly, Malacarne (2017) has argued that socioeconomic background impacts friendship formation through status hierarchies more than homophily: he found that students with a high socioeconomic status (SES) receive more friendship nominations in US secondary schools, but that the share of

12 I am personally skeptical about the methodology of this article because the comparison is based on a relatively large geographic area. For this reason, it seems difficult to argue that the relevant differences between areas with high or low inter-school stratification pertain *only*, precisely, to this stratification. For example, areas with high inter-school stratification may have higher population densities, different employment rates, predominantly urban or rural residential structures, etc., all of which may confound the relationship between regional stratification and within-school homophily. The article offers no reason as to why certain areas are highly stratified and others are not, which makes it difficult to anticipate what hidden correlated factors could be at work. Nevertheless, apart from these issues of causal identification, the article still shows that the strength of ethnic homophily, as well as its relation to individuals' ethnic identification, strongly varies across contexts, which is an important result in and of itself.

13 There may be an issue of reversed causation here: students with a more active sociability may be more inclined to ask their parents for money in order to go out with their friends.

same-SES peers in the school is unrelated to the number of friends. His results, however, do not rely on SNA methods, but on regression models that predict the total number of friendship nominations received by individuals, which makes for an indirect measure of homophily at best (the models do not consider whether friendship ties are emitted by same- or different-SES peers within each school). In an ethnographic study conducted among 8- to 12-year-old US students, Adler and Adler (1998) also found richer individuals to have a higher status on average, particularly among girls.

In these studies, socioeconomic background is essentially conceptualized through economic resources and consumption power, with a focus on relative deprivation: poorer youths would not be able to participate in the same social activities as their friends because they do not have the money to do so (e.g. going to the movies, shopping or wearing fashionable clothes). This is very different from the usual conceptualization of ethnic homophily, where students are seen as belonging to discrete groups, each potentially having specific cultural or behavioral traits, with theoretical inspirations drawn for example from intergroup contact theory (Allport 1954). As we shall see below, this also differs from how socioeconomic background is thought of in other research traditions; it is an empirical question as to which conceptualization better captures its impact on students' relationships.

1.2 Sociability and Social Reproduction

The second strand of literature that I will discuss is rooted in the sociology of education, as well as of social stratification and inequalities. Unlike studies of the previous strand, which despite their diversity, can be grouped under the common label of Social Network Analysis, these studies lack a clear set of methods or research questions that would neatly delimit a research tradition. I would argue, however, that they share a key interest in social reproduction, that is, the perpetuation of macro-social hierarchies from one generation to the next. Therefore, and although it does not do justice to the number of other puzzles that these studies address, I will refer to them as "social reproduction studies", for lack of a better name.

There is a considerable literature on social reproduction in the sociology of education, but not all of it looks at student peer groups. In fact, family socialization, on the one hand, and the school institution, on the other, have attracted the most attention. Therefore, youth sociability has mostly been addressed in as much as (a) it relates to family or school socialization and (b) it contributes to social class, ethnic or gender inequalities.

Note that by this definition, the SNA studies on ethnic homophily that were the focus of section 1.1.3 could fall within the scope of the present section as well. It is not always easy to pinpoint what exactly separates two groups of research; but I feel confident in saying that social reproduction and SNA studies, although they both tackle schooling inequalities and class/ethnic divides, belong to almost unrelated traditions. One important dimension of this split is the relation to so-called critical theories of education, that is, theoretical perspectives that see school primarily as an institution of cultural arbitrariness and social reproduction (e.g. Bernstein [1977] 2003; Bourdieu and Passeron [1970] 1990; Willis 1977; Young 1972). According to Duru-Bellat (2015), these views have been very influential in the sociology of education and inequality, especially in France and the UK since the 1970s'; and while most contemporary researchers are less radical than, say, Bourdieu and Passeron or Willis, the critical framework remains influential overall – it is part of the field's intellectual background. This is however, not the case for the aforementioned SNA studies, as illustrated by their policy recommendations and overall adhesion to academic criteria (resisting teachers is a "problem behavior", academic performance is desirable in and of itself, etc.)[14].

Altogether then, these are studies of family and schooling inequalities and their relation to youth sociability that are strongly marked by critical theories of social reproduction. In turn I will discuss studies centered around the role of the family and the school system.

1.2.1 Working-Class Culture, School Segregation and Parental Socialization

Many ethnographic studies agree on the role of youths' own sociability in the formation of working-class identities. Willis's *Learning to Labor* (1977), a classic of the field, describes the formation of socially homogeneous peer groups among British working-class adolescents. Mixing the working-class values found in their family with the youth culture of

14 Presented as such, it seems as if research strands are over-determined by political concerns (pro-school SNA studies vs critical works denouncing school as an institution of social reproduction). While this is not necessarily true for all individual studies, I do believe that at the level of scientific traditions as a whole, there are indeed political and moral positions underpinning certain research choices. This does not imply a pure reductionism of scientific production to political opinions though: there are a number of other reasons that can contribute to scientific fields growing in isolation and developing different methods and theories. These include national or regional traditions (SNA studies are very much present in American, Swedish and Dutch literature, while critical traditions are mostly to be found in France and the UK), methodological constraints (network analysis is a demanding statistical method that often calls for heavy specialization) or simply different properties of the researched objects (as I said, SNA researchers are deeply interested in formal network mechanisms, while critical theorists have a more instrumental approach to sociability, their primary focus being social reproduction; thus it makes sense that the former would call for methodological sophistication in network analysis, and the latter for a deeper theoretical view on the school institution and its role).

the 1970s (built around rock music in particular), they develop what he calls a "culture of resistance" that contributes to drive them away from school, eventually reproducing their socioeconomic condition. Within these peer groups, social pressure and mutual socialization keep individuals in line with the group's norms. Thus, for Willis, though occupational background and family socialization remain the primary cause of children's social dispositions, it is the interaction among similar peers that reinforces these dispositions and crystallizes them into scholarly deviant youth subcultures.

These mechanisms take place in a strongly segregated working-class environment. As such, socioeconomic and/or ethnic segregation at the town or school level is central to explaining the social homogeneity of student peer groups. Its most basic, yet powerful component is spatial segregation: places of residence strongly correlate to social class. As for school specifically, middle- and upper-class families often carefully choose where they live in order to avoid dominantly working-class schools[15]. They are also able to exploit administrative opportunities to bypass public schools' sectors, or more simply resort to private schooling (Oberti et al. 2012; van Zanten 2015).

Furthermore, even when they do attend diverse schools, families can attempt to control the relations of their child. A straightforward option consists in directly monitoring their friendships, ranging from subtle incentives to avoid certain peers ("I don't want you to keep bad company", "do not hang around with them"...) to sheer forbiddance. What Coleman called *inter-generational closure* (Coleman and Hoffer 1987), that is, for parents to know the parents of their child's friends, can also be used to exert control over relations (e.g. the child is only allowed to go to a friend's place if parents first agree upon it between themselves). Van Zanten (2015) talks of a "surrounding strategy" of upper-class families, particularly those who work in the private sector and/or whose capital is predominantly economic rather than cultural, who would tend to keep tight control over their child's social environment.

However, based on interviews with parents, van Zanten argues that this surrounding strategy is not applied to the same extent in families whose capital is primarily cultural and/or that work in the public or associative sectors. Indeed, these families typically adhere to left-wing values of diversity and social mixing. Yet if mixing is valued on a theoretical level, in practice members of lower-classes or ethnic minorities are still perceived as threatening.

15 In France, students that go to public schools (~80% of the population) have to attend the school of their district, although there are certain exemptions to this rule (for clarity, I use the term "public school" to refer to schools funded and ran by the State, as opposed to private schools). Regarding the choice of school, I only cite literature on the French case here, because institutional rules can make things very different across countries.

Consequently, these families struggle to find a balance between their ideological beliefs and concerns for their children's education (Ball, Vincent, and Kemp 2004; Brantlinger, Majd-Jabbari, and Guskin 1996; van Zanten 2003)[16]. Furthermore, they fit Lareau's ideal type of the "concerted cultivation", where parents engage in argumentation and negotiation with the child and try to avoid frontal coercion (Lareau 2003). For both these reasons, they are much more reluctant to directly engage in monitoring their child's friendships.

These families have, however, more subtle ways – conscious or not – to maintain some control over their children's sociability. According to Lareau, in the model of concerted cultivation, parents constantly try to orient their child toward educationally profitable activities, through incentives, affective proximity and "soft" control of their desires and interests. They also dedicate a lot of their own time to shared activities with their child. As a result, upper-class youth tend to have very tight schedules, doing many extra-curricular activities and moving in-between several environments and peer groups (the friends of the tennis club, of the music school, those met during summer camps, etc). Through time constraints and a fragmentation of friendship circles, the possibility for horizontal socialization among children is greatly reduced. This contrasts with the second educational model of Lareau's typology, "natural growth": according to her, working-class parents tend to value child autonomy and not to meddle with their social life. Thus, it might be the case that peer-to-peer socialization is stronger relative to direct parental socialization among children and adolescents in working-class families.

Of course, parents do not only impact children's sociability through external constraints: the transmission of internalized dispositions also plays a key role. Many of the individual characteristics out of which students will perform homophilic sorting when choosing their friends might be impacted by their education and the family context they grew up in (e.g. musical tastes, moral values, ways of speaking, favorite games etc). This may happen in an involuntary fashion, as children mimic their parents more or less spontaneously – which should concern all socioeconomic groups –, but it can also be the product of conscious and directed educational efforts – this time with socioeconomic variation in the willingness and/or ability to conduct such effort. Coming back to the educational model of the "concerted cultivation" that is prevalent among the upper-classes (particularly those with high cultural capital), the emotional proximity between parents and child favor the internalization

16 Note, though, that researchers see this balance as clearly tilted; for them, in most cases, the concern for social distinctiveness will prevail, and exposure to other social groups will be kept to moderate and unthreatening levels.

by the child of the expectations, values and interests of their parents. In turn, this leads the child to behave according to their parents' desires (to an extent at least), notably in the choice of friends. Van Zanten (2015) talks of a "cooptation" strategy, which she contrasts with the "surrounding strategy" I mentioned above.

With that said, the extent to which dispositional transmission within the family impacts the social life of children is debated among researchers. A strand of thought known as the childhood studies (by analogy with cultural or gender studies) emerged in the 1970s as a reaction to dominant paradigms in the social sciences which saw children as essentially white sheets waiting to be socialized (such as Durkheim and Parsons in sociology, or Piaget in psychology). On the contrary, childhood scholars emphasize children's agency, their capacity to act and choose as autonomous social actors (Garnier 2015; James and Prout 1997). Children are seen as having a culture of their own; they are active in forming or dissolving relationships, following a logic that cannot be reduced to the influence of adults (Delalande 2003). In a similar vein (and though not directly part of childhood studies), Pasquier (2005) has argued that recent generations of high school students were characterized by a high level of autonomy from adults, a rejection of "classical" cultural hierarchies (i.e. highbrow vs lowbrow culture), the dominant influence of peer-to-peer socialization, and a form of individualized eclecticism in terms of their cultural tastes. According to Pasquier, generational differences are now more important than socioeconomic differences in explaining teenagers' cultural practices.

However, these positions have elicited some critics. Boyer (2006) contests Pasquier's conclusions that generational differences now prevail over socioeconomic ones; upon closer inquiry, she argues, there are still sharp socioeconomic divides among adolescents in terms of their cultural practices[17]. As for childhood studies, Lignier and Pagis (2014) argue that their focus on children's agency have led them to neglect the role of adult mental categories in children's social activity. For Lignier and Pagis, while children are indeed actively constructing their social relations and, more generally, making sense of the world around them, they do so by borrowing moral and aesthetic categories, tastes and distastes, from their adult-directed environment. In their study, they ask primary-school students to talk about the other pupils they dislike. They found that academic motives are often advanced to justify antagonisms (of the type "I don't like X because he is stupid, he does not know how to answer in class"), as well as what they call "domestic schemes": children refer to the values

17 Boyer (*ibid*) also suggests that the strong theoretical claims put forward by Pasquier are not sufficiently supported by the empirical evidence she provides – something that I tend to agree with.

emphasized by their education to denigrate others. This is the case of hygiene ("he is dirty") and of certain socioeconomic or racial prejudices that the authors interpret as coming from domestic socialization (without direct proof, though). They conclude that children mobilize schemes acquired from teachers and parents when determining who they like or not[18]. This idea that adult categories strongly shape children's perceptions is supported by a series of other studies which show that from a young age, children exhibit a "social sense", defined as a capacity to rank jobs and to perceive – sometimes confusedly – the social status attached to different occupational positions (Lignier and Pagis 2012; Zarca 1999; in these studies, children are asked by researchers to rank occupations from the least preferred to the most, and to discuss their choices).

Yet Lignier and Pagis's result is itself contradicted by other studies. In research on gentrified neighborhoods in Paris, London and San Francisco, Lehman-Frisch, Authier, and Dufaux (2012) found that in public primary schools with socioeconomically diverse classrooms, children commonly pair with friends from a different background (although they did not provide quantitative estimates similar to those of SNA studies). In this case, geographic propinquity – being in the same classroom – seems to result in effective friendship ties, which would suggest that internalized traits and mental categories are not (yet?) an obstacle to friendship formation among dissimilar children. By extension, this would point to spatial segregation as being the main driving factor of socioeconomic homophily when considering the entire population of children (i.e. including those that do not live in highly mixed contexts). Note that this does not mean that there is no homophily whatsoever in children's friendships: Lehman-Frisch et al. also found that, outside of school, sociability contexts are much more segregated, with children mostly coming into contact with same-background peers (primarily because of extra-curricular activities for middle- and upper-class students, and extended family networks for working-class ones). As a result, out-of-school friendships are much more homophilic that in-school ones.

Altogether then, it remains unclear to what extent children use adult categories of judgment related to macro-social positions – ethnicity and class – to organize their relations. This may also strongly depend on age: it is generally considered that children are much more

18 I find the methodology of this study to be problematic, in that the authors do not propose a systematic examination of the different schemes used by students, nor do they try to assess their overall frequency; mostly, they reason from a few selected examples. As a result, while they show that children *sometimes* resort to mental categories imported from adult discourse, it is hard to know how important these are relative to categories that are endogenously produced within kids' peer groups (or that are imported from other sources, such as mass media or the internet).

dependent on their parents than adolescents, both materially (e.g. parents do not let them go out alone) and symbolically (they believe in and care for what their parents tell them). Octobre et al. (2010) mention that one key feature of adolescence is the shift from parents to peers as the main reference group in terms of values, opinions and tastes – the same idea can also be found in Adler and Adler (1998). Yet it could also be argued that class-specific psychological traits increasingly develop with age, which means that young children may have an easier time crossing socioeconomic boundaries in their friendship choices. In fact, there is simply a lack of empirical evidence supporting either position.

1.2.2 Peer Hierarchies and Academic Status

Not all working-class children fit the ideal type of the rebellious "lads" studied by Willis. Ethnographic studies have also looked at how a minority of students from working-class backgrounds endorse school norms and actively try to break apart from their non-school oriented peers, with the desire of upward social mobility. For these students, sociability is organized by and around the school: they choose friends with similar academic interests and most importantly, restrict their sociability circles to classmates, contrary to scholarly deviant students who often develop ties outside of school, among neighbors and older peers. Whyte ([1943] 1993) already distinguished "college boys" from "corner boys" in an American slum of the 1930s'. Willmott (1966) in the 1960s' London or van Zanten (2012) in contemporary French suburbs found similar patterns. For academically oriented working-class students, school is essential in shaping their sociability, and more generally their behavior – they are after all, attempting to break away from their native social environment. Things are less clear for scholarly deviant youths. As van Zanten (2000) discusses, two positions exist in the literature. Some scholars consider that the attitudes and values of rebellious students primarily come from the family and from peer groups formed in the neighborhood. In other words, street sociability matters more than school socialization. This is notably the position of Willis, discussed above. However, for others, the academic ranking performed by the institution plays a key role in activating these street-oriented dispositions: it is because they feel ill-adapted and rejected at school that students turn toward other sociability circles. Cohen (1955) argues that working-class students first encounter the material and moral ideals of the middle-class at school; yet they are not able to reach these ideals by conventional means (i.e. school success), which pushes them to find illegal ways to gain money in order to reach middle-class living standards. Other research has also described the polarization of students between "pro-" and "anti-school" groups as directly resulting from academic rating: initial grade differences serve as a basis for individual differentiation and friendship selection, resulting in segregated peer groups which further aggravate the gap in academic performance (Ball 1981; Hargreaves and Hargreaves 2006; Lacey 1971).

We find a similar opposition in the literature on the formation of ethnic identities. While some insist on out-of-school experiences of discrimination and on street sociability, others stress the role of teachers who, by punishing or reacting to minor breaches of the school order, tend to harden, so to speak, the ethnic identification and opposition to school of minority students (see van Zanten 2000 for a review). An interesting position is also that of Bonnéry (2006): for him, there is a perverse effect to the positive promotion of ethnic diversity that schools sometimes perform. Indeed, and precisely as a reaction to perceived problems of racism and ethnic integration, teachers in schools with a high proportion of immigrants frequently organize events related to the culture of the countries of origin of their students (e.g. traditional meals at the school party, learning flags and country names in the language classes, literature texts about tolerance and diversity...). In primary school, these strategies are quite efficient in raising children’s interest and positive feelings toward the class. However, as they grow up, some of them start to lag and see their results collapse (for more generic reasons linked to their social background). At this point, ethnic identification categories have been made available to them, and they interpret their academic failure through these categories – calling the teachers racists, distancing themselves from well-performing white students, etc. Strikingly, Bonnéry describes the case of French schools; even though the French system is known for its universalist moral principles, which in principle preclude the recognition of children’s ethnic or religious differences. In British or American schools, where ethnicity is more freely recognized and discussed (Raveaud 2003), the backlash of positive ethnic labeling that Bonnéry theorizes might be even stronger.

These studies resonate with the theme of the mediating factors of homophily in SNA studies. Remember that Tuma and Hallinan (1979) found that racial segregation was stronger in classrooms that resorted to ability grouping, and Moody (2001) in schools where extra-curricular activities were strongly segregated by race. A similar mechanism might explain these patterns: the more ethnic categories are made apparent by features of the context, the more students use them to sort out their peers and select their friends.

In any case, scholars tend to agree that the dominant force exerted by schools on students’ sociability is one of divergence: “good” and “bad” students are pulled further apart by academic labeling. What they do not agree on is how school rankings shape the relational structure along its vertical dimension. Remember that SNA studies essentially depicted a situation in which academic performance was depreciated in students' subcultures: psychologists associate dominance and aggression with deviance regarding adult norms, Coleman (1961) deplored the athletic rather than academic orientation of youth groups and McFarland (2001), reversing the causal relation, argued that popular students were more likely to defy teachers. In the ethnographic literature, van Zanten (2012) similarly found that, in a predominantly working-class school, "good" students remained somewhat reserved so as not to attract jealousy from their peers, who could sometimes mock or harass them.

However, in other studies, we find a fully reversed pattern, academic status being now positively associated with peer status. Delalande (2005) and Gayet et al. (2003) argue that socially isolated students frequently come from lower socioeconomic backgrounds and have difficulties at school. Lignier and Pagis (2014), as we saw, also consider that grades are an important status criterion for children; according to them, even pupils who do not perform well in school acknowledge the legitimacy of this ranking, and recognize their peers as more "clever" than they are. Furthermore, Palheta (2012) conducted interviews with students in vocational high schools, who passed through special classrooms for low-performing students when they were in middle school[19]. When recalling this period, students frequently describe teasing episodes from "regular" students, who would for example call them "stupid" or "retarded".

Therefore, the case for the academic performance/peer status relation is not a clear one. Once again, patterns may strongly vary with age. According to Cousin and Felouzis (2002), middle school students progressively develop a "critical distance" from teachers and school rules –e.g., they observed more frequent mockery or criticism toward teachers in older grades. As for high school, both Dubet and Martuccelli (2014) and Rayou (1998) observed that students denied – at least formally and in front of adults – that grades were a legitimate means of differentiation. Openly mocking someone for their bad grades would thus be seen as illegitimate behavior within peer groups, and increasingly so as students grow older.

19 In France, all students are supposed to attend the same track in middle school, as the split between general and vocational training officially begins in high school. However, it is increasingly the case that middle schools create special tracks for students who are not adapted to regular schooling, which effectively act as pre-vocational schooling. The main such track is called SEGPA: it was historically designed for students with intellectual disabilities, but is now used more broadly for students with severe school difficulties. Working-class students are disproportionately represented in these programs (Palheta 2015).

2. Tensions in the Literature and Directions for Research

As we have seen across this review, social mixing and adolescent sociability have been approached from various angles. The two main strands of research discussed converge on a number of results, but they also reach opposed conclusions regarding certain questions. More importantly, a key limitation of the field as a whole is that SNA and social reproduction studies essentially stand parallel to one another, with very little discussion across them, such that their disagreements remain empirically unaddressed. Moreover, the methodological gap between them is large enough that certain diverging conclusions are simply difficult to compare, because they rely on different types of evidence. Finally, certain research questions fall outside of both fields' primary interests, in which case the issue is not that conclusions differ, but simply that the question has not been frontally tackled.

In the coming sections, I will try to pinpoint the research questions that suffer from this gap in the literature. For clarity, they are classified into two large groups: those pertaining to homophily (2.1.) and those touching upon conflict, status and hierarchy among students (2.2.). These two themes will not, however, be equally covered in the rest of the dissertation,. Homophily is the core subject of the present work and will be addressed at length in the coming chapters. On the other hand, issues of status and hierarchy are not directly part of my research questions. Nevertheless, they will be touched upon at some point (notably chapters 4 and 6), as they are still related to homophily and friendship formation. Thus, I hope to provide a few empirical insights on these questions as well, although in a more exploratory way than regarding homophily.

2.1. Socioeconomic Homophily among School Students

2.1.1. Expected Levels of Homophily

The fact that students tend to form peer groups that are homogeneous with regard to important attributes appears to be widely accepted. The exact term of "homophily" is mainly used in the SNA literature, but the same idea is found in social reproduction studies (e.g. van Zanten 2000). The strength of socioeconomic homophily specifically, however, remains difficult to anticipate, as there are simply few studies that investigate it. We know that occupational and educational homophily are important in adult friendships (McPherson et al. 2001), but it is unclear whether this is also the case for students. SNA studies that have touched upon this question have generally found low levels of socioeconomic homophily among students (Kruse et al. 2016; Malacarne 2017; McFarland et al. 2014). On the other hand, social reproduction studies tend to see social segregation as an essential feature of the school system, and even of individuals' life courses as a whole; students are thus expected to have friends that are generally similar to them (Herpin 1996; Willis 1977). With that said, the focus of the two strands of literature are different: SNA studies consider the state of friendships within schools, whereas social reproduction studies are strongly interested in inter-school segregation. Therefore, there may be high levels of homophily when considering the entire national population of students, but low levels within each school.

This is a tempting interpretation, but one that should be considered with caution, because the result of low within-school homophily is a rather frail one. To begin with, these findings were obtained in three Western countries (Germany, the Netherlands and the US), but there may be cross-national variations in patterns of homophily, especially considering that national school systems strongly differ. Moreover, the studies that reached this result have three key methodological limitations, which all stem from their lack of interest for socioeconomic background (remember that it is generally used as a control variable in models meant to investigate a different question). First, the variable capturing students' backgrounds may be relatively imprecise, with high levels of missing values and/or an inappropriate definition and operationalization of the concept of socioeconomic origin[20]. Second, estimates come from multivariate models, which means that the potential mediating impact of other covariates on socioeconomic homophily is not considered (see below). Finally, the socioeconomic composition of the studied schools is not taken into account. Yet we know that, in the case of ethnic background, the composition of the population has a strong impact on levels of homophily (Kruse and Kroneberg 2019; Moody 2001): in homogeneous schools with a large majority group, homophily tends to be weaker[21]. Therefore, and assuming a similar pattern for socioeconomic origin, the overall lack of homophily found across US, Dutch or German schools might precisely be due to the fact that most of these schools are rather homogeneous in socioeconomic terms. In this case, one might find higher levels of

20 For example, McFarland (2014) reports high levels of missing values on the variable of parental occupation in one of the two datasets; and there is simply no socioeconomic information whatsoever in the second. Moreover, both McFarland (2014) and Malacarme (2017) use an arguably rough measure of socioeconomic background (five-step ordinal scales for parental occupation and education, which are then used as continuous variables).

21 To be clear, homophily estimates are typically computed as relative odds of friendship ties to be observed, which means that they are not sensitive to the sheer size of the different groups in the population. This suggests that the relationship between homophily and the composition of the population is not just a mechanical consequence of uneven group sizes, but a substantial change in individuals' relational behavior (though untangling individual behaviors and structural effects is always tricky, especially since individuals usually change their behaviors in response to changes in structural opportunities).

homophily when focusing on highly mixed schools, such as the ones considered in the present dissertation.

2.1.2. What is Socioeconomic Origin Made of? Economic and Cultural Dimensions of Social-Class Belonging

Another difference between SNA and social reproduction studies pertains to their conception of students' socioeconomic backgrounds. Indeed, there are a variety of ways to define socioeconomic positions for adults; for example, "social class" and "socioeconomic status" are related but nevertheless distinct notions (Chan and Goldthorpe 2007). This leads to different methods for building categories or scores of social positions, with scientific and political debates surrounding the statistical institutions that produce these tools (Penissat and Rowell 2012). Crucially, this is not only due to theoretical disagreements between researchers, but also to the fundamentally complex and multidimensional nature of social stratification. Furthermore, the ways in which parents' socioeconomic positions are transferred onto their children and impact their friendships, is itself complex; certain definitions of socioeconomic positions may prove more heuristic than others in investigating adolescent homophily specifically.

One way to think of the different components of socioeconomic positions is to distinguish between economic and cultural capitals (Bourdieu [1979] 2016)[22]. These have different implications in terms of relational processes. The economic capital of parents affects children's relations through conspicuous consumption, residential segregation, and access to specific activities, either within (e.g. charged cafeteria) or outside of school (e.g. going to the movies). Cultural capital, on the other hand, may shape students' preferences, tastes, values, or interests, which in turn will affect their relational behaviors and choice of activities. As a result, a given measure of socioeconomic origin may be either more or less appropriate for testing different hypotheses, depending on the relative weights it attributes to these two forms of capital[23].

22 Closely related notions are Coleman's economic and human capitals (Coleman 1988). Note that both Bourdieu and Coleman consider a third type of capital, social capital, which depends on one's relations to others with valuable resources. This may also be relevant for socioeconomic homophily among children, in which case one would want to look at the impact of parental networks on children's friendships. However, doing so by integrating measures of social capital within the individual-level variable of socioeconomic background would be a rather inefficient way of using the information about parental networks. Rather, it is best to rely on dyadic measures that directly consider the impact of relationships among parents on the relationships of children. For this reason, I only consider economic and cultural capitals in the discussion of how to construct an individual-level score of socioeconomic background.

23 Of course, economic and cultural capital tend to correlate rather strongly (Bourdieu [1979] 2016), which means that a measure focusing on a single one of them would nevertheless make for a decent proxy of the other. Still, this would introduce additional noise to the measurement and thus make it harder to detect rele-

In that regard, several SNA studies lean toward an interpretation of socioeconomic homophily among school students in terms of economic inequalities, rather than cultural differences. This is particularly apparent in cases where the measure of socioeconomic origin is directly based on material wealth, or when researchers hypothesize that lower-class students suffer from a lack of social integration because they are not able to participate in activities that require money with their schoolmates (Hjalmarsson and Mood 2015; Olsson 2007). In other studies, some measure of parental education is used, but the interpretative frame remains that of economic resources (e.g. Malacarne 2017, whose article is quite tellingly called “Rich Friends, Poor Friends”). On the other hand, social reproduction studies, though they do not reject interpretations in terms of economic resources, generally pay significant attention to cultural and academic capitals. This is apparent in their interest in youths’ cultural tastes (e.g. Willis 1977 with rock music), as well as in the focus on students’ academic results and on the corresponding homophily (e.g. van Zanten 2000). The debate around the internalization of adult categories of perception by children (Delalande 2003; Lignier and Pagis 2014; Zarca 1999) also comes closer to a perception of social background as primarily made of cultural capital, since researchers are interested in students’ psychological traits, rather than external properties like consumer goods.

I will present my own measure of socioeconomic origin in chapter 3 (section 1.2.3). Compared to those found in the SNA literature, it gives more weight to cultural capital and less to economic capital. Indeed, I have reason to believe that socioeconomic homophily primarily operates through cultural rather than economic factors, at least in the studied schools.

2.1.3. Mediating Factors of Socioeconomic Homophily

This leads us to the broader question of the mechanisms through which socioeconomic origin can shape students’ friendships. Observing socioeconomic homophily is not in itself, indicative of discriminatory friendship choices. Students from similar backgrounds are more likely to come into contact with one another, to share certain tastes, values or other traits, and to be embedded in the same social networks, all of which can push them toward befriending similar peers without them directly caring for their friends’ background. This has been largely discussed in the SNA literature regarding ethnic homophily. Tuma and Hallinan (1979) have shown the importance of academic achievement, Moody (2001) of the composition of schools, Stark and Flache (2012) of shared tastes and opinions and Kruse and Kroneberg

vant effects.

(2019) of the broader local context in inducing higher or lower levels of ethnic homophily among school students. One may expect these mechanisms to be relatively similar in the case of socioeconomic homophily. As for social reproduction studies, they have emphasized the role of academic results and attitudes (Cousin and Felouzis 2002), residential segregation (Oberti et al. 2012) and cultural tastes (Lepoutre 1997; Willis 1977) in fostering same-background friendships, this time for both ethnic and socioeconomic homophily. However, unlike SNA studies, they do not offer quantitative estimates for the strength of these inducing effects.

This question of the mediating factors of socioeconomic homophily is one of the main threads of the dissertation. Identifying them can prove tricky from both a theoretical and methodological point of view. Therefore, I will not go into more details here; chapter 2 is entirely dedicated to this topic and will review the relevant literature in more depths.

2.1.4. Autonomy of Adolescent Groups or Internalization of Adult Criteria?

Finally, an interrogation that runs through a large part of the literature – both SNA and social reproduction studies – is that of the autonomy of adolescent subcultures. On the one hand, it is clear that students' friendships are not independent from properties of their household, nor from adult-enforced principles of differentiation (grades and school regulations in particular). As we saw, Lignier and Pagis (2014) argue that students sort their peers through "tools of thought" imported from the school or family; this seems coherent with the ethnic and academic homophily consistently observed across SNA studies, even when potential confounders are controlled for. Several ethnographies also found academic results and attitudes to be an important element of differentiation among peers, with relatively frequent conflicts between "good" and "bad" students (Cousin and Felouzis 2002; van Zanten 2012). Some studies also argue that students with low results are generally mocked or marginalized within peer groups (Delalande 2003; Gayet et al. 2003; Lignier and Pagis 2012), which would point to an overall internalization of teachers' expectations and morals by most students.

On the other hand, many studies point toward the ability of children and adolescents to create cultural worlds of their own, independent from – and occasionally opposed to – adults' rules and expectations. Remember that Coleman (1961) insisted on youth peer groups' search for autonomy, while McFarland (2001) saw resistance to teaching as a somehow natural inclination of students, regardless of their social background. This stands directly opposed to the previous conclusion that students internalize and conform to academic hierarchies. We also saw that Cousin and Felouzis (2002) speak of a "critical distance" toward teachers that middle school students progressively develop, and that Dubet and Martuccelli (2014) as well as Rayou (1998) observed that high school students were reluctant to openly classify each other based on their grades, as these school-based categories appeared illegitimate to them. Moreover, one study that can be associated with social reproduction research and that I have not mentioned yet is Merle's book *The Humiliated Student* (2015). Merle argues that most French students have felt mocked, belittled or unfairly punished by teachers at some point. Consequently, resentment toward school is not restricted to "bad" students nor to working-class ones. Altogether, what these various contributions point toward is the fact that differences within the student population may sometimes be over-shadowed by another opposition, that of adults and teenagers, or of teachers and students. Thus, the school context may, at the same time as enforcing differentiation through academic rankings, also contribute to constructing socially different students as fellow classmates and peers.

This may lead students to reject adult norms, by either reversing them (e.g. rebellious students are seen as cool) or, perhaps more simply, by relying on other means of differentiation, rooted in youth subcultures and mostly orthogonal to adult categories (e.g. musical tastes). Indeed, researchers in the sociology of culture insist on the fact that children and adolescents share certain references or tastes specific to their age group, supported by communication media and by a cultural industry that specifically target them; all of which is relatively obscure to adults (Octobre et al. 2010; Pasquier 2005). It has even been argued (controversially so) that peer-to-peer socialization and generational trends now dwarf socioeconomic background in determining adolescent cultural practices (Pasquier 2005). We also saw that childhood studies emphasize children's ability to create their own shared cultures, codes and values (Delalande 2003).

Altogether then, it is not clear to what extent students rely on parents' and teachers' mental categories in organizing their sociability. This may depend on age: as they transition from children to adolescents, students should be increasingly concerned with the opinion of their peers rather than parents, and value autonomy from adults in general (Adler and Adler 1998; Octobre et al. 2010). This should in turn impact patterns of socioeconomic homophily: one would expect the development of age-specific cultures, or the shared opposition of students to teachers, to attenuate social distances. On the contrary, a strong internalization of academic rankings, or of the tastes, values and beliefs of one's family, might pull socially different students further apart, as they tend to differ in these regards. Interestingly, the rejection of adult categories may also contribute to increasing social distances, in cases where these categories are reversed rather than ignored: the hostility of “bad” students toward academically dedicated peers (“teacher’s pet”, “smart-ass” etc.) should foster academic homophily and thus, indirectly, socioeconomic homophily as well.

2.2. Peer Status and Socioeconomic Background

2.2.1. Homophily or Hierarchy?

According to Malacarne (2017), socioeconomic background tends to shape friendship networks along a vertical dimension, rather than a horizontal one: in other words, status prevails over homophily. He argues that upper-background students receive more friendship nominations overall, regardless of the social composition of the school they are in. He contrasts this pattern with that of ethnicity, for which, on the contrary, students tend to receive more nominations in schools where their ethnic group is larger. Consequently, while ethnicity, as a social principle, would primarily consist of discrete groups, with most students holding a preference for in-group friends, socioeconomic origin would scale individuals along a continuous axis, with higher positions being more prestigious from the point of view of all students. This is coherent with the argument made by Hjalmarsson and Mood (2015) and Olsson (2007) that poor students tend to be marginalized because of a stigma of poverty, as well as an inability to take part in consumer activities. Interestingly, it seems that this hierarchical perspective is most prominent among researchers that conceive of socioeconomic positions as primarily made of economic capital. Indeed, if students are differentiated by how much money they own, then one can clearly be thought of as having “more” or “less” than others, in a strictly hierarchical fashion.

However, other studies perceive socioeconomic origin as somehow analogous to ethnicity, in the sense that students can be differentiated horizontally, and that they exhibit homophilic inclinations. This is notably the case of McFarland et al. (2014), who propose a generic theoretical framework for the contextual moderation of homophily, which is then applied to both ethnicity and socioeconomic origin (more on this in chapter 2). Ethnographic studies in working-class segregated schools also report a rather intense sociability, with a shared culture and a clear sense of group identity among working-class adolescents (Lepoutre 1997; Willis 1977). Although the question of homophily appears rather abstract in these cases because there is very little opportunity for cross-class friendships in segregated schools, we would nevertheless expect these adolescents to exhibit some form of in-group preference should they attend more diverse settings. In any case, under this view, socioeconomic positions are not only differentiated by a greater or lower quantity of some resource, like money; they are also characterized by a substantive cultural “content”, i.e. norms, values, tastes or beliefs specific to certain positions.

Altogether then, it is an empirical question whether social origin primarily elicits homophily or status asymmetry among students (or a mixture of both). Fortunately, the network models that will be introduced in the empirical chapters are particularly well-fitted to answering this type of questions.

2.2.2. Majority Patterns and Non-Linear Effects of Socioeconomic Background

The above discussion would not be complete without considering the impact of the local school context on homophilic and statutory patterns. Most notably, this context depends on the socioeconomic composition of schools. Indeed, many ethnographers found conformity to the group to be a dominant feature of youth subcultures, with popular students being the ones that best incarnate shared norms or values. For example, Coleman (1960) showed that popular students are often those that represent their school in sport competitions. In a different vein, Cousin and Felouzis (2002, pp.90-94) suggest that well-liked students are the ones that best navigate between two contradictory aspirations of the student body, that of doing well at school and that of belonging to the peer group (by opposition to adults). Students that are perceived as extreme, i.e. deviant, in either aspect (being too dedicated to school work, or on the contrary too opposed to teachers) will thus suffer rejection from the group. Applying this idea of a “majority rule” to the case of socioeconomic origin, one would therefore expect status hierarchies to depend on the school’s composition – which may also explain why different views are found in the literature as per the horizontal or vertical nature of socioeconomic divides. It may be the case, for instance, that working-class students tend to be dominated symbolically and, thus, “upward looking” (i.e. seeking the friendship of higher-background peers), in schools where they make for a minority of the student body. On the contrary, they may develop a positive group identity with strong homophilic inclinations when they are in a majority.

Another important point is that the effect of socioeconomic origin is not necessarily linear. By this, I mean that the relationship between socioeconomic background and sociability patterns does not have to follow a homogeneous pattern from very low to very high origins, but that there may be threshold patterns. For example, it may be the case that the effect of economic deprivation on social isolation only appears for students that come from particularly precarious households (e.g. both parents are disabled and unemployed). At the same time, students coming from relatively established working-class families (e.g. factory workers with a stable job) may be equally integrated as middle- or upper-background students, as they would not suffer from the stigma of extreme poverty, or simply because they would be numerous enough to form cohesive groups based on homophilic bonding. In such a case, the effect of socioeconomic origin on network centrality would appear weak overall (many working-class students would be as popular as upper-class ones), but it would be concentrated on a few students for whom it makes a big difference (very poor students are much more likely to be isolated than most of their peers).

Beyond threshold effects, it is also possible that socioeconomic origin relates to network centrality following a curved pattern. It may be the case that adolescent subcultures penalize extreme origins, just like they do for academic attitudes; conformity to the majority of the population would matter more than standing at the extreme of a status scale. Thus, both working-class and upper-class students may be at a statutory disadvantage, as they both represent extreme and minority groups in a context of social mixing. This would imply that middle-class students are best integrated within peer groups. Moreover, socioeconomic homophily could reinforce such a pattern: middle-class students are structurally closer from most of their peers in terms of socioeconomic origin, so they may simply have more people relatively close to them with whom to form homophilic ties, resulting in more friends overall. Once again, this is a relatively straightforward prediction to test, which will be tackled in chapter 4.

2.2.3. Peer Status and Academic Results

Finally, it is interesting to consider the interplay of peer status, socioeconomic background and academic results and attitudes. Here as well, there appears to be disagreements within the literature: some researchers found high grades or dedication to school rules to be depreciated in student groups (Coleman 1960; McFarland 2001) while others argued that “bad” students were commonly mocked by peers (Delalande 2005; Gayet et al. 2003; Lignier and Pagis 2014). This is strongly linked to whether student subcultures are perceived as autonomous, or on the contrary aligned on adult expectations. In any case, as mentioned in 2.2.2. above, this tension can be resolved by considering that peer status and academic results are associated following a non-linear pattern, with students generally valuing a certain amount of adhesion to academic criteria while rejecting the peers considered as "brainies" who by-pass this implicit limit (Adler and Adler 1998; Cousin and Felouzis 2002; van Zanten 2012). Moreover, certain academic attitudes may be ambiguous in terms of status: different statutory principles could co-exist within peer groups – particularly in mixed schools – such that a given attitude may be valued by some and depreciated by others. This forces us to envision peer status as being multidimensional, i.e. irreducible to a constant and unambiguous hierarchy that would rank individuals from most to least prestigious. Furthermore, it echoes with the distinction between social acceptance and social impact in social psychology, which underlines the multidimensional nature of peer status. Troop-Gordon et al. (2011) found high academic results to be positively correlated to acceptance but negatively correlated to impact.

Regarding socioeconomic origin, one may expect its relation to peer status to be relatively similar to that of academic results, given the pronounced correlation between both (higher-background students tend to do better at school). Nevertheless, Passerieux (2010) argues that good students from a lower-class background are particularly sanctioned by peers when they appear as overly dedicated to school-work – an attitude that, on the contrary, is seen as normal among upper-background students. If this is true, then the relationship between socioeconomic background, academic results and peer status is a complex one: upper-background students are more likely to exhibit very high academic results, which tend to be depreciated within peer groups, but they may also suffer less from this in statutory terms compared to the (fewer) lower-background students that have similar such high results. However, once again, there is a relative lack of empirical evidence supporting this view, which should therefore be considered with caution at this point.

Conclusion

Many of the gaps in the literature come from the lack of discussion between SNA and social reproduction studies. On the one hand, SNA research on adolescent friendship has not frontally tackled the question of social class belonging and of the socialization attached to it, nor did it question the role of the school institution in activating or exacerbating social distances between students. On the other hand, social reproduction studies have a sometimes crude approach to sociability, and they lack the quantitative tools to compare relational features across different contexts in a standardized way, as well as to isolate the effect of specific mechanisms on tie formation using multivariate modeling. With that said, some of the theoretical disagreements pointed above are substantive ones, which sometimes pertain to the very conceptual foundations of the different approaches[24]. Theoretical syncretism always bears the risk of reaching contradictions, and my point is not to blindly mix incompatible approaches. Still, there is room for studies to draw insights from these different traditions. This is particularly true on the methodological side. Network analysis and ethnographies are not so frequently combined, which is in part due to the strict technical specialization that each of them demand. Yet in the family of quantitative methods, network analysis is probably one of the most appropriate to mixed-method research, since it best applies to the “meso” level, i.e. intermediate social structures made of a few tens or hundreds of individuals.

This integration of different approaches appears as necessary in order to answer some of the main research questions of the thesis. Explaining the apparition of socioeconomic homophily in adolescent friendship demands that we consider the many properties that are associated with socioeconomic origin, including cultural ones, but also to account for complex network processes that structure relationships above and beyond individual characteristics, which require dedicated statistical models. It is also important to place student relationships back into the context of the school, and to consider the impact of academic hierarchies and of their relation to students’ background. Finally, considering socioeconomic homophily within certain schools may lead to misleading conclusions if the type of socioeconomic diversity found in this school, and its relation to inter-school segregation, is not taken into account. The fact that many upper-class parents try to maintain some control over their child’s relationships, and that they attempt to avoid excessive levels of mixing in their schooling choices, means that the population of schools is not selected at random: not any household ends up in socioeconomically diverse schools. For this reason, the strength of socioeconomic homophily cannot be thought of independently from the composition of a school, nor from broader mechanisms of social segregation that operate at the inter-school level. Within-school homophily and inter-school segregation are two complementary dimensions of socioeconomic homophily and of social reproduction at the global level, which calls for greater discussion between their respective bodies of literature.

24 Although I have not directly mentioned it, because it can vary quite significantly depending on the authors, the broader theoretical frameworks found in the two strands of literature are also rather different. For instance, Rational Choice Theory has been very influential in SNA research (Wittek, Snijders, and Nee 2013), whereas part of the field at least in social reproduction studies has clearly rejected it. The distinction between positivism and constructivism can also be relevant to qualify certain studies in the SNA and social reproduction literature respectively, though not in all cases (there are also strongly positivist or “objectivist” approaches in the sociology of education, in a Durkheimian and/or Bourdieusian vein).

Chapter 2: Context and Dispositions: a Theoretical Framework for Understanding Homophily

Socioeconomic homophily can be the result of several different social processes. In order to explain how and why it appears, it is therefore necessary to identify and disentangle these processes, assessing which are operative and which are not, as well as the ways in which they can mutually reinforce or inhibit one another. While this is the task of empirical research, it demands a clear theoretical framework for identifying these different processes and understanding the methodological conditions to their empirical study. Consequently, the objective of this chapter is to propose a typology of explanations for homophily and a model of their relations to one another.

This is a topic that has been extensively discussed in the Social Network Analysis (SNA) literature – much more so than in social reproduction studies –, and the typology I will propose largely overlaps with those commonly used by researchers in that field. Nevertheless, some of the concepts typically used are either incomplete, or overlap with one another. Different researchers also use different names for closely related concepts. As a consequence, we lack a unified and unambiguous framework. To remedy those issues, I try to integrate various notions of network analysis within the broader theory of human action that has been proposed by Lahire (2005)[25]. More than a theory strictly speaking, it is fair to call Lahire's work a grammar: it offers basic concepts for describing human action, while remaining relatively agnostic as per the substantial postulates that one is willing to make about the state of the social world. By borrowing these basic concepts and the related vocabulary, it is therefore possible to propose a typology of homophily-inducing processes that is flexible enough to integrate the processes typically discussed in the SNA literature, while being rigorously built and using unambiguous concepts and categories. In that sense, the model proposed in this chapter is not new per se, but rather a new shell put around existing notions.

This model will then serve as a basis for the rest of the dissertation, as the different chapters will tackle different aspects of the friendship formation process. Nevertheless, not all

25 Note that Lahire's work could be attached to the field of social reproduction studies described in chapter 1, even though his theory has a very large scope and is therefore not only meant to study social reproduction (nor does it focus on school settings). In any case, as I mentioned before, what I called "social reproduction studies" make for a less coherent and more heterogeneous field overall than SNA studies.

the elements raised here will be met with an empirical answer, as the number and complexity of the social processes underlying friendship formation far exceed the scope of this dissertation. Therefore, the objective is also to open a discussion with the rest of the field, and to suggest directions for future research.

The chapter is organized as follows. Section 1 introduces the proposed model of friendship formation, starting with basic notions commonly found in the SNA literature, then integrating them within Lahire's framework. While this model is somewhat generic and could apply to several different topics, most of the discussion will be geared toward adolescent friendships. Then, section 2 applies the model to the case of socioeconomic homophily among adolescents and reviews the empirical evidence that exists in the literature for the different theoretical processes identified. Finally, section 3 proposes a short discussion of some methodological implications of particular interest. The conclusion of the chapter describes how the rest of the dissertation will fit within the proposed framework.

1. A Generic Model of Friendship Formation

The question of how homophily emerges can be thought of as an instance of a broader problem: why do individuals engage in some relationships rather than others? How can we explain the fact that ties are not formed at random? The concept of homophily refers to a relational configuration of particular interest, where ties are more frequent among similar nodes: it is therefore one particular outcome of the social processes by which relationships are formed. As such, any explanation of homophily relies, either implicitly or explicitly, on a model of friendship formation[26].

In this section, I first present the basic model of friendship formation implicit in a large part of the SNA literature (1.1) and discuss the issues it raises (1.2). Lahire's theory of human action is then introduced (1.3), and the solutions it offers to the previously identified issues are discussed (1.4 and 1.5). From there, I propose a new model for friendship formation, first in a basic (1.6) then extended (1.7) form.

1.1. Naive Model

SNA researchers commonly distinguish between two large classes of explanations for homophilic patterns: those based on differential meeting opportunities, and those based on the

26 Note that homophily does not only refer to friendship – one could for example, consider homophily in advice, gossip or professional networks. However, I restrict the discussion here to the case of friendship, which is the main theme of the dissertation.

preferential selection of others with specific characteristics. First, not all individuals are equally likely to meet, nor are they equally likely to meet in contexts that are favorable to tie formation. The concept of *foci* (Feld 1981) refers to all the entities that structure social interaction, including activities, institutions or simply geographical spaces[27]. People usually meet and bond around specific foci; and most foci tend to gather individuals with particular properties relative to the general population (e.g. leisure activities or workplaces are foci that tend to be strongly segregated by social class). Second, and net of the access to particular foci, individuals try to satisfy certain subjective preferences in forming relationships. In most cases, this implies a preference for similar others, as similarity on relevant characteristics would make agreement and understanding easier (McPherson et al. 2001); increase the predictability of other people's behavior (Hamm 2000); and facilitate the emergence of positive group identity, as argued by social identity theory (Turner 1987).

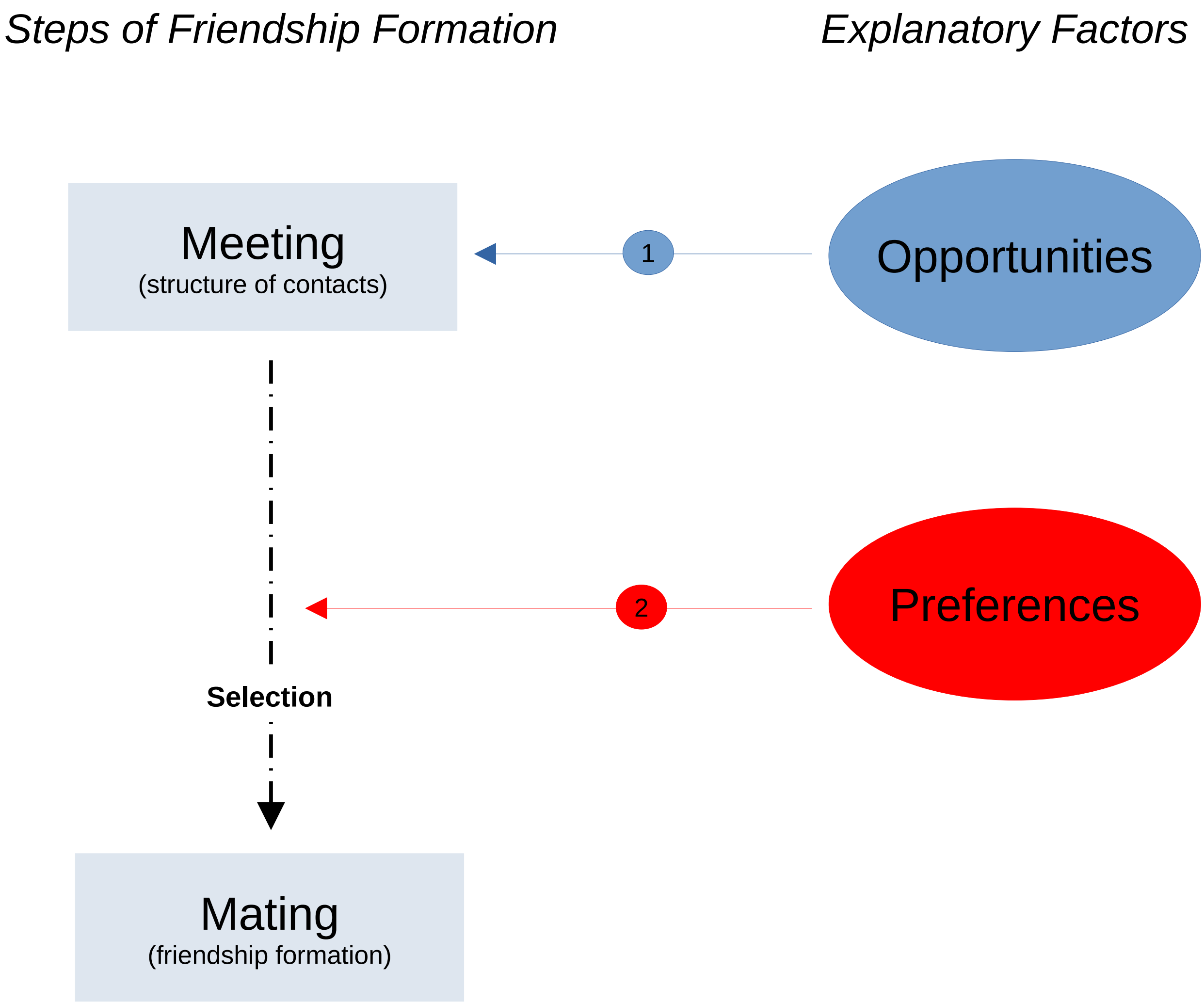
This general dichotomy can be found under different names depending on the authors, such as structural opportunities vs individual preferences (Blau 1994; Hallinan and Teixeira 1987), induced homophily vs choice homophily (Kossinets and Watts 2009) or baseline homophily vs inbreeding homophily (McPherson et al. 2001). For now, I will use the terms "opportunities" and "preferences". This terminology makes it clear that the field of social network analysis is largely influenced by rational-choice theory, which makes abundant use of these two concepts (see Snijders 2013, for an example of how rational-choice theory is articulated to network modeling).

What model of friendship formation does this suggest? For a friendship to be formed, individuals first need to meet, and then to select one another as friends. Verbrugge (1977) coined the terms "meeting" and "mating" for these two steps. In its simplest form, the opportunity/preference couple can therefore be applied to these two steps separately: individuals meet certain others because of structural opportunities, then select certain friends within that pool based on their preferences. This is illustrated in Figure 2-1[28].

27 Somehow equivalent concepts are that of "social circles" (Simmel and Hughes 1949), "social settings" (Pattison and Robins 2002) or "social situations" (Block 2018).

28 I must express my thanks to Marion Hoffman for having first thought of this way to represent meeting and mating processes.

Figure 2-1: Simple Model for Friendship Formation



A few words of caution as per the interpretation of this model. First, in reality, meeting and mating are not two distinct, well-bordered moments; each may be spread over time, and they may overlap (e.g. two persons ca n meet several times in different contexts, and the process of becoming friends can be diffuse over their encounters). Additionally, meeting should not be seen as binary (having met or not), but rather as continuous: the more two persons are put into contact, the more likely it is for them to become friends (though probably with decreasing marginal returns; Block and Grund 2014). This is also referred to as "propinquity" (Kruse et al. 2016), where the likelihood of a tie depends on the physical proximity between two persons. Therefore, the "selection" arrow between the "meeting" and "mating" boxes in Figure 2-1 indicates a logical and causal connection (it is necessary to meet in order to become friends) more than a strict temporal ordering.

Second, "friendship formation" should be understood in an extensive sense, as referring to all the processes that explain why a friendship exists between two individuals at a given time. As such, it does not only pertain to the initial formation of a friendship, but also to its ability to persist over time and, consequently, to its odds of dissolution. Typically, some of the attributes relevant to preference-based selection (process 2 on the figure) may only become apparent in the course of a relationship, thus impacting one's willingness to maintain an existing friendship or not.

1.2. Limitations of the Naive Model

This initial model raises a number of issues, that are well-identified in the SNA literature. First, as Cepić and Tonković (2020) note, the distinction between preference-based choices and meeting opportunities gets blurred in cases where individuals self-select in certain foci. For example, homophily on smoking behaviors might come from smokers' increased contact opportunities when they gather to smoke outside: even if they do not particularly look for smoker friends, they self-select into a context that will make it easier to meet fellow smokers. In this example, homophily is induced by individual behaviors (smoking), but this is not a relational preference strictly speaking, and the final tie-formation mechanism at work pertains to contact opportunities rather than the direct selection of certain friends. It is therefore necessary to distinguish between imposed meeting contexts (such as schools or classrooms for teenagers) and chosen meeting contexts. While the first are independent from individuals' preferences, the second are not.

Second, individuals are socialized within foci, or through the relationships that they already have. This in turn affects their preferences, creating retro-active loops between meeting opportunities, friendships and homophilic preferences. In the network literature, this is often known as the "selection/influence" problem (Brechwald and Prinstein 2011): does homophily emerge because people select similar others as friends, or because friends influence one another into becoming similar? To carry on with the example of smoking, evidence has been found for both types of processes among high-school students: smokers tend to become friends, and friends of smokers also tend to start smoking (Mercken et al. 2009; note that in this study, selection effects appear stronger than influence ones).

Third, elements of the general context impact the expression of individuals' preferences. For example, McFarland et al. (2014) argued that the propensity of US secondary-school students to choose same-ethnicity friends depended on the educational climate of the school: in adverse or hostile environments, students would feel more threatened, and therefore increasingly look for others who they consider similar to them. I will use the expression *contextual moderators* (McFarland and colleagues talk of "ecological moderators") to refer to these properties of the context that affect individuals' capacity or willingness to act upon their latent preferences.

Finally, relationships are not independent from one another. The existence of a tie depends on the state of the network in which that tie is embedded, through mechanisms such as reciprocity (people look for friends who also want to be their friends) or triadic closure (the principle that "the friends of my friends are my friends"). These are called network endogenous processes ("network processes" for short), and they tend to produce non-linear relationships between individual preferences and the aggregated level of homophily that is measured in a network. For example, Goodreau et al. (2009) found that triadic closure bolstered ethnic homophily among US teenagers, by aggravating the dyad-level propensity to select same-ethnicity friends.

1.3. From Opportunities and Preferences to Context and Dispositions

These four problems point to the limitations of the opportunity/preference couple: certain factors affecting the formation of relationships simply do not fall into either category. The solution I propose is to enlarge these categories, which implies redefining "opportunities" and "preferences" as subcategories of two other concepts, namely "context" and "dispositions", borrowed from Lahire's theory of human action (Lahire 2005)[29].

One advantage of the opportunity/preference couple is to distinguish between factors that are external to the individual (the opportunities they are met with, that pertain to the state of the world) and those that are internal to them (psychological traits). "Context" and "dispositions" are a generalization of this idea. In Lahire's framework, any individual action is the joint result of the external context to which an individual is exposed at time *t* and of their psychological dispositions, which can be either biological endowments or traces of past socialization experiences. These dispositions include relational preferences, but also the preferences for certain activities or spaces, as well as any form of habits, inclinations or

29 All further descriptions of Lahire's model are taken from this same reference.

abilities. As for contexts, they include all the elements that are external to the individual at time *t* – including, from one's perspective, other individuals and their behavior.

Crucially, context and dispositions are not separate forces, but constantly interact. Not only are individuals socialized into the contexts that they experience, thus affecting their dispositions, they also self-select into certain contexts based on their dispositions. Moreover, one of the prominent ways in which contexts shape behavior is by activating certain dispositions within individuals. Lahire sees most humans as dispositionally heterogeneous, in the sense that numerous, sometimes contradictory dispositions can co-exist in a single person. Contextual incentives then determine which of those dispositions are activated in a particular situation.

What advantages do these broader concepts offer in order to explain friendship formation? First, context includes both structural opportunities for meeting, and what I referred to as *contextual moderators*; that is, elements of the context that moderate the expression of individuals' dispositions. These moderators differ from structural opportunities in that they do not affect the frequency of physical contact among individuals, but rather, given these contacts, affect selection behaviors: as an explanatory concept, "opportunities" pertains to the "meeting" step of friendship formation, whereas contextual moderators intervene in the "mating" process of selecting certain available partners over others. For example, even if we assume that many individuals have a latent, intrinsic preference for same-ethnicity peers, it should be clear that the strength of the selective behaviors entailed by this preference will vary depending on the salience of ethnic labeling and identities in any given situation. The types of activity available to individuals might also impact homophilic behaviors: two persons with few cultural references or interests in common might find it hard to hold a long, face-to-face conversation when they first meet, but this might not matter as much if the context provides them with a shared goal or interest (e.g. sport teams).

Second, dispositions include a number of traits that are less conscious or purposive than the notion "preference" generally implies. For example, friendship formation might be facilitated among individuals that have the same ways of speaking (slang, accent, etc), because it eases communication; yet this is not strictly speaking a preference. Moreover, differentiating between dispositions at large, and dispositions that are (at least partly) explicit, conscious and goal-oriented – i.e. "preferences" in the narrow sense – is valuable from a methodological point of view. Indeed, only the latter can be directly identified within individuals' discourse about their relational choices ("I am friends with X because..."). As such, preferences can be investigated through declarative methods like survey questionnaires or interviews, whereas non-conscious and non-purposive dispositions will generally need to be inferred from observed practices.

1.4. Joint Determination of Meeting and Selection

We can now apply these concepts of context and dispositions to the issues raised above with our naive model of friendship formation. The key limitation of this simple model was the monotonic association of opportunities and preferences with, respectively, meeting and selection. Instead, the notion of dispositions being activated by a given context implies that both meeting and selection are joint products of contextual and dispositional factors.

In that regard, it is important to differentiate between latent and expressed dispositions. *Latent dispositions* refer to individuals' intrinsic traits, which are virtually independent from the particular context they are part of at any given moment. Nevertheless, latent dispositions are never expressed in a void, but always through their encounter with a particular context. *Expressed dispositions* are therefore the joint product of latent dispositions (which generally cannot be observed directly) and of contextual moderators – so to speak dispositions that have been filtered through contextual elements and translated into specific behaviors. Note that this distinction between latent and expressed dispositions applies to any type of disposition, whether explicit preferences or more subtle and unconscious traits.

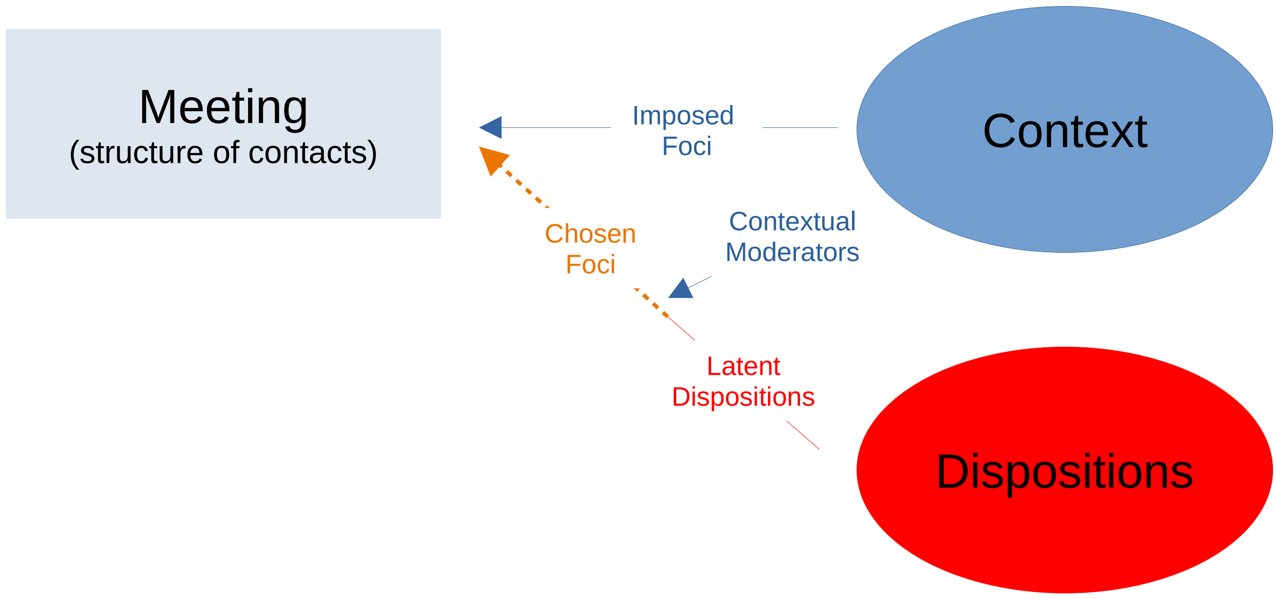
With this in mind, we can describe the processes leading to a certain structure of contacts (meeting) in a population. Individuals' contact opportunities depend on some foci that are outside of their control – *imposed foci* – but also on the social situations they self-select into – *chosen foci*[30]. Whereas imposed foci are independent from individuals' dispositions, i.e. a purely contextual factor, chosen foci are expressions of individuals' dispositions, thus a combination of latent dispositions and contextual moderators. Contextual moderators include the available foci among which individuals can choose, but also any contextual element making the dispositions toward certain foci either more or less salient.

For example, in a given school, imposed foci can be classrooms or ability groups (to the extent that they are imposed upon students by the school administration, which is generally the case in France). Chosen foci may be clubs that offer extra-curricular activities:

30 The use of the adjective "chosen" does not necessarily imply that participation in these foci reflect individuals' desires or expectations. Consider, for example, the case of school students that are in detention after breaching school rules. Attendance to this foci (in which friendships certainly do form) is a form of self-selection, because it depends on individuals' behavior, which is disposition-driven. Yet it is clear that students did not "choose" to be detained in the usual sense of the word.

students decide to attend those based on their own interests or desires (dispositions). This choice is of course constrained by the type of clubs effectively available; but it will also depend, for example, on which clubs are seen as prestigious in the school, or on whether teachers try to push certain students toward a given club (e.g. a music teacher may try to enroll students with a nice voice in her choir). All of these are contextual moderators: they do not really deprive students from their ability to pick the foci they want, but they affect how their various dispositions will activate and combine into a given behavior (e.g. a student may both like football and have a desire to be popular, in which case whether the football club of the school is prestigious or not may result in a different club choice even from the same set of initial dispositions).

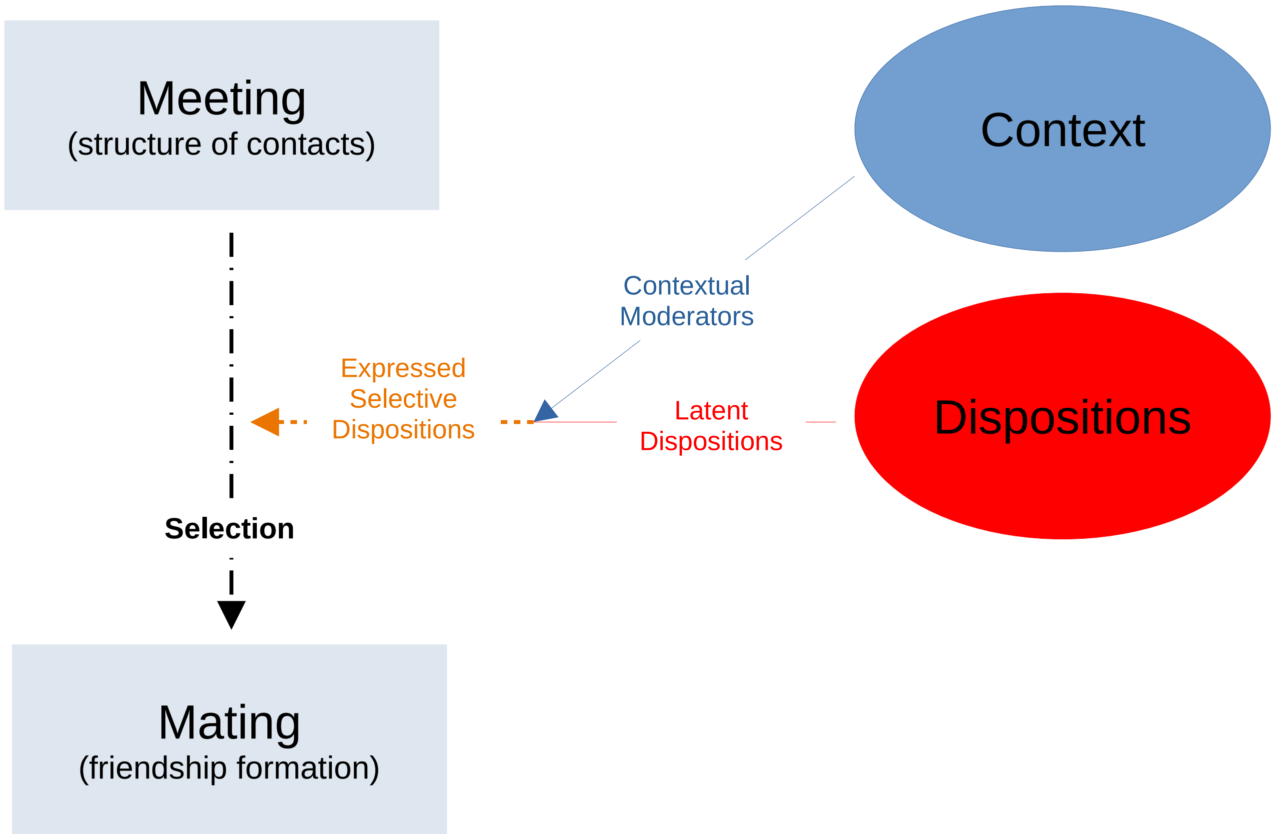
Figure 2-2: Determinants of the Structure of Contacts in a Population



Moving on to selection processes. These are, like chosen foci, a form of disposition-driven behavior, thus the joint result of latent dispositions and contextual moderators. I call *Expressed Selective Dispositions* those expressed dispositions that are relevant to processes of friendship selection. Note that by convention, when using the expression "expressed dispositions", I will implicitly refer to expressed selective dispositions exclusively – chosen foci, though also a type of expressed dispositions, will be called by their specific name. Similar to the basic model of section 1.1., selection operates given a pool of "candidates", i.e. given the structure of contacts (meeting). This is represented in Figure 2-3 below.

For example, one type of expressed selective dispositions may be the importance given by students to ethnicity in selecting their friends at school. This will depend on their intrinsic beliefs or habits regarding ethnicity, such as whether they hold racial prejudices (latent dispositions); but also on how salient ethnic categories are in the school (contextual moderators). This could include how older students refer to each other using ethnic labels, whether certain teachers implicitly or explicitly mention students' origins, or how strongly ethnic background is correlated to academic results (if the correlation is high, then prejudices associating certain origins and academic behaviors may appear more relevant to students, reinforcing their belief in these prejudices' usefulness for selecting friends).

Figure 2-3: Determinants of Selection Processes in a Population



There is an asymmetry with the previous Figure 2-2: whereas context could sometimes directly impact meeting, regardless of individuals' dispositions (imposed foci), there is no equivalent here. Indeed, dispositions are never expressed outside of a particular context. It could be the case, though, that some latent dispositions are so strong and consistent that they will have very little sensitivity to contextual moderators (e.g. strong racial prejudices, where one tries to avoid persons of color regardless of the situation). In general, the relative importance of contextual moderators and latent dispositions can vary considerably.

1.5. Network Processes

Another key issue I raise concerning the opportunity/preference pair is that of endogenous network processes; that is, the impact of the current state of the network on individuals' relational behaviors. How does this class of explanations sit within a dispositional-contextual framework? From the perspective of any given individual, the surrounding network structure, made of other people's relations, is clearly something external, thus a feature of the context. Nevertheless, there are various ways in which this contextual element can impact friendship formation.

First, some authors have treated network ties as being essentially a type of foci (e.g. Block 2018). Shared friends can act as the medium by which two persons meet. Conversely, two individuals who inhabit different sociability circles are unlikely to come into contact. In other words, the structure of the network impacts contact opportunities among individuals (one may speak of a "relational propinquity").

Second, it has been argued that network effects reflect psychological traits and/or cultural norms in friendship formation. This is best illustrated by the idea of "balance": closed configurations (either reciprocal dyads or transitive triads) would be more stable because they would offer more comfort and security to individuals (Cartwright and Harary 1956). For example, it would be uncomfortable for a person to befriend someone that does not reciprocate their positive attitude; eventually, they would probably stop behaving in a friendly way toward the other person. Similarly, open triangles (*i* is friend with *j* and *k*, but *j* and *k* are not) would lead to tension within groups: for example, if one invites two friends over to their house but these two friends dislike one another, the situation might feel unpleasant for all three individuals. There might also be forms of transitivity of individuals' attitudes and beliefs regarding others: if a friend one considers to be trustworthy believes that another person is trustworthy (as signaled by their friendship toward that person), then the first person would be inclined to consider the third trustworthy as well (thus facilitating friendship formation).

In such cases, relational configurations only impact friendship formation inasmuch as individuals hold specific dispositions – for example, the cultural expectation that friendships should be reciprocal. Therefore, network processes act by activating or inhibiting certain latent dispositions; and they can incite individuals to select certain others as friends among the persons they have already met (mating process). This would classify them as contextual moderators rather than foci.

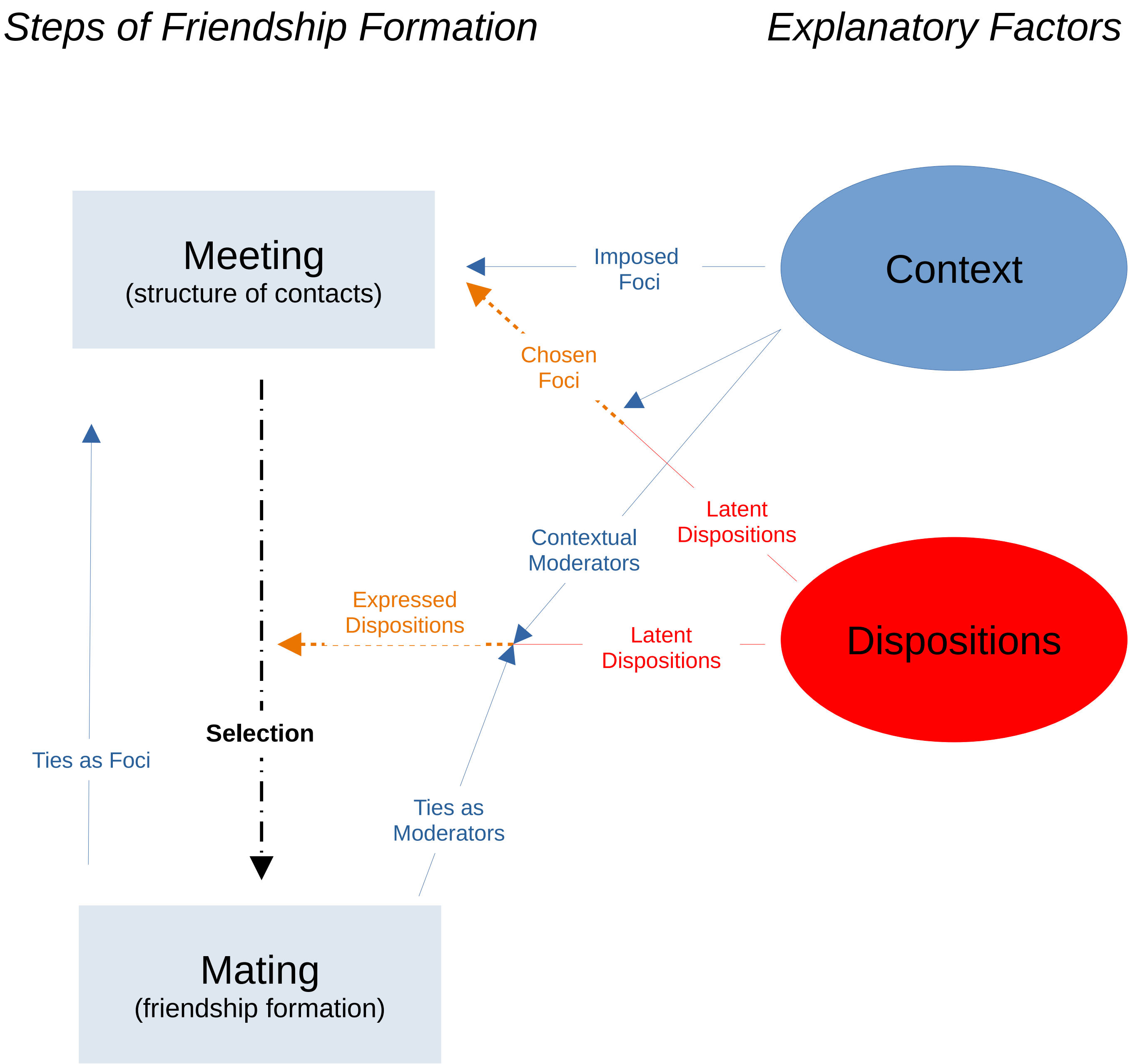
These two types of processes – balance or "relational propinquity" – are both equally plausible, and most likely operate simultaneously in many empirical contexts[31]. Nevertheless, they are extremely hard, if not impossible to disentangle from one another. Distinguishing network processes as a whole from other contextual factors, however, is possible, and is in fact one of the main advantages of statistical network models. As a matter of fact, network processes are originally a statistical, rather than theoretical category. By this, I mean that concepts such as reciprocity or triadic closure are coherent from the perspective of a statistical network model: reciprocity is the effect that captures the over-representation of reciprocated ties, and triadic closure the over-representation of closed triangles (compared to a baseline model where ties are distributed randomly). Therefore, for methodological reasons, I consider it preferable to group both types of explanations (ties as foci and ties as contextual moderators) under the category of "network processes", and to make it a specific class of processes within the proposed model of friendship formation.

1.6. New Model of Friendship Formation

It is now possible to integrate the different elements discussed thus far into a more complete model of friendship formation. This is illustrated in Figure 2-4. The logical connections among the different elements (arrows on the figure) are now more numerous and provide a detailed typology of tie-formation factors.

31 Though I have restricted the discussion to reciprocity and transitivity here, the same applies to other network endogenous effects, such as popularity (people with many friends have more opportunities to meet more people, and given their high-status due to their large number of friends, will encounter more people who are motivated to try and become their friends).

Figure 2-4: Contextual-Dispositional Model of Friendship Formation



In this model, like in the previous one, individuals meet with other individuals, among whom they select their friends. Four classes of tie-formation processes are identified that intervene at different steps of friendship formation.

1 **Propinquity through imposed foci.** Individuals attend certain foci that induce uneven contact frequencies among them. This partly determines the pool of available candidates for friendship formation ("meeting"). Importantly, these are opportunities that are independent from individuals' dispositions; that is, foci whose attendance is out of individuals' control (at least in the short-run).

2 **Propinquity through chosen foci.** Individuals' dispositions also impact contact frequency through self-selection in certain foci (e.g. activity groups).

3 **Expressed Selective Dispositions**. Individuals' dispositions determine who becomes friends or not among the persons that have met. This selection depends on individual relational preferences (explicit, partly conscious inclination toward certain others) as well as any psychological trait that may favor or impede the emergence of certain relationships (e.g. ways of speaking, ease of communicating with others, pro-social attitudes, etc.).

4 **Network endogenous processes** include all the processes by which existing relationships impact the formation, dissolution or maintenance of other relations. They can do so by impacting meeting opportunities ("tie-as-foci", e.g. meeting new people through shared friends) or as a form of contextual moderators of individuals' dispositions (e.g. the search for balanced relationships). This can be considered a subcategory of contextual factors but is specific enough to be treated as a distinct class of processes.

In addition to these four types of social processes, two types of underlying explanatory factors can be identified, contextual and dispositional ones. *Imposed foci* is the sole category of process that is determined by a single type of factor (context). All other processes can be split between the influence of individuals' latent dispositions and of contextual moderators. This can result in several subcategories based on all relevant combinations of processes and factors: latent dispositions toward selective behaviors (e.g. racial prejudices), latent dispositions toward chosen foci (e.g. preference for a type of sport), contextual moderators of the dispositions toward selective behaviors (e.g. the local salience of racial categories), and contextual moderators of the dispositions toward chosen foci (e.g. the prestige of a sport club). However, it is important to understand that these subcategories are typically impossible to observe. Whereas foci and expressed dispositions are essentially forms of behavior (e.g. attending an activity club or behaving a certain way toward peers), contextual moderators and latent dispositions are theoretical *explanans* of these behaviors.

The most crucial consequence of this is that *latent dispositions and contextual moderators are only meaningful when simultaneously considering different individuals and different contexts*. As Lahire points out, dispositions are never directly observed, being essentially hidden within individuals' minds; their existence can only be inferred from observed behaviors. Moreover, these behaviors are never pure expressions of one's dispositions, but always occur from their encounter with a specific context. This raises the issue of whether the concept of disposition has any empirical use whatsoever. "*Indeed, if an actor's disposition is only visible during the time of action, through multiple behaviors, then one could wonder whether the distinction between "dispositions" and "behaviors" is truly useful*" (Lahire 2005, p.99; I translate). Lahire's answer is that dispositions are needed to account for the regularity of certain behaviors across multiple contexts: one may be considered to "be" shy – as opposed to "behaving" shy – if she behaves in a consistently shy way in several, unrelated situations. The same is therefore true of contextual moderators, which are only meaningful inasmuch as a given individual (i.e. fixed dispositions) behaves differently across contexts. Consequently, whereas *expressed dispositions* can be defined from a single situation (X behaves in a certain way), *latent dispositions* need several separate situations to be inferred. This has important methodological implications, to which we shall come back to below.

1.7. Time, Socialization and Context Transformation

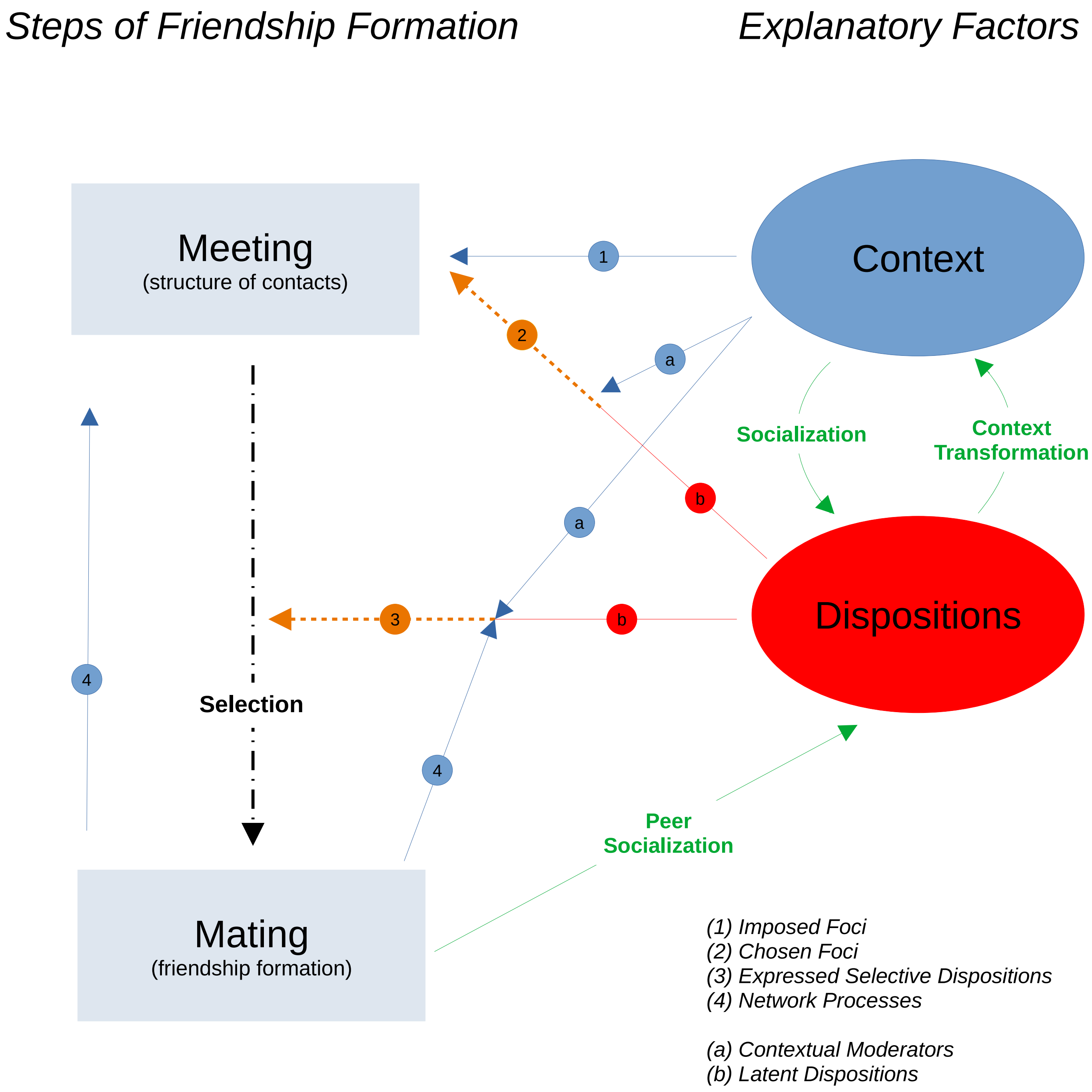
Among the four problems raised in section 1.2., there is one that has yet to be addressed: individuals become socialized within the contexts they attend. Additionally, a reciprocal problem is that of the context being transformed by the behavior of individuals.

These two problems bring us to an essential aspect of Lahire's theory. Over time, retro-active loops imply that contexts and dispositions are effectively consequences of each other, as contexts emerge from disposition-driven behaviors and as dispositions form within contexts, eventually blurring the distinction between the two. However, the key assumption that grounds the entire theory is that these processes of mutual influence are (in most cases) relatively slow. Humans are characterized by a certain amount of psychological rigidity, such that they do not adjust instantly to any change in their environment. Conversely, most social contexts are outside the immediate grasp of any one individual. As a consequence, both context and dispositions can usually be treated as approximately constant *within a certain time window* (the length of which depends on the considered case and of the researcher's interests). For example, in the case of the present study, it is clear that the organization of the school system, or students' place of residence, is hardly impacted by the dispositions that they may acquire when interacting with one another. Similarly, even though horizontal socialization among students can occur, it seems reasonable to assume that it does not over-write the entirety of their past family socialization.

To be clear, it is still possible to investigate processes of self-selection and of socialization; but the time frame considered will generally be different from that of other relational processes. This leads me to treat socialization and context change differently from other processes in the proposed model of friendship formation. The model presented in Figure 2-4 can be understood as a “static” or “short-term” model. There is still a notion of time in some processes, particularly network endogenous ones, but they can operate while context and dispositions remain essentially constant. Consequently, this static model treats context and dispositions as being exogenous, as shown by the fact that no arrow points toward the corresponding boxes. This may or may not be a reasonable assumption depending on the considered case. In any case, from a methodological point of view, all the elements of Figure 2-4 could be investigated through cross-sectional models (with network processes being reconstituted from observed tie configurations).

However, it is also possible to extend the model further by integrating additional processes that can unveil over longer periods of time, as shown in Figure 2-5. In this extended model, context and dispositions are not exogenous anymore, but are subject to feedback processes from other elements of the model.

Figure 2-5: Extended Model Integrating Feedback Processes



Three further processes are therefore added:

1 **Peer Socialization** is an instance of socialization that specifically occurs through inter-personal relationships. It comes close to the concept of "influence" often used in longitudinal network modeling, which refers to the propensity of individuals to

become similar to their friends over time, though with some differences that we will come back to in section 3.

2 **Socialization** refers to the changing of an individual's dispositions following his exposure to a certain context. Note that peer socialization is not included here, by convention (a more precise name might be "socialization from aggregated properties of the context").

3 **Context transformation** refers to the fact that individuals can modify the context they are part of through their behavior, itself driven by their dispositions.

The distinction between *chosen foci* and *context transformation* may sometimes prove tricky. To give an image, chosen foci pertain to the orientation of individuals within a somehow stable landscape, when context transformation pertains to a modification of the landscape itself. In that sense, self-selecting into a foci need not change the aggregate properties of that foci. For example, an individual may adhere to a club with little to no impact to the club's functioning or overall atmosphere. This self-selection would thus primarily impact the sociability of the considered individual; club-mates would only be impacted in as much as they form relationships to that individual. By contrast, an individual trying to change the rules of a club would have repercussions for the sociability of all club members, beyond their ties to the considered person. This is a distinction of degree more than of nature, as an individual who self-selects into a context may have various implications as per the general properties of this context; but I still expect the distinction to prove useful in many empirical cases.

2. Application to Socioeconomic Homophily among Adolescents

Having formally introduced a model of friendship formation in the previous section, we now turn toward applying it to the case of socioeconomic homophily among teenagers. For each of the classes in the typology, I consider the specific processes that may impact homophilic patterns among middle-school students, and review the empirical evidence found in the literature. First are imposed and chosen foci (2.1); followed by expressed selective dispositions (2.2); then network effects (2.3); and finally influence, socialization and context transformation (2.4).

2.1. Imposed and Chosen Foci

2.1.1. Socially Segregated Foci

First, regarding imposed foci, residential segregation and differential school choices by parents tend to produce highly segregated meeting opportunities among socioeconomic groups (Merle 2012; Oberti et al. 2012). Moreover, even within a given school, same-background students should on average live closer to one another. Indeed, social mixing at school typically occurs because of a catchment area that covers several neighborhoods with different social compositions, rather than there simply being no residential segregation whatsoever (*ibid*). Previous network studies found residential homophily among secondary-school students while controlling for socioeconomic background, ethnicity and network effects, which strongly suggests that geographic propinquity has a direct impact on friendship formation (Kruse et al. 2016).

Second, at school, students can be sorted among different tracks or ability groups. These orientations are strongly impacted by social background, even once academic performance is accounted for (Palheta 2012). Furthermore, even without formal tracking, certain mixed schools choose to create unofficial “good classrooms”. This can be a way to guarantee upper-class families a form of protection or special attention for their children, so as to incite them to stay in the local school rather than opting for private schooling (Delvaux and Van Zanten 2006).

Parents exert a certain amount of control over their children’s sociability, which, from the perspective of the child, constitutes an imposed feature of the context (though they may have ways to escape this control, and increasingly so with age). One prominent way in which they do so is through inter-generational closure (Coleman and Hoffer 1987). Friendship between parents facilitate contact among children, as well as allowing parents to extend their control over their child outside of the home. For example, certain parents will only allow their child to go to a friend’s place if they know that friend’s parents. Moreover, this also applies to cases where parents initially met through their children. While it is true that parental friendships are partly endogenous to children’s social activity, it is not enough that two children are friends for their parents to become so as well – they also need to get along themselves. That a friendship between children would “transfer” to their parents can then provide opportunities for this initial friendship to be deepened (e.g. organizing shared vacations). Given the socioeconomic segregation of adult friendships (McPherson et al.

2001), inter-generational closure should on average favor socially homophilic friendships among students.

Moreover, parents are likely to be more sensitive to social status and hierarchies than their children are. Some upper-class parents have been found to worry about their child's exposure to social diversity – which they fear might lead to decreased academic performance or "bad" behavior in general (Ball et al. 2004; van Zanten 2015). Although such attitudes should by definition be under-represented in mixed contexts (as the more "segregative" parents would simply avoid these altogether), we may still expect certain parents to hold them, which may result in strategic efforts to orient their children's friendship choices toward socially similar peers.

Moving on to chosen foci, we can also expect those to depend on students' social background. Most youths take part in out-of-school activities, frequently in clubs or associations. Since teenagers from different backgrounds have different cultural and leisure tastes (Octobre et al. 2010), the resulting meeting opportunities ought to be socially segregated. Furthermore, these activities are also partially under parental control. Van Zanten (2015) found upper-class parents (particularly managers from the private sector) to exert what she calls a "surrounding strategy": by pushing their children toward socially homogeneous environments, they indirectly favor socially homophilic friendships. Though sheer forbiddance or obligation would fall into the category of imposed foci, parents may also use more subtle forms of influence over their child's choices. In that sense, and from the perspective of children, there can be a somewhat blurry continuum between "chosen" and "imposed" activities.

2.1.2. Contact Opportunities and Multidimensional Homophily

When considering the impact of contact opportunities on socioeconomic homophily, one should not only look at the distribution of socioeconomic backgrounds across the considered foci, but also to its joint distribution with other relevant attributes. Indeed, as Blau (1994) pointed out, individuals do not generally consider a single attribute when performing homophilic selection; since attributes essentially come as a "package" within individuals, friendship formation is ultimately determined by joint similarity across several characteristics. This is sometimes known as "multidimensional homophily" (Block and Grund 2014), although in the present framework it should be called "multidimensional selection" instead. In cases where the pool of available partners does not include all possible combinations of the relevant attributes, it might prove impossible to satisfy all of one's relational inclinations at once, i.e. finding the "perfect" friend. This may force individuals to arbitrate between their selective dispositions, effectively repressing some of them to satisfy others. Based on this, McFarland et al. (2014) predict lower homophily in small and/or homogeneous population, where there are too few available "packages" of dispositions among the potential partners for individuals to act upon most of their selective dispositions.

2.2. Expressed Selective Dispositions

2.2.1. Latent Dispositions and Selection Processes

What type of disposition-driven selection could induce socioeconomic homophily among teenagers? To begin with, it should be kept in mind that selection processes are not necessarily driven by a search for similarity. For example, some attributes tend to be sought for by many persons (like physical attractiveness in romantic relationships), in which case selection does not imply homophily. I will talk of *homophilic selection* to refer to selection processes that are based on a search for similarity. In the literature, this is sometimes known as *assortative mixing* as well (Goodreau et al. 2009). By extension, I will sometimes also talk of *homophilic dispositions* to refer to the dispositions that underlie homophilic selection, and of *homophilic behavior* to refer to the manifestation of these dispositions.

The simplest case would be that of a direct, explicit preference for same-background friends – that is, discrimination in the usual sense. For instance, Eder (1985) conducted an ethnographic study in a US secondary school in the 1980s', where she found socioeconomic background to be a strong determinant of indigenous status. Wealthy students openly mocked "grits" (slang for destitute persons) and avoided socializing with them. However, more recent ethnographic studies in French schools did not report such a clear element of socioeconomic discrimination, suggesting this largely varies according to time and place (e.g. Cousin et Felouzis, 2002 ; Pasquier, 2005 ; van Zanten, 2012)[32].

The preference for same-background peers does not necessarily imply conscious discrimination. Socioeconomic origin entails a large number of derivative properties, such that it can be mediated by similarity in tastes, beliefs, or ways of speaking and behaving

32 Assuming that something does not exist based on the fact that ethnographers do not mention it is a risk, at least in cases where they do not clearly state that they looked for said thing (which is not the case here: socioeconomic discrimination is simply not mentioned). However, socioeconomic origin and school segregation are important themes for these authors, so we may expect that they would have reported cases of blunt, salient socioeconomic discrimination at the very least, even if it was not initially part of their research questions.

(Bourdieu [1979] 2016). In that regard, the correlation between adolescents' social background and their cultural and leisure tastes (Octobre et al. 2010) implies that taste-based selection of friends should, on average, entail socioeconomic homophily. Furthermore, the same is true of other types of dispositions which, although less conscious or purposive than the concept of "preference" generally implies, impact friendship selection. Archetypal examples are ways of speaking and body-language. Social and ethnic groups develop specific forms of language (Labov 1993) and gesture (Efron 1941) (both cited in Lahire 2005, pp.110-111). Similarity on these might favor positive attitudes and likability among individuals, even without them realizing it.

This extends to other attributes which, though conceptually distinct from socioeconomic background, are nevertheless correlated to it. In such cases, homophilic selection on these attributes will entail socioeconomic homophily through a compositional effect. For example, we have much evidence for selection of same-ethnicity friends among teenagers (Hallinan and Teixeira 1987; Moody 2001); and, at school, students also tend to select friends with similar results and attitudes (Flashman 2012). Both ethnicity and grades are correlated to socioeconomic background, meaning they should act as inducers of socioeconomic homophily (Duru-Bellat 2015; Ichou 2018)[33].

Therefore, the level of homophily observed in a given context depends on its dispositional pool; that is, the types and frequencies of dispositions in the population of interest. For this reason, particular attention should be paid to selection biases in this population[34]. It is important to differentiate this from the effect of contextual moderators: selection biases do not act by activating or inhibiting certain dispositions, but by affecting the types of dispositions found in a population[35]. For adolescents, different schools may attract

33 Certain authors refer to this as a "by-product" homophily (e.g. Smith, Maas et van Tubergen, 2014). The term suggests that this homophily is somehow fortuitous, being an unintended side-effect of another form of homophily. However, in a framework such as that of Bourdieu ([1979] 2016), mediation through cultural tastes or academic results is essentially instrumental to socioeconomic distinction, at least at the macro-level (the very reason that artistic tastes matter so much is because they act as a sign of one's social position). The difference between saying that "academic homophily *mediates* socioeconomic homophily" and that "socioeconomic homophily is a *by-product of* academic homophily" thus pertains to the causal relation that one is willing to postulate at the macro-level (in this case, whether socioeconomic background is the prime determinant of academic performance). At the level of a single school, though, both have essentially similar implications: X is correlated to socioeconomic background, thus a selection of friends based on X will induce socioeconomic homophily – regardless of the exact causal link between X and socioeconomic background.

34 Note that the term "selection bias" is unrelated to the notion of selection in the model of friendship formation. It is used in the usual sense in statistics, i.e. referring to the mechanisms through which a particular population has been constituted.

35 In a sense, one may say that selection biases are features of the context (e.g. the pedagogical orientation of a school, which attracts specific families). Nevertheless, in our current framework, these selection effects are not directly part of the friendship formation model: they pertain to the social processes that led to the current

different type of families, resulting in unobserved heterogeneity across schools as per individuals' dispositions. In particular, the fact that certain upper-class families try to avoid diverse schools incidentally implies a selection bias for the population of these schools: the families most hostile to socioeconomic mixing – and thus potentially the students with the most "segregative" dispositions – should be under-represented in diverse contexts. Therefore, while school avoidance certainly contributes to homophily at the macro-level through inter-school segregation, it might also (to an extent) favor heterophily locally, by selecting for pro-mixing families within diverse schools.

2.2.2. Contextual Moderators

Dispositions are never expressed in a void, but always through their encounter with a particular context. In that regard, contextual moderators can activate, or on the contrary inhibit, the expression of students' dispositions. McFarland et al. (2014) found the size, tracking system, composition and emotional climate of US secondary schools to be associated with differing amounts of socioeconomic and ethnic homophily (which they interpreted, perhaps too quickly, as reflecting students' propensity for socioeconomic and ethnic selection). According to them, large schools would offer a broader choice of partners; non-tracked systems would make individual traits more important (as opposed to tracking affiliation); heterogeneous population (in ethnic and social terms) would highlight dissimilarity; and negative climates (i.e. tense or adverse settings) would favor selection based on ascribed characteristics; all of these resulting in higher levels of socioeconomic and ethnic homophily. Though I have reservations about the article's methodology and results[36], the broader argument that school settings impact selection behaviors is an important one.

In the school context, one important moderator may be the salience of academic attributes. Homophilic selection among adolescents based on academic results and attitudes is likely enhanced by teachers' effort to label and rank students based on their performances

state of dispositions in the population, which are treated as exogenous factors (the boxes "context" and "dispositions" in Figure 2-4 could of course be taken as phenomena to be explained as well, but this shifts the focus away from the settings where friendship formation is considered).

36 Apart from the fact that they are somehow quick to conflate observed relational patterns with individuals' preferences (e.g. centralized networks indicate a preference for hierarchy, homophily indicates a preference for similar others), some measurement choices are also debatable (e.g. schools' emotional climate is measured through average GPA scores, which seems to suggest that "good" schools in academic terms are also necessarily "nicer" from an emotional point of view). Furthermore, the entire argument of the article is that the individual-level psychological traits most relevant to friendship formation, which they identify as being "balance", "dominance" and "choice homophily", are roughly similar across all US students, which in turn implies that any difference across schools should be attributed to contextual moderators. This allows them to disregard any potential selection bias that would result in different schools having different populations, and thus different pool of latent dispositions.

(Sirota 1988). It is also clear that the conflicts that emerge among students concerning legitimate classroom order (e.g. while one student wants to talk in class, the other wants to follow the lesson) are a direct consequence of the classroom settings themselves (in a small room with many students and a single teacher, it is necessary for all students to be quiet for any of them to be able to follow the course). As mentioned before, academic selection may then induce socioeconomic homophily.

However, school settings might also provide moderators that inhibit socioeconomic selection. In research conducted in a French middle-school during my masters degree, I found “good” and “bad” students to sometimes express a sense of solidarity toward one another, especially in front of teachers that were judged as unfair. Students also seemed to hold a kind of egalitarian ethos, commonly reasserting during interviews that they were all comrades or “all the same” (this generic discourse not preventing, of course, numerous animosities and hierarchies in practice) (Chabot 2017). This identification as fellow students or classroom members, as well as shared experiences in the school context (e.g. a particular teacher, exam, test, activity, school trip, etc), might provide ground for bonding among academically and socially dissimilar students, thus toning down socioeconomic selection altogether.

Supporting this idea, Block et Grund (2014) argued that there is a decreasing marginal impact of multidimensional similarity: similarity on some characteristics might make dissimilarity on other ones less important. Therefore, if students from different socioeconomic backgrounds share numerous other traits, socioeconomic dissimilarity might become marginally irrelevant to friendship formation. The impact of the school context on socioeconomic homophily may therefore be more complex than a simple exacerbation of socioeconomic distances, and it may vary significantly depending on the exact pedagogical configuration found in a given school.

2.3 Network Endogenous Effects

The two main network processes consistently found in friendship networks are reciprocity and triadic closure (Block 2015). While neither of these induce homophily on its own, they can aggravate existing homophilic tendencies. Homophily aggravation is a well-known effect of triadic closure. Goodreau, Kitts et Morris (2009) found the estimates for ethnic homophily in US secondary schools to be reduced by 5% to 15% when controlling for triadic closure (using Exponential Random Graph Models, a class of models that I will introduce in chapters 4 and 5). According to them, this is because triadic closure in an already homophilic network has a higher chance of connecting similar rather than dissimilar nodes: "*if two ties A-B and B-C are both within an attribute category, the tie created by triad closure A-C will also be within category"* (*ibid.*, p.4). More than a social property of friendship ties, this seems to be a formal implication of network structures: stylized simulations show that increased triadic closure consistently results in high levels of homophily for even modest levels of baseline homophily, at least for simple models (Asikainen et al. 2020; Foster et al. 2011).

However, on the side of social processes, Block (2018) has argued that there is usually a negative interaction effect between triadic closure and propinquity processes associated with similarity on a given characteristic. The idea is that there are decreasing returns to contact opportunities: two students that meet regularly are much more likely to form a friendship than two students that rarely or never meet, but there is a smaller difference between meeting regularly and meeting very frequently. As a result, having shared friends should matter more, in terms of friendship formation, for dissimilar individuals. Indeed, shared friends act as a foci offering contact opportunities, and similar individuals already have such opportunities due to their shared attributes; the additional foci provided by shared friends is therefore less important to them (e.g. if two upper-class students go to the music school together, they already have an opportunity to form a tie, such that having or not having a shared friend would not make as much of a difference for them). At the level of the network, this means that transitive closure may sometimes attenuate the effect of uneven meeting opportunities and, as a result, reduce homophily[37].

To be clear, this does not directly contradict the previously mentioned argument that triadic closure fosters homophily. The formal argument was that triadic closure aggravates homophily *even when the triadic closure effect is independent from dyadic similarity* (i.e. the likelihood of a triangle closing is the same regardless of whether that triangle is homophilic or heterophilic). Block's point, on the other hand, is that, in friendship networks, the triadic closure effect should be less pronounced when forming homophilic triangles. Therefore, if this interaction has a small effect size (the probability difference between closing a homophilic or heterophilic triangle is small), it may not be enough to compensate for the

37 Block does not really talk of "transitive closure" and of "propinquity", but of "transitivity" and "homophily". However, he uses these terms in a sense slightly different from my own, so I rephrased his idea within the framework developed here. Also, note that even though his idea pertains to contact opportunities, a similar argument could be made about dispositional selection (decreasing returns of dispositional similarity, i.e. the idea that having a few common points with someone is sufficient, psychologically speaking, to become friend with that person).

structural effect of triadic closure, in which case the interaction theorized by Block would simply moderate the homophilic impact of triadic closure. On the other hand, if the interaction is large enough, it might reverse the pattern entirely, such that triadic closure would now induce heterophily at the aggregate level. Nevertheless, the empirical evidence mentioned above suggests that this is not the case: triadic closure does seem to increase homophily in students' friendship networks. Therefore, Block's hypothesized interaction is either non-existent, or simply reduces the homophilic impact of triadic closure without canceling it entirely.

So far, we have covered the main processes that will be investigated throughout the dissertation, i.e. those pertaining to the static model of Figure 2-4. Table 2-1 provides a summary of the discussion so far.

Table 2-1: Summary of the Social Processes in the Static Model of Friendship Formation

Name	Social Process	Explanatory Factors	Examples for Socioeconomic Homophily among Adolescents
1. Propinquity through imposed foci	Increased contact because of common participation to imposed foci	Structural Features of the context, outside individuals' grasp	Socially segregated schools; tracking at school; social composition of the neighborhood; parental networks
2. Propinquity through chosen foci	Increased contact because of common participation to chosen foci (self-selection based on individuals' dispositions)	Dispositions leading to self-selection into specific social foci (e.g. preference for an activity)	Sport or art clubs; public facilities (libraries, sport fields)
3. Expressed Selective Dispositions	Selection of friends with particular properties (given structural meeting opportunities)	**Latent Dispositions**. Individuals' psychological traits, inherited from past socialization.	Shared cultural tastes; easiness to communicate with similar others; discrimination and prejudices; preferences for similar others
		Contextual Moderators. Elements of the context that impact the expression of internalized dispositions (apart from structural opportunities)	School climate; emphasis put on grades and academic attitudes by teachers; opinions and comments of parents on friends
4. Network processes	Ties as chosen foci + ties as contextual moderators	Feedback loops between the network state and further friendship formation	Aggravation of homophily through triadic closure; moderation of homophily by negative interaction between selection and triadic closure

Note: the explanatory factors of chosen foci can also be split between latent dispositions and contextual moderators. Here the focus is placed on these two explanans only for expressed selective dispositions, because chosen foci will be the subject of less empirical developments in the dissertation.

2.4. Influence, Socialization and Context Transformation

2.4.1. Peer Socialization

The notion of *peer socialization* in my model largely overlaps with the concept of *influence* used in the network literature (see section 3. below). It is often presented as one of the main processes that explain homophily, together with selection (Brechwald and Prinstein 2011). For example, smokers select each other as friends, and friends influence one another into smoking. However, this does not really apply to socioeconomic homophily: social background is a fixed property of adolescents, such that it cannot depend on the state of their friendship network. Having upper-background friends does not make one upper-background as well.

Nevertheless, peer socialization may operate on the dispositions underlying homophilic selection. For example, an upper-background student that has some lower-background friends may acquire from them various cultural or psychological traits, which in turn will increase their chances of forming more friendships with lower-background peers in the future. This is a bit different from the influence usually spoken of in the network literature, since in this case influence operates on unobserved characteristics and does not induce homophily or heterophily directly, but rather, does so indirectly, through furthering homophilic selection effects. In any case, this should essentially have an aggravating effect on either homophily or heterophily: individuals would tend to reproduce the type of social contacts that they already have (someone with similar friends would acquire further dispositions toward similarity, and someone with dissimilar friends would acquire further dispositions toward dissimilarity)[38].

An interesting implication is that this may result in phenomena of path dependency: the initial state of one's personal network might determine their future relational choices, with potentially strong divergences over time from even minor differences in the initial state. For example, adolescents arriving in a new social context where they do not know anyone might, at first, pair-up almost at random, having no information about the others around them. Mutual influence might then over-write (to an extent) their previous dispositions, such that these initial friendships would be reinforced over-time, and new pairings would emerge with peers similar to the initial friends. In the case where initial affiliations were random, this should result in an absence of homophily. Of course, in real-life contexts, various factors would likely act as "pull back" forces that limit such path-dependency effects (e.g. an upper-class student that starts acquiring working-class dispositions from their friends might get

38 The picture of socialization that I propose here is arguably simplistic, notably because I assume that dispositions are simply transferred from one person to the other in a virtually identical form: my friend smokes, so I smoke as well. In practice, transmission may be imperfect, with changes in the underlying disposition, or it may even pertain to what is known as *repulsive influence* (Flache et al. 2017): a person I despise smokes, so I decided never to smoke myself. In any case, the discussion proposed here is not meant to investigate socialization processes specifically, but simply to anticipate the average impact they may have on homophily, hence this simplification.

scolded by their parents). However, in certain situations, this may explain otherwise surprising relational patterns; we shall see an example of this in chapter 8, when considering the case of a summer camp.

2.4.2. Socialization and Large-Scale Context Change

Socialization is conceptually very similar to influence, the sole difference being that it does not require direct inter-personal relationships among individuals. For example, in socially mixed schools, the overall context of the school might contribute to the socialization of individuals into pro-mixing dispositions, net of their specific friendships. If anything, this should result in a decrease of homophily over time. Nevertheless, I would generally expect influence to prevail over socialization – that is, individuals likely learn more from friends and within their close social circles than simply from the broad environment in which they evolve. Socialization is also more challenging from a methodological point of view than influence is, at least for network analysis, because the broad context surrounding students is much harder to objectify than their relationships. This type of socialization is probably best captured by qualitative methods, especially ethnographic observation, than by quantitative network modeling, which means that I will generally have less tools to address it in the present dissertation.

Finally, *context transformation* is mostly treated as null in the current study. The prime reason is that adolescents are rarely in control of their sociability contexts: schools, in particular, are administered in a top-down fashion, with students being left very little power. Certainly, there can still be certain forms of adaption of the school staff to the student population, for example by changing school regulations or pedagogical practices when these are excessively rejected by students (van Zanten 2012). Nevertheless, such processes take several years and student generations to occur, and their identification would demand an investigation of adults' practices at school. This would be outside of the scope of the present study which focuses on students. For these reasons, I will treat the school context as given, and as being roughly constant over the three years of the study.

3. Some Methodological Implications

To conclude this chapter, it is interesting to briefly consider some methodological implications of the theory discussed thus far. The most important one pertains to the feedback processes of the extended model, namely peer socialization, socialization and context transformation. As I mentioned before, it is generally necessary to use longitudinal data in order to identify these processes. This is clearly acknowledged in the literature, notably regarding selection/influence models (Brechwald and Prinstein 2011). However, another key implication is that the notion of socialization (either from peers or from the general context) is only meaningful when considering several different contexts.

I already mentioned Lahire's point that dispositions are never directly observed, but reconstituted from observed behaviors. As a consequence, the separation of dispositions and contextual moderators demands variance on at least one of the two; i.e. observing different individuals in the same context, or the same individual in different contexts. Consequently, the notion of socialization can only be distinguished from contextual moderators when considering the behavior of the same individual across separate contexts. Indeed, socialization is the idea that an element of the context generates a new disposition within someone. Thus, the only way that this latent disposition can be identified is by observing its expression outside of the context from where it was formed.

This also applies to peer influence. Consider the common finding that teenagers' friends have an influence on delinquency behaviors. If we observe, in a given school, that students start to perform acts of delinquency after having formed a friendship with delinquent peers, this does not necessarily imply that they formed new dispositions: it is possible that the influence of their peers only apply when these peers are present (e.g. pressuring them into performing certain acts). In that sense, friends act as a contextual moderator, but not really as a source of socialization, leaving the individual's inner dispositions virtually unchanged. We would say that someone has been socialized into delinquency only if they keep on performing delinquent acts in other contexts, outside of their friends' reach, which may or may not be the case. In that sense, the literature's definition of "influence" is, from the point of view of the model proposed here, a mixture of socialization and contextual moderation[39].

For the present dissertation, this is a serious limitation, given that the students from the main sample are only followed in a single context, that of their school. One solution would be to study other adolescents in a different context, with a design inspired by the methodology of natural experiments (see chapter 8). Nevertheless, it is also possible to investigate individuals' dispositions within a single school, though with less confidence than via a comparison of different contexts. Indeed, by using qualitative methods, one can study the "tool of thoughts"

39 Note that in some cases, theoretical reasoning may be enough to plausibly infer socialization, even when considering a single context. However, the implicit reference that lets us define socialization remains as one's behavior in other contexts, even if it is postulated rather than observed. Typically, when considering friends' influence on smoking behaviors, our knowledge of smoking addiction lets us assume that individuals will likely keep on smoking even when they are not with their smoker friends.

used by students when sorting their peers; that is, the perceptive categories that they use. This will be done in chapter 6. In essence, the idea is to try to look at their dispositions in a more direct fashion, i.e. not through observed behaviors (of which friendship networks are an approximation) but by analyzing their discourses in semi-directive interviews. Despite various limitations (discussed in the dedicated chapter), combining reported friendship networks with the perceptive categories identified in the interviews will provide more leverage in inferring students' latent dispositions and contextual moderators[40].

In any case, a large share of the dissertation will be dedicated to the four main processes from the static model, i.e. imposed foci, chosen foci, expressed dispositions and network processes. The separation of latent dispositions and contextual moderators, on the one hand, and the study of feedback processes (socialization and context transformation), on the other, will be explored in a more suggestive fashion, hopefully paving the way for further research in this direction.

Conclusion

The model of friendship formation proposed in this chapter will be used throughout most of the dissertation, not necessarily in any and all analyses, but as a generic thread that will connect the different chapters. Most empirical results will come with a short-discussion of the information they bring as per the different classes of processes discussed here.In the conclusion of the dissertation, these different results will be brought together in light of the theoretical model, in order to provide a relatively holistic (though still incomplete) picture of socioeconomic homophily among students.

The driving question of most investigations will be that of the respective roles, and relative importance, of contextual and dispositional factors in explaining homophily among students. To what extents are friends "chosen" based on selective affinities, and to what extent are they "found" based on contact opportunities, path dependency, or simply by chance? Are selective affinities strongly rooted into students' mind, reflecting early socialization within the family and, as such, relatively insensitive to contextual moderation? Or are they strongly dependent on their local activation or inhibition, which would open space for pro-mixing

40 One way to think about this is to treat the interviews as a source of information about different contexts (e.g. in the interviews students can also talk about out-of-school situations). We can also see them as a specific context of their own. Indeed, when answering the investigator, the student is not submitted to the same external moderators as when interacting with peers. This is often seen as a methodological issue (the questions might create answers that have no external validity) but can also be thought of as a resource (answers are not directly impacted by ordinary contexts of social interaction).

policies and highlight the decisive role of the school context in shaping homophilic patterns? To answer these questions, it will be necessary to methodically examine the different types of processes identified in this chapter.

In the rest of the thesis, chapters 5, 6 and 8 in particular will be dedicated to identifying specific processes of homophilic friendship formation. Chapter 5 relies on generative network models that are designed to find the set of low-order processes that best account for the observed structure of a network. They are therefore powerful tools for disentangling several friendship formation processes, even though the connection between the model's statistical effects and effective social processes among students is not always straightforward. Chapter 6 adopts a complementary approach, based on the qualitative analysis of semi-directive interviews with students. Whereas network models try to infer relational behaviors from observed friendship patterns, the aim here will be to investigate the perceptive categories that students use in their social interactions and that should be guiding these behaviors. In a sense, it is about finding hints regarding students' dispositions apart from their expression through relational behaviors. Finally, chapter 8 shifts the focus outside the school and studies the sociability of 60 teenagers during a summer camp. In this peculiar context, some social processes are structurally absent (imposed foci in particular), which makes it easier to narrow down the effect of other processes, notably expressed dispositions.

Chapter 3: Fieldwork, Data and Methods

In this chapter, I present the research design in detail: the fieldwork and cases chosen, the data collected, and the main methods used throughout the dissertation. I also present the four schools in which the study was conducted.

Section 1 covers the different aspects of the research design. It starts with a presentation of the different data sources that make up the empirical material (1.1), then introduces a few variables that will be of particular interest in several chapters (1.2). Finally, the methods through which the data will be analyzed are presented (1.3). Section 2 then presents the four schools where the research took place. First, their socioeconomic composition is presented and put into perspective with national data on all French schools (2.1). Then, I discuss the two main dimensions on which the schools differ from one another, namely their schooling sector (public/private) and geographic location (urban/rural) (2.2). Finally, a short overview of each school is proposed (2.3).

1. Research Design

1.1. Data Collection

1.1.1. A Cohort of Students followed throughout Middle School

The main study focuses on four different French middle schools, which I will refer to as Paris 1, Paris 2, Savoie 1 and Savoie 2. These schools have in common that they are all socially mixed, that is, tthey welcome students from working-, middle- and upper-class backgrounds in close proportions. However, they differ on many other aspects, including size, sector, location, ethnic composition and average academic level. In addition to this, a further study was conducted in a summer camp in July 2019, with no relation whatsoever to the four schools. Chapter 8 is dedicated to this auxiliary study, where it is presented in detail. Here, I will focus on the main in-school study, which will be used in all other chapters.

In each school, the entire cohort of students that started middle school in the Fall of 2017 was followed for a total of 3 years, from the end of their first year (Spring 2018, age 11-12) to the end of their fourth year (Spring 2021, age 14-15). In total, 861 students are part of the sample, although not all of them were present for the entire duration of the study, as many students left or joined the cohort across the three years (about 55 leavers and 35 joiners per year across the four schools). On the first wave of data collection, the cohort was made of 749 students (695 in wave 6).

In France, middle school lasts for four years, meaning that the study covers most of the students' time there, except for a large portion of the first year[41]. As a schooling level, middle school stands well apart from both primary and high school in the French system. Indeed, the continuity between primary and middle school is almost non-existent in public schools, with entirely different buildings, teachers, school administrations and, for the most part, classmates (in each school, on average, students used to be in the same primary school as 13% of the cohort at the start of the first year). For private schools, some students were part of the same school complex in primary school, but they make for a minority of their grade-level (see section 2. below). As for high school, the split is also clear: middle school ends with a national exam (*Brevet des Collèges*), and the first (official) orientation process in the French system takes place at that time, separating general and vocational schooling and distributing students across several different high schools depending on their career plans and/or academic level. Therefore, middle school years make for a rather coherent "chunk" of students' life. Consequently, the present study offers valuable insights on the evolution of relationships during a complete and well-bounded stage of the life-course; but, conversely, it lacks information about transitions or discontinuities in life trajectories and their impact on relationships[42].

1.1.2. Calendar of Data Collection and Impact of the COVID-19 Crisis

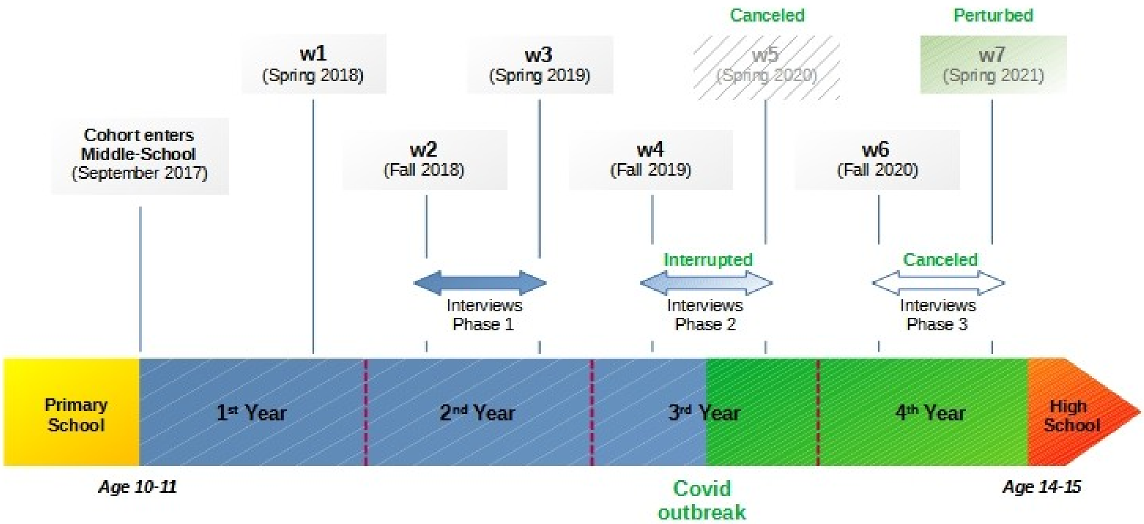
During these three years, two types of data were collected, quantitative and qualitative. Although the initial research design was meant to be rather balanced between these two, complications arose during the collection process due to the COVID-19 pandemic (see below). This particularly hindered the qualitative part of the data collection, which demanded more time and organization than the questionnaire survey. As a result, and although the study still adopts a mixed-methods approach, there is now a clear asymmetry between the quantitative data, which has been collected in a longitudinal and systematic way, and the qualitative one, which is sparser and, for the most part, cross-sectional.

41 The first wave of data collection (Spring 2018) took place around April/May; the school year runs from early September to the end of June. Therefore, most of this first year (September to April) is not covered by the study, but a small part of it is (April/May to end of June).

42 One exception may be the case of students that arrived in the schools in the course of middle school, and that were included in the sample at the time. However, I did not conduct targeted analyses at this subgroup yet – though the data would allow for it.

The quantitative data is made of questionnaires that all students of the cohort (i.e. the entire grade-level) were asked to complete every six months. In total, seven waves of questionnaires were planned, from the Spring of 2018 to the Spring of 2021 (Figure 3-1). However, the outbreak of the COVID pandemic in early 2020, and the various lock-down episodes and sanitary restrictions that have been repeatedly implemented in France ever since affected this process. Wave 5 (Spring 2020) was canceled entirely, as it was scheduled during the first lock-down following the epidemic outbreak. Moreover, wave 7 was affected as well, though less drastically: in two out of the four schools, the data collection was completed online, instead of the usual in-person collection, which clearly impacted the quality of the answers (higher rates of non-response). In the other two schools, questionnaires were done in person, but later in the year than usual (June instead of April/May), and with a lower response rate as well (some students did not come back to school immediately after their re-opening). By contrast, wave 6 took place somewhat normally, as it fell in-between two periods of school closing (Fall of 2020)[43]. Nevertheless, it should be clear that the pandemic and lock-down had an impact on students' social lives, meaning that the relational patterns observed in wave 6 may still differ from what would usually be observed among students of that age (or what would have been observed without the epidemic). Therefore, waves 1 to 4 can be thought of as taking place under "normal" settings, while waves 6 and 7 were rather exceptional.

Figure 3-1: Timeline of Data Collection



43 In France, schools closed during the first epidemic wave for approximately three months, from March to June 2020 (re-opening took place gradually over a month between May and June). They closed a second time for about a month during the third epidemic wave, in April and May 2021. Note that classrooms could be closed outside of these periods when a cluster was detected.

In the rest of the dissertation, quantitative analyses will rely on waves 1 to 6. Indeed, the perturbations of wave 7 were important enough that comparing it to other waves would demand particular precautions and a dedicated exploration of the biases at work. Moreover, it was collected near the end of the PhD, leaving little time to integrate it into the analyses. Altogether then, it seemed more prudent to leave it aside, especially given that the other waves proved to be enough regarding my research questions[44]. Wave 6, on the other hand, is particularly important: it almost doubles the total time span of the study compared to stopping at wave 4 (two years and a half instead of one year and a half), and lets the data cover all four years of middle school (from the last trimester of the 1st year to the first trimester of the 4th year). It was also collected under (relatively) regular conditions, so its inclusion in the analysis is justified. Of course, since it still took place in a very special period, particular care will be taken in interpreting results from this wave.

With regard to the interviews, 50 were conducted in total, including 38 in phase 1 (i.e. between waves 2 and 3 of the questionnaires). They concerned students from three out of the four schools (all but Paris 2, for logistic reasons). Follow-up interviews with the same students were planned between waves 4 and 5 (Phase 2), then again 6 and 7 (Phase 3), as well as first interviews with Paris 2's students. However, the pandemic started in the course of phase 2, such that only 12 interviews could be conducted, most of them in a single school (Paris 1, 8 interviews out of the 12). After this, the sanitary restrictions in the schools and the two lock-downs made it too difficult to resume interviews, leading to phase 3 being entirely canceled. Therefore, most qualitative analyses will rely on phase 1 interviews, although a few phase 2 interviews will be used marginally as well.

1.1.3. Questionnaire Items and Response Rates

Questionnaires were made of one permanent and two swing modules. The permanent module was asked at each wave, whereas the swing modules consisted of two different sets of items that were integrated alternatively in one wave out of two (module 1 for waves 1, 3 and 7, module 2 for waves 2, 4 and 6[45]). The permanent module included the name generators used to construct friendship networks (see below), as well as questions about popularity,

44 The integration of wave 7 would have been decisive if the previous waves hinted at certain time trends that needed to be confirmed, and/or if important results were too frail by lack of statistical power. However, it turned out that most features of interest were relatively stable over time between waves 1 and 6; and those that experienced a particular trend between waves 1 and 6 did so in a strong enough way for the trend to be clearly apparent without the need for an additional wave. See chapters 4 and 5.

45 A few questions from module 1 were integrated into the wave 6 questionnaire, to compensate for the canceling of wave 5.

bullying and disliking – in other words, in-school relational items. The swing modules comprised questions pertaining to individual attributes, notably cultural tastes, leisure practices, social media use and personality traits, as well as name generators for out-of-school relationships (e.g. naming the friends that one has already invited home at least once).

In addition, a few items pertained to time-invariant characteristics, namely: number of siblings, parental occupation, ethnic background, place of residence and former primary school. Most of these were asked in wave 1 (or on the first wave of entering the sample for students joining the study later on), but questions about parental occupation and place of residence were asked a second time in wave 4, in order to get more robust answers (students came to know more about their household situation as they grew older) and to account for potential changes in household situations. Moreover, for these items, I would check on the answers that were missing or imprecise after each wave, in order to specifically target the concerned students at the next wave and ask them for further details orally. For parental occupation in particular, this proved essential in order to improve the quality of answers (see below).

Response rates per school and wave are displayed in Table 3-1. These are high overall, being above 95% in more than half of the cases and above 90% in all but two cases (Savoie 1, waves 1 and 2). They tend to be a bit lower in Savoie 1, because here there were a dozen parents who refused to allow their child to participate in the study, against only one or two in the other schools. Other missing values come from students being absent when the questionnaires were passed out[46]. Note that missing value imputation can be performed rather reliably for most items, because the data is both longitudinal (answers from previous and future waves can be used for imputation at time t) and relational (the answers that others gave about student i can be used to infer what i would have said about others). Consequently, the effective sample size is somehow better captured by the proportion of students that answered questionnaires at least once (last column of Table 3-1). When using imputed values of some sort in the treatments, I will indicate it explicitly (in chapter 6 in particular).

46 I would usually organize catch-up sessions for students that missed a wave, which means that those who did not answer missed (or skipped…) the catch-up session as well. This implies that students who have disciplinary issues, are at odds with teachers and the school administration, have unstable family situations, or have recurrent health issues, tend to be over-represented in the missing answers, since they are absent more frequently.

Table 3-1: Number of students that answered the questionnaires, per school and wave

	Nb of students that answered the questionnaire / total nb of students in the grade-level					Nb of Students that answered at least once out of the six waves
	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6	
Paris 1	76 / 83 (91.6%)	81 / 82 (98.8%)	78 / 82 (95.1%)	79 / 80 (98.8%)	78/85 (91.8%)	100/104 (96.2%)
Paris 2	262 / 264 (99.2%)	254 / 259 (98.1%)	255 / 257 (99.2%)	238 / 243 (97.9%)	228/235 (97.0%)	297/300 (99.0%)
Savoie 1	158 / 183 (86.3%)	170 / 192 (88.5%)	173 / 192 (90.1%)	178 / 192 (92.7%)	175/193 (90.7%)	200/214 (93.5%)
Savoie 2	218/219 (99.5%)	200/208 (96.2%)	200/204 (98%)	192/194 (99.0%)	177/182 (97.3%)	240/243 (98.8%)

1.1.4. Interview Grid

In total, 50 interviews were conducted with 44 students. These include 38 interviews (40 students) in phase 1 and 12 interviews (14 students) in phase 2 (a few interviews were conducted with two students at once). Most phase 2 interviews were done with students that had already been interrogated in phase 1 (10 out of 12). Moreover, interviews are not distributed evenly across schools: there are 21 in Paris 1 (13 and 8 in phases 1 and 2), 2 in Paris 2 (0 and 2), 11 in Savoie 1 (9 and 2) and 16 in Savoie 2 (16 and 0). The reason is that the school administrations set different rules for organizing the interviews, making it either more or less difficult to see many students efficiently[47].

The main questions pertained to students' sociability and relationships with schoolmates, such as naming one's friends, describing how they first met, or naming the people one dislikes. To a large extent, interviews were meant to investigate the discrepancy between questionnaire nominations and actual friendship practices: the idea was to get a more detailed description of the type of relationship that would then be analyzed quantitatively (for instance, whether one refers to the same persons as being their "friends" in the questionnaires and in the interview[48]). Questions also pertained to some aspects of peer relations that are

47 One school (Savoie 2) allowed students to skip classes to come to the interview, thus the high number of interviews in phase 1. In the other two (Paris 1 and Savoie 1) it was necessary to find a gap in their timetable (e.g. because a teacher was absent). As for Paris 2, negotiations with the school administration took longer, such that interviews could only start in Phase 2. Furthermore, when the epidemic broke out, phase 2 had started in Paris 1 and, to a lesser extent, Paris 2 and Savoie 1, but not yet in Savoie 2. This timing explains why most phase 2 interviews concern Paris 1 (8 out of 14).

48 As it turns out, they usually do not (more extensive nominations in the questionnaires).

difficult to capture through person-to-person nominations, such as the general mood of the classroom, social norms enforced by the group, or popularity structures. Finally, additional questions – which were sometimes not asked due to a lack of time – were meant to get individual information about the interviewed student: musical tastes, leisure activities, academic attitudes, etc. These were used to identify the most relevant distinction categories among students' tastes, which then served as a basis to construct certain questionnaire items (see 1.2.6).

1.1.5. Impact of the COVID crisis on Students' Sociability

As much as it disrupted data collection, the COVID crisis also created a new case study of great interest. Indeed, because students were surveyed through the same items before and after the lock-down and sanitary restrictions, it is possible to look at the impact that this had on their networks. Moreover, in waves 6 and 7, special items were added to the questionnaires that specifically investigated their relationships during this period (how often they were in contact with schoolmates during the lock-down, and with whom). Nevertheless, from a methodological perspective, such an investigation can prove tricky, because there is no control group that was not exposed to the treatment of interest (i.e. COVID). Yet the evolution of friendship networks between waves 4 and 6 could also be attributed to the normal aging process of students, which cannot be separated from the effects of the pandemic in the absence of a control. There are ways around this issue – notably by trying to see whether the changes observed between waves 4 and 6 are particularly sudden or unexpected compared to those observed from waves 1 to 4 – but this would demand a dedicated study. Unfortunately as it stands, I lacked the time to properly analyze this part of the data, and indeed the space to include it as a distinct chapter in the thesis. Therefore, this question of the impact of COVID is left aside from the body of the text, except as a potential bias to be wary of when drawing conclusions from the data about socioeconomic homophily in general. The interested reader may nevertheless refer to Appendix 3A, where some simple, exploratory treatments are proposed. Future works shall be dedicated to exploring this question in greater detail.

1.2. Main Variables of Interest

In this section, I go over a few variables that will be used several times in the treatments, and whose construction and theoretical foundation deserve particular attention. Note that some technical details may seem overly specific at first, but these will be important for the analyses of later chapters. Indeed, the aim of this section is to lighten the text of the empirical chapters, by beginning to introduce key measures. It is therefore possible for the reader to skim over the section here, and to come back to it later on; it is organized as a sort of catalog, where one can look for specific variables without having to read all subsections in order.

1.2.1. Relational Measures: Friendship Nominations

In order to capture students' relationships with one another, the questionnaires contained several name generators. A name generator is an item where the respondent is asked to nominate other persons, either from a closed-list or through open questions. In this case, all generators were restricted to students' grade-level, that is, to all the other students that were part of the study. These generators were of two types. For some of them, the questionnaires contained an exhaustive list of all possible nominees, and the respondent could tick boxes near their names to indicate a nomination. There was no limit to how many persons could be nominated this way. Alternatively, some generators asked students to name a maximum or 5 persons, in which case there was no list, but simply a blank space for the respondent to write the names in full letters. Moreover, some generators were repeated at each wave of questionnaire, and others one wave out of two (Table 3-2).

First, an exhaustive list presented at each wave offered three possible choices alongside each name: "very good friend", "friend" and "I dislike him/her". It was also clearly specified that respondents could choose not to check any, if they had no opinion on the person, did not know them, or considered them to be a mere classmate[49].

49 During the form-filling it appeared that, despite explicit warning (which was written on the sheet and that I would also insist on orally when presenting the questionnaires), some students tended to see the three answers as being exhaustive options: they would check "friend" for anyone that they would somehow get along with, and "dislike" for most others, including some that they in fact did not know.
This partly comes from a comprehension bias of the notion of "disliking". In French, the expression used in the questionnaires was "*Je ne l'aime pas*", which literally translates as "I do not like him/her" (rather than "I dislike him/her", which has no strict equivalent in French). The expression is therefore ambiguous: it can relate to having actual negative feelings toward someone, or simply not having positive feelings. Still, the first sense (negative feelings) is more common and, from an adult perspective, was clearly the meaning implied by the formulation in the questionnaires – even more so since I would explicitly indicate that this was the correct understanding, both on the questionnaire sheet and orally. However, some students kept interpreting it as "not being friend" even under this precision. When I asked them to correct their answers, some would even argue that the two ideas were similar (e.g. me asking "*Is it that you don't like them, or just don't know them?*" and the student answering "*Yes, I don't know them, I don't like them*"). This was more common in the first waves of data collection, while students were still young; but it also was far more common among students from the working-class. The fact that some of them would actively *resist* the classification of relationships that I asked of them leads me to believe that this was not only a mistake or a communication issue but was somehow related to the very ways in which these students conceive of friendship.
In cases where this bias was too strong – i.e. there was hardly any name left unchecked – the "dislike" nominations were treated as invalid in the analyses.

The “friend” and “very good friend” nominations are the two main dependent variables that will be used throughout the dissertation, as they define the friendship networks that will be analyzed throughout most treatments. Therefore, they can be seen as our basic measure of positive relationships among students. “Friendship” offers the advantage of being a relatively holistic concept, which encompasses various aspects of positive relationships among individuals: trust, pleasure, time spent together, etc. The main idea behind the use of this notion is that it captures the subjective importance that students attach to their relations: each respondent is effectively free to balance between the various dimensions of her relations so as to decide which ones meet the “friendship” threshold. For example, one may call someone they see only once a year a “friend”, if their emotional investment in that relation is strong enough and rooted in other important components of the concept (e.g. it is someone they trust a lot). There are, of course, some issues associated with a strictly subjective measure (see below), but this subjectivity is an intrinsic part of students’ attachments to one another and should therefore be frontally tackled.

The use of two levels of nominations, “very good friend” and “friend”, is of particular importance: strong and weak relationships may exhibit different features and follow a different logic, regarding socioeconomic homophily in particular. In that regard, many friendship nominations correspond to what may be considered relatively weak ties. In Table 3-2 below, it can be seen that the average outdegree (number of nominations emitted by a respondent) is pretty high for both types of nominations (10.9 and 19.8 respectively). This may be a declarative bias linked to the prestige and status attached to the number of friends in youth subcultures: students may tend to over-estimate the size of their personal networks because it provides them with a positive self-image. It may also pertain to the very definition of friendship among young adolescents. According to Bidart (2010), the definition of friendship as an intimate tie, based on elective affinities and restricted to a small number of persons, becomes more prevalent as individuals get older. Among children and adolescents, friendship is more frequently envisioned through the common belonging to large social groups.

However, it is important to take into account that there is some inter-individual variance in students’ conceptions of friendship, with some coming closer to a restrictive/elective conception and others to an extensive/collective one. This is shown by the high standard deviations of the outdegrees (which also captures effective differences in social integration, though): 11.6 and 14.0 for “friend” and “very good friend” nominations respectively (Table 3-2). Moreover, the median of the outdegrees is lower than the means (8 and 17). Indeed, some respondents emitted extremely high numbers of nominations – a large part of which were not reciprocated –, considerably pulling the mean upward.

In any case, and contrary to what an adult mind may expect, even “very good friend” nominations do not necessarily correspond to exceptionally strong or intimate relationships[50]. This underlines the importance of considering different types and intensities of friendship ties in order to get a more complete picture of students’ social lives. For this reason, another item pertaining to friendship was included in the questionnaires. Students were asked to nominate the friends that they spend the most time with at school, up to a maximum of five persons (hereafter referred to as “5-friend nominations”). This offers two advantages compared to “friend” and “very good friend” nominations. First, by imposing an upward limit on the number of nominations, it makes it possible to identify the strongest relationships of each student, as well as compressing the differences in the emitted number of nominations that come with students having different understandings of the notion of “friendship”. Second, by giving an objective criterion (the time spent with each friend, which is not necessarily equivalent to emotional attachment), it should also reduce the amount of inter-personal variation in the understanding of the item. Altogether then, this is meant to be a forcefully standardized measure, which is complementary with the uNo.ounded and fully subjective nominations from the previous list. In general, this upward limit on the number of emitted nominations makes network modeling a bit more difficult, and it results in a loss of information compared to unlimited nominations[51]. Therefore, this item will be primarily used for robustness checks of the results obtained from the other friendship nominations. However, on a few occasions, it will also reveal differences in friendship structures depending on the type of nominations considered.

1.2.2. Relational Measures: Out-of-School Relationships and Negative Ties

A second exhaustive list, which was only in the questionnaires in one wave out of two, was meant to investigate respondents’ out-of-school relationships with one another. Next to the name of each peer in their grade-level, they could tick up to four boxes to indicate

50 This can also be seen from the large number of “very good friend” nominations that are not repeated from one wave to the next – roughly 40% across waves and schools (see Appendix 3B).

51 In particular, a network with many unobserved ties may lead to bias in the estimation of transitivity processes: two students may be inserted into a triangle that does not appear in the network of the five closest friends, because one of the ties that forms this triangle just fell short of falling into that category.

different relationships. They could indicate if (a) they had already invited that person to their home (b) they had already been invited to that person's home (c) they had already met (intentionally) with that person outside of school and outside of their homes (including out-of-school activities, parks, public facilities, movie theaters, etc.) and (d) their parents knew the parents of that person. These measures are interesting in several regards. To an extent, they may capture unobserved heterogeneity in the strength of friendship ties, since the friends seen outside of schools are likely closer than in-school friends on average. Nevertheless, there may also be different constraints shaping students' sociability outside of school, resulting in different mechanisms of friendship formation and maintenance (e.g. residential proximity may matter more). Finally, students' declarations about their parents knowing other parents can be used to measure inter-generational closure, which may be an important element in understanding socioeconomic homophily (cf. the discussion in chapter 2).

Another question asked students to nominate up to five persons that bullied or mocked them[52]. Together with the "dislike" nominations mentioned above, these bullying nominations provide information about the tensions, conflicts or power relations that exist among students. These are a bit different conceptually from the friendship or out-of-school nominations discussed above: so-called "negative ties" do not necessarily correspond to relationships in the strongest sense. Indeed, they may rest on sparse or indirect interactions (e.g. disliking someone because of their behavior toward another person), and they are usually not stabilized by a clear set of norms and of mutual obligations that are known to both parties and are publicly signaled to others. Indeed, a friendship does not only engage two friends: their relationship is known to other students and comes with a set of expectations about how each one of them should behave, which can be evaluated and sanctioned by other peers (betraying a friend will be seen as wrong by most classmates). It is not clear to what extent this is also the case for negative ties. Consequently, typical network processes such as reciprocity, transitivity or popularity do not necessarily apply to negative ties. This implies that theories or interpretations pertaining to positive relationships should not be extrapolated to negative ones too quickly. Furthermore, it is also interesting to distinguish between disliking and bullying nominations. The first can be seen as both more demanding and more strictly relational than

52 There is no strict equivalent to the word "bullying" in French. The expression used was "*des élèves qui se moquent de toi ou qui t'embêtent*", which can be translated as "students that mock you or bother you". The French term used for bother, which is "*embêter*", has a slight childish connotation, which may make it sound less serious or important than "bullying" does in English. Nevertheless, the average number of emitted nominations was small (0.4), which suggests that students did consider relatively serious or repeated offenses in their answers.

the second, in the sense that disliking pertains to an opinion, whereas bullying or mocking implies an effective interaction between the two students.

The items discussed thus far are summarized in Table 3-2. For each of them, I indicate the average, median and standard deviation of the outdegree distribution (averaged over all waves). Additional descriptives can be found in Appendix 3B. Moreover, the questionnaires contained other relational items that will not be exploited in the present dissertation, due to a lack of time and space. These pertain to popularity nominations on the one hand, and to the contacts that students kept during the lock-down episodes of the COVID pandemic on the other. They are briefly presented in Appendix 3A (lock-down measures) and Appendix 3C (popularity measures), along with simple exploratory treatments to see how these relationships may relate to students' socioeconomic background.

Table 3-2: Main Questionnaire Relational Items

Type of nominations	Format of the question	Waves	Item	Average number of nominations	Median number of nominations	St. dev. for number of nominations
Basic in-school relationships	Tick names within exhaustive lists	All	Very Good Friend	10.9	8	11.6
			Friend	19.8	17	14.0
			I dislike him/her	6.9	4	9.4
Special in-school relationships	Name up to 5 persons	All	Friends you spend the most time with at school	4.2	5	1.1
			Persons that bully or mock you	0.4	0	0.9
Out-of-school relationships	Tick names within exhaustive lists	1, 3, 7 (module 1)	Invited him/her at home	2.6	2	2.8
			Have been to his/her home	2.6	2	2.6
			Seen outside of school and home	4.4	3	5.3
			Parents know his/her parents	3.9	3	4.4

Note: "tick name within exhaustive lists" means students were presented with a list of all possible nominees in the questionnaires, with no limit on the number of nominations. "Name up to 5 persons" means there was no list, and they had to write up to five names themselves. In both cases, nominations were restricted to peers in the grade-level.

1.2.3. Socio-Academic Scores: A Custom Measure of Socioeconomic Background

In order to measure students' socioeconomic background, I have constructed a continuous index, which I call Socio-Academic Scores (also referred to as "SAS" or "SA scores"). In short, it is a measure of the academic performance predicted by the occupation of students' parents. In other words, this is a way of ordering occupational groups based on the expected school performance of their children, which can be seen as capturing the average academic or cultural capital of occupational groups. Other measures of occupational backgrounds were tested, notably ISEI, but this custom measure became the eventual preference. I will now detail the reasons for this choice and the construction of the measurement.

In the questionnaires, students were asked about the job of their parents (or guardians) twice, in waves 1 and 4. Moreover, those whose answer was too vague in wave 1 (~25%) were specifically targeted in wave 4, and I would go to them and ask for details orally[53]. This was done again in wave 6 for the few students whose answer was still imprecise (~5%). Based on these declarations, the occupations of both parents were classified into two distinct occupational typologies: the French Socio-Professionnal Categories (PCS) at level 3 (42 categories), and the International Standard Classification of Occupations (ISCO; Ganzeboom and Treiman 1996). I then proceeded to convert each of the two categorical variables into a continuous index. Indeed, in order to study homophily through network models, continuous variables hold important advantages: occupational groups are ordinal in nature, whereas the effects that rely on categorical variables in network models assume that the groups are not ordered; and continuous variables generally allow for more parsimonious models, with less parameters needed to accommodate the same volume of information.

In the case of ISCO, this was fairly straightforward: ISCO categories were converted into the International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom, De Graaf, and Treiman 1992), which is the continuous measure typically associated to ISCO[54]. One limitation of ISCO/ISEI, however, comes from the fact that it is an international index:

53 Particular attention was paid to (a) the employment status (business owner vs waged worker) (b) the public sector/private sector distinction and (c) the hierarchical position for waged workers and employees (e.g. manager, head of a team, of an office, field supervisor, etc.). When the student was unable to give such indications, or simply did not know the name of a parent's occupation, I would still ask for concrete details that they were more likely to know (e.g. "does he/she work alone or with other people?", "does he/she always work for the same company?", "does he/she work in an office, in a factory or outside?", etc.).

54 Another alternative would have been the SIOPS index of occupational prestige (Ganzeboom and Treiman 1996), but this is a somehow more specific measure which is more rarely used in the network literature.

they are extremely useful for cross-country comparisons but, applied to a single country, they might miss some specific features of the local labor market, that are better accounted for by country-specific classifications (in the French case, this notably concerns the distinction between the public and private sectors, which is reflected in the PCS but not in ISCO). Moreover, most French national statistics use the PCS; I therefore needed a PCS-related measure to compare the socioeconomic composition of my sample to that of other French schools (see section 2.1 below).

To the best of my knowledge, there is no agreed-upon method for converting PCS into a continuous score. I therefore adapted an idea initially proposed by Rocher (2016) to create a variable well-fitted to my current objective. I used the PCS categories of the two parents to predict students' expected performance at the *Brevet des Collèges*, the national exam that takes place at the end of middle school. To do so, I resorted to a nationally representative sample of French middle school students, with information on both parents' occupations as well as individual students' results at the national exam[55]. A regression model was built, where each student's performance at this exam was predicted by the PCS of both parents (note that the regression model implied no interaction effect among the occupation of the father and mother)[56]. The predicted performance was standardized in the national sample (using sampling weights). I then used this model for out-of-sample prediction: the students from my four schools (which were *not* included in the estimation of the regression model) got attributed an expected performance at this national exam, based on their parents' PCS[57].

This variable is called Socio-Academic Scores (SAS). It should be stressed that this is a measure of occupational background, not of academic performance; it tells us nothing about students' individual grades. Rather, it is a means of ordering occupational groups along a continuous scale, based on their average school performance at the national level.

55 See Anon (2017) for a presentation of this database, called the DEPP Panel of Secondary Education.

56 There are 14 different grades in this exam (for the different disciplines of middle school). I used a Principle Component Analysis to get a factor score summarizing these 14 grades, with students' overall performance being given by their position on the first axis of the factorial space (62% of the total inertia).
In his own measure, Rocher (2016) constructs a factorial space using variables related to academic performance, but also to social background directly (parental income and education). This creates a tautology in the final regression, given that the dependent variable (the factor score) partly includes information about the independent variable (parental occupation). By contrast, I only used academic variables to construct the factor space whose first dimension is used as a dependent variable.

57 Students with a single parent, or for whom one of the parent's occupations was too vague, were attributed a score based on a separate regression model that only considers the occupation of a single parent. Moreover, only household parents were considered, i.e. parents living with the child. This is due to the key importance of cultural transmission in the variable's theoretical rationale (see below) and turned out to improve the variable's predictive power on relationships compared to a version where both parents were always included.

Out of the 861 students that were part of the study at some point between wave 1 and 6, a SA score could be attributed to 826 of them (4% missing). In my sample, the measure ranges from -1.74 to 2.43, with a mean of 0.25 and standard deviation of 0.98. Remember that the measurement is standardized in the national sample: therefore, positive scores mean that students come from an occupational background that performs better at school than the national average, and negative scores signifies that they come from a background that performs worse (the distribution is skewed to the right because the occupational groups which perform best at school are smaller and have more homogeneous results than those which perform worse). As an indication, students whose parent of reference is classified as a worker or laborer (category 6 of the PCS) have an average score of -0.96, and students whose parent of reference is classified as a professional (PCS 3) of 1.48. Thus, the typical distance between working-class and upper-class students can be evaluated at 2.44 points, which will be used as a reference for the interpretation of results in various analyses.

From a theoretical perspective, SAS have several interesting properties compared to ISEI (apart from the fact that they are derived from the PCS rather than from ISCO, which I have already mentioned). First, they imply that occupational groups are ordered based on their average academic performance, disregarding all of their other characteristics. In particular, the expected income of occupational groups is not taken into account. By construction, only their cultural capital – or, more precisely, academic capital – is considered. This is a strong theoretical commitment, but which I am quite confident in: students' socioeconomic homophily is better understood in terms of cultural capital rather than in terms of economic capital (though the two are strongly correlated in any case). As we will see in chapter 6, conspicuous consumption does not appear to be a means of social distinction among students, whereas academic results as well as cultural tastes are. More generally, the fact that students interact in a school context means that school-related elements of their family socialization are particularly salient in shaping their behaviors and perceptions of each other.

Second, most existing measures of socioeconomic background are in fact measures of parental occupation (e.g. occupational group of the head of the household, or average socioeconomic status of both parents). By contrast, SAS are built around the idea of the transmission of academic capital: what is considered is not the occupation of parents *per se*, but the expected impact that it may have on their children (as measured through academic results). This also affects how the occupation of the two parents is synthesized: the regression model used to predict academic results take the occupational group of both the father and mother as independent variables, which means that it attributes a different coefficient to the same occupational group depending on whether it is that of the father or mother. Going through the estimated coefficients (see Appendix 3D), one can see that the occupation of the mother generally has a stronger impact than that of the father, which is coherent with the idea that mothers tend to contribute more to the education of young children and to the transmission of cultural capital (Lahire 2017).

To sum up then, Socio-Academic Scores are better fitted to the French national labor market than ISEI, as they are derived from PCS rather than ISCO; and they are built around the idea of the transmission of cultural/academic capital, which seems particularly relevant to the study of homophily in a school context. On the other hand, ISEI is an established measure and is broadly used in the literature, which may give it more credit than SAS. In the end, however, both measures are strongly correlated: their Pearson's R is of 0.86 in my sample, which means that, for the most part, they offer a similar picture of the French occupational stratification. The reason SAS were eventually preferred is that their predictive power on students' relationships was slightly higher than that of ISEI, especially regarding homophilic patterns (dedicated tests are presented in Appendix 3D). Still, ISEI was used for robustness checks: the key results of the PhD were replicated by replacing SAS by ISEI (notably in the models of chapters 4 and 5). This did not affect the conclusions, though SAS generally yielded slightly larger effect sizes.

1.2.4. Academic Achievement: Grades and Competences

In Paris 1 and Savoie 1, grades were obtained from the school administration for each school subject and trimester of the four school years. For each network observation, the trimester average obtained at the closest time *before* the wave is used. In Paris 2, however, grades are only available for the first year of middle school, because the school refused to communicate them afterward (following a change in the national regulation at the time). Therefore, only grades from the first wave are used; they are treated as a constant attribute throughout all waves. Finally, in Savoie 2, students do not receive numerical grades. They are evaluated by "competences" for each school subject (e.g. knowing how to extract information from a document, or knowing how to write in a clear way), with a four-post ordinal scale for each competence (very good, satisfactory, fragile and unsatisfactory). The "grade" variable is therefore a construct of my own meant to capture students' overall academic level: I used a factor analysis on all competences, then kept the coordinates on the first axis in order to get a continuous score indicative of overall performance.

1.2.5. Ethnic Categories Built from Migratory Background

Students' ethnicity was coded by combining their first and last names and the foreign languages that they declared being able to speak in the questionnaires. These serve as proxies for the property of interest, namely the country of birth of students and of their parents, which is not known. It is therefore not possible to distinguish between first- and second-generation immigrants. The final variable aggregates alleged origins into 13 categories, that correspond to large geographic areas (e.g. Eastern Europe, South-Eastern Asia or Sub-Saharan Africa; see the complete list in Appendix 3E). In some cases, I will also use a higher level of aggregation with only 5 categories (native French, Middle East, Sub-Saharan Africa, Europe except France and other origins).

As Felouzis (2003) remarks, migratory origin differs from ethnicity, defined as a subjective construction that depends on both individuals' feeling of identity and of the labeling processes performed on them by others. In that regard, the name and languages spoken are indicative of objective migration background rather than ethnicity strictly speaking. That said, the final categories that I use come very close to the type of subjective categories that students mobilize in their relationships. In particular, the three largest categories – native, Middle East and Sub-Saharan Africa – largely overlap with the racial categories often used within peer groups ("Whites", "Arabs" and "Blacks"). Moreover, the subjective significance of one's country of origin may also be correlated to the type of declared first name (immigrant parents may choose to give a French name or one from the country of origin) and to the languages spoken by the child (e.g. whether parents speak French or not at home probably impacts the subjective construction of ethnic identity). Therefore, in practice, the measure should capture a combination of processes that pertain to both migration (objective) and ethnicity (subjective). For this reason, I will use the term "ethnic background" or "ethnic origin", which should be understood as a practical approximation.

More generally, there are many debates in the sociology of immigration about how ethnicity should be conceptualized and measured[58]. The use of first and last names as a proxy for migratory background is certainly not an ideal solution in that regard. However, given that the primary focus of this dissertation is socioeconomic homophily, with ethnic origin being mostly used as a control variable, I believe that the level of imprecision associated with my

58 See for instance the recent controversy in France around the definition of descendants of immigrants (through either patrilineal or plurilineal ascendance) in the Trajectory and Origins survey (Coulmont and Simon 2021; Mignot 2021).

measures can be deemed reasonable. A lengthier discussion of the coding rules I used, their potential biases and the reasons behind this methodological choice are in Appendix 3E.

1.2.6. Cultural and Leisure Activities

At each wave, various questions asked students about their cultural practices, including musical tastes, book-reading, internet videos, sport, and artistic activities (Appendix 3F). These questions differed between the two swing modules of the questionnaires. In module 1 (wave 1, 3 and 7), students were asked about the frequency of several activities (book reading, listening to music, doing sport…). They could then name the exact activities they did through open questions (for example, "cite 3 music artists that you like", "cite 3 video makers that you like" or "what sport do you do outside of school?"). By contrast, in module 2 (wave 2, 4 and 6), they were presented with two closed lists, one for music artists and one for video makers, and had to give their opinion about each proposed name ("I like it", "I do not like it", "I have no opinion" and "I do not know it"). These closed lists were meant to capture important elements of differentiation in adolescent tastes, notably those pertaining to differences in gender and social origin. They were based on the names that came up frequently in students' answers to module 1, as well as on the relevant cultural distinctions that students reported during the semi-directive interviews[59].

There are two ways in which these answers will be used, which correspond to different understandings of the impact of cultural tastes on friendship formation. First, it is possible to adopt a matching approach: two students are considered to be culturally close inasmuch as they share specific tastes or interests. This suggests that cultural similarity may facilitate friendship formation through shared interests, discussion topics, activities or foci. Second, it is possible to adopt a positional approach: tastes are taken as indicative of latent positions in a cultural space. This is coherent with theories of culture that see tastes as signals of one's macro-social position (Bourdieu [1979] 2016; Coulangeon 2010). In this perspective, shared tastes are not necessarily directly related to friendship; rather, they serve as proxies for unobserved attributes – most notably cultural capital – that are the ones assumed to drive friendship formation. This second approach does not care about students' having the exact same practices, but rather about them having similar ones. Similarity is thus defined in relation to an external reference, in this case the notions of "highbrow" and "lowbrow" culture.

59 This was so mostly for wave 4, as no interview was conducted before wave 2. Nevertheless, for wave 2, I also relied on interviews conducted in a previous study, during my master's thesis (Chabot 2017).

According to the first view, two students may for example become friends because they both play tennis; however, if one plays tennis and the other sails, there is no reason that this should facilitate friendship formation, as these are quite different sports. For the second view, however, both tennis and sailing are typical upper-class sports, therefore constituting a hint about students' social origin. Homophily may therefore be expected for related, but different sports, as approximating some form of cultural capital. Note that this remains true even if socioeconomic background is controlled for: first, because occupational groups do not fully capture a household's class positions (in that sense, cultural tastes may capture the heterogeneity of cultural capital and of educational degrees within occupational groups); and second, because cultural capital may be acquired by students independently from their parents (e.g. though other family members, teachers, or friends).

To accommodate these two views, I built two different types of cultural variables. The two of them are built using entirely different sets of questionnaire items. The first, called *cultural agreement*, is based on the closed lists of musicians and video makers proposed in module 2 of the questionnaires. For each pair of students, I use the square-root of the total number of shared opinions (only 'like' and 'dislike' are considered), for musicians and video makers separately, which result in two distinct covariates of cultural agreement. These covariates therefore relate to the idea of matching on tastes[60]. Note that this is a dyadic covariate, measured at the level of student pairs, not individuals.

The second variable is called *cultural status* (by analogy with socioeconomic status): students' declared tastes and practices in terms of music, book-reading and artistic activities were used to construct a continuous score that distinguishes between highbrow and lowbrow practices. Unlike cultural agreement, this is an individual variable. In the questionnaires (module 1), open questions asked students to nominate up to three music artists they liked, as well as the artistic activities they did, such as painting, dance, music, drawing etc. (either by themselves or in a club). I went through all the reported musicians to classify them in terms of musical styles (28 categories in total); then each style, as well as each declared artistic activity, was attributed a number between 1 and 5, with 1 indicating lowbrow practices and 5 highbrow practices[61]. Together with the frequency of book-reading (ordinal scale from 1 to 4,

60 Out-of-school activities were not considered for this variable, because the number of activities declared by students is too large, such that there are very few instances of exact matching (except for highly popular activities like football or drawing, but this raises other issues in terms of highly skewed distributions of the matching covariates).

61 The idea that a given musical genre may be associated to a certain socioeconomic position has been challenged by recent works, that underline the musical eclecticism of upper-class groups in particular (Coulan-

treated as a continuous score), these were then averaged in order to obtain students' individual scores. Details about the coding process and construction of the variable are in Appendix 3F.

This variable is markedly correlated to socioeconomic background, but nevertheless distinct from it (Pearson's R of 0.40 across schools). Even though it is conceptually derived from the typical cultural practices of different social groups, it is an individual level variable; meaning that a student may have cultural practices that appear to be dissonant with their occupational background. For example, a working-class student may culturally behave in a way that, from a macro-perspective, is over-represented among the upper-class.

To be clear, both measures are somehow problematic from a methodological point of view. Cultural agreement was defined from a relatively restricted set of opinions (two closed lists of music artists and video makers); and cultural status implied several arbitrary coding decisions, that relied on my own subjective perception of what is a highbrow or lowbrow genre[62]. Nevertheless, theory-wise, similarity in cultural tastes may be an important mediator of socioeconomic homophily (this is actually a view that is implicit in many works in the sociology of culture and of social inequalities, notably in Bourdieu's influential work *The Distinction*; Bourdieu [1970] 2016). Therefore, it seems important to try and consider its impact on students' relationships, even with imperfect measurements.

1.3. Methods

In this section, I briefly go over the type of methods used in the dissertation, namely Social Network Analysis (1.3.1) and mixed methods based on qualitative and quantitative data (1.3.2). Note that the exact treatments and statistical models used will be presented gradually, in the corresponding empirical chapters; this is more of a general overview, meant to lay out the basic rationale of the methods and the way in which they fit within the research design. It also introduces some key terms specific to network analysis, which will be necessary to understand future treatments.

geon 2003). In that regard, the notions of "lowbrow" and "highbrow" taste may not be appropriate to describe musical tastes in recent years. With that said, these works nuance this simple model of cultural highbrowness more than they refute it, as well as part of musical eclecticism seems to be explainable by social mobility and by the massification of higher education (Coulangeon 2010) – in which case it does not equally apply to middle school students. Pragmatically, however, the scores that I built prove to be a good predictor of friendship ties, as well as correlating to students' socioeconomic origin. In essence then, they can be seen as a decent proxy of students' cultural capital, even though they are not based on "pure" categories from a theoretical perspective.

62 To mitigate this issue, I tried to make sure that my classification was coherent with the actual socioeconomic background of the students who listened to the different musical genres, or that practiced certain activities. Moreover, I also relied on the interviews to figure out what musical distinctions were relevant from students' perspective. Still, this remains a subjective process, rather than one that would be based on fully explicit criteria.

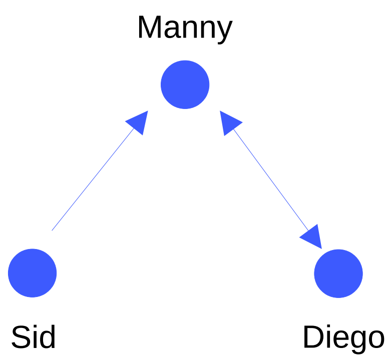
1.3.1. Social Network Analysis

Social Network Analysis (SNA) is a set of statistical methods derived from graph theory in mathematics. Its basic rationale is to treat relationships among individuals as being as important as their individual characteristics in explaining social phenomena, if not more important (Mercklé 2016).

Formally, a network is represented by a **graph**, usually in the form of a matrix called **sociomatrix** or **nomination matrix**. Each individual in the studied population is attributed a row and a column, and cells take the value 1 or 0 depending on whether a tie exists between the individuals of the corresponding row and column (the diagonal of the matrix is set to 0 when self-nominations are impossible)[63]. The network is said to be directed if some ties are not reciprocal – Claire declared being friends with Asma, but not the other way around. In this case, the matrix is asymmetrical, and ties are directed. By convention, the sender is called *i* and is placed in the row, and the receiver is called *j* and placed in the column. Thus i→j (i sends a tie toward j) does not imply j→i (j sends a tie toward i). The term “network” can also be used to refer to the visual representation given by the sociomatrix. Individuals (called **nodes** or **vertices)** are pictured by a dot, and ties (called **edges**) by a line connecting these dots. For directed networks, this can be an arrow, going from the sender to the receiver.

Figure 3-2: A Simple Sociomatrix and Network Representation

	Manny	Diego	Sid
Manny	0	1	0
Diego	1	0	0
Sid	1	0	0



Networks are said to be **complete** when all relationships of a certain type are known within a clearly bounded population. In other words, it implies that the pool of individuals whose emitted ties we know of – for example all the students that answered the questionnaires

63 For binary networks. It is also possible to have a continuous variable instead of a binary one, for example to indicate the strength of ties, but the present dissertation only focuses on binary networks where a tie either exists or does not (the strength of relationships is tackled through the use of different types of ties). This tends to be the standard practice in a large part of the SNA literature, because weighted networks (sometimes also called count networks) are harder to analyze. In addition, many relationships are better represented as a binary outcome (for example, it would be strange to ask students to rate the strength of their friendships along a continuous scale).

– is the same as the pool of individuals that could receive these emitted ties – all the students that could be nominated in the questionnaires. This is the case for the present study, as name generators were submitted to all the students of a grade-level within each school which allowed them to nominate other students exclusively from this grade-level. Importantly, this does not mean that true friendship networks are limited to this population; it is clear that there are several relationships across grade-levels in most schools. Speaking of complete networks is a methodological, not an ontological statement, although the boundaries of the studied population should try to follow existing social boundaries as much as possible (in this case, grade-levels are a sensible unit to consider, because the majority of students' friendships at school are with same-grade peers).

Let us now review some important terms in network analysis, which will be used a lot in the following chapters. **Nodes** and **edges** are somehow synonymous to individuals and ties, but they refer more directly to the statistical object of the graph, rather than to the social reality that it is meant to represent. A **dyad** is any pair of nodes and has three possible states: empty (no edge between the nodes), asymmetric ($i \rightarrow j$) or mutual ($i \leftrightarrow j$). I will also talk of **directed dyads**, indicating that (i,j) is not the same as (j,i) (even if there is no edge among them). A directed dyad has only two possible states, 1 or 0 (there is an edge or not). Similar to a dyad, a **triad** is any group of three nodes. There are 16 different types of triads depending on the state of directed edges between the nodes; the triad is said to be empty or null when it has no edge. By contrast, a **triangle** or **triplet** indicates three nodes that are connected to one another, although some of the component dyads may be asymmetrical (meaning there are different possible types of triangles, it is a generic term).

The **degree** of a node is the number of edges attached to it. The **indegree** is the number of received edges, and the **outdegree** of emitted edges. **Centrality** is a broader concept that refers to how central or influential a node is in the network. It can be operationalized in different ways: the indegree can be used as a measure of centrality, but is not the only option (e.g. betweenness centrality or eigenvector centrality; Prell 2012). Importantly, centrality is not the same as **popularity**, which is more of a polysemous term. Indeed, in the sociology of education and of adolescent sociability, it can refer to the indigenous notion of "popularity" in adolescent subcultures, or more simply to a generic notion of prestige and power among peers. On the other hand, in network analysis specifically, popularity is the social process by which nodes with high degrees tend to attract more nominations – that is, a type of "Matthew effect" where the rich get richer.

Finally, when analyzing network structures, the relevant unit of observation is often the dyad, more so than the node. This means that certain independent variables are defined as dyadic variables, not nodal ones. For example, gender is a property of individuals (male or female), but gender matching is a property of dyads (same-gender or different gender). By convention, I will refer to variables located at the level of the dyad as **covariates**, whereas those located at the level of the individual will be called **attributes** (the expression "individual covariate" may be used by certain authors, but I will avoid it for clarity). As we shall see later on, network modeling relies on covariates, not attributes *per se*, although covariates are frequently built from attributes.

1.3.2. Social Network Analysis and Semi-Directive Interviews: a Mixed-Methods Approach

As was mentioned earlier, the collection of the interviews was perturbed by the COVID crisis more so than that of questionnaires, resulting in an imbalance in the original research design. For this reason, SNA methods applied to the questionnaire data within a "classical" quantitative framework now stand as the main method used throughout the dissertation. With that said, the interviews remain a precious source of information, and they will be used as well, albeit in a smaller number of chapters (mostly chapters 6 and 7). Crucially, the point is not to conduct two parallel studies, one qualitative and the other quantitative, but to use both data sources in conjunction to answer the same research questions – that is, a mixed-methods approach.

The first important way in which the two methods interact is by using one to control, interpret or contextualize the data obtained from the other. On the one hand, the interviews were used to examine how respondents understood questionnaire items, by comparing, for the same students, the answers given in the questionnaires to those made to similar questions in the interviews. For example, this helped to identify the rather extensive interpretation made of the term "friend" in the questionnaires: students, when asked to name their friends during the interviews, generally offered far less names than they did in the questionnaires. On the other hand, the quantitative data was used to plan and conduct interviews. Respondents were selected in order to diversify relational profiles, as measured through the questionnaires of the preceding wave (having some central students, marginal ones, students with very homophilic personal networks or with very heterophilic ones, etc.)[64]. Moreover, before an interview, I

64 This was done systematically for the interviews conducted in Savoie 2 and Paris 1, but less so in Savoie 1 (the reason is that interviews were planned in advance in the first two schools after I had selected the students I wanted to talk with; in Savoie 1, on the other hand, many interviews were organized "on the fly"

would check the respondent's questionnaire data, as well as the structure of networks within their classroom. This allowed me to orient questions on peers that I knew had been declared as friends or as someone disliked in the questionnaires, or simply to follow more easily the description that the respondent made of their friendship circles. During the analysis, it also provided context for the interpretation of what the respondent declared, typically by considering their position in the school's network (cf. chapter 6).

The second approach to mixed methods that I tried to implement in the dissertation is to integrate quantitative and qualitative empirics to the same line of reasoning. Whereas the first approach aimed to create a dialog between the two sources of data, here it is about combining the conclusions that can be derived separately from both. This is done based on the theoretical framework presented in chapter 2. Assuming that friendship formation is partly driven by the expression of students' internalized dispositions, interviews may be used to examine these dispositions, given appropriate methodological precautions (of course, students do not simply say what dispositions they have when asked about it). Indeed, dispositions are, by definition, engraved within individuals, meaning that they should somehow persist throughout different contexts or situations (which is what separates them from contextual moderators). This means that it should be possible to detect them outside of their usual context of expression – in this case, during an interview. Moreover, quantitative network modeling can also be used to infer the role of expressed dispositions in friendship formation, by separating selection from other relational processes (network mechanisms, propinquity, etc.). Therefore, network modeling and interview analysis can provide insights about the same phenomena, namely the disposition-driven selection of specific friends (in that case, socioeconomically similar ones). It is then possible to look for corroboration, complementarity, or contradiction between the two methods. In the end, any account made of homophilic selection among students should be coherent with the evidence derived from both methods. Chapters 5 and 6 thus aim to study selection processes through, respectively, network modeling and interview analysis.

2. Presentation of the schools

I will now provide a general presentation of the four studied schools. The aim is to lay out the scene for subsequent analyses, by providing important elements of context that will help in the interpretation of results throughout the empirical chapters. First, the

when students had some free time – e.g. because a teacher was absent – so it was harder to select a sample of interviewed students in advance).

socioeconomic composition of the four schools is examined in light of national data on all French school (2.1.). I then discuss the main dimensions on which the four schools differ from one another (2.2.) and propose a qualitative overview of each school in turn (2.3.). Finally, I present basic descriptives on the friendship networks of the schools (2.4.).

2.1. Socioeconomic Mix in French Middle Schools

In order to get an idea of the distribution of socioeconomic diversity among French schools, I rely on a national data set from 2015[65]. Each student is attributed a continuous score of socioeconomic background – the Socio-Academic Score (SAS) introduced above – based on the occupation of the head of the household. Then, for each school, I simply compute the mean and the standard deviation of students' individual scores within that school, which provides a straightforward measure of, respectively, the overall socioeconomic "level" of the school (mean) and the amount of socioeconomic diversity within it (standard deviation).

Figure 3-3 shows the distributions of these two measures for all French middle schools in 2015 (note that the size of each school is not taken into account, each counts as a single observation). Considering all middle school students in France, SA scores have a mean of -0.10 and a standard deviation of 1.11[66]. On Figure 3-1, the part of the distribution that is equal to or greater than these values is colored in red. If students were distributed across schools independently from their socioeconomic background, the two distributions should be centered on -0.10 and 1.11, meaning that half of each graph should be colored in red. Yet one can see that this is not the case. Only 42% of schools have a mean SAS higher than the mean value among individuals, which is due to smaller schools having, on average, a higher proportion of students from lower backgrounds. More importantly, only 13% of all middle schools exhibit a standard deviation of SAS higher than that of the entire population (instead of 50% if there was no socioeconomic segregation among schools). It seems clear that French middle

65 Data comes from the Central School Base (*Base Centrale Scolarité*) of 2015. It is provided by the ministry of education and is an exhaustive census of all schools and students in the country for a given year. For each student, the occupational category of the household head is available (using the French Socio-Professional Categories at the level 2).

66 In principle, if the DEPP data (used for constructing SA scores) and the Central School Base (used for computing the scores of all French students) were both perfectly representative of the entire population of French students, the mean should be 0 and the standard deviation 1 (since SA scores, by construction, are standardized within the DEPP data). This small difference can be explained by coding or sampling biases that differ across the two datasets, as well as by the fact that the DEPP data has information on both parents whereas the Central School Base only has information about the head of the household.

schools are less mixed than they would be if students were distributed at random among them[67].

Now what about our four schools? In each graph, their value is indicated by dotted vertical lines. Regarding the mean SAS, all four exhibit values higher than the majority of French schools: 0.42 (Paris 1), 0.26 (Paris 2), 0.17 (Savoie 1) and 0.34 (Savoie 2). Thus, all of them can be considered to have a middle-upper profile – particularly Paris 1 and Savoie 2 – due to the presence of a manager or professional in a large number of the surveyed households. However, now considering the standard deviation of SAS, the four schools fall roughly in the middle of the distribution: right on its mean for Savoie 1 (0.99), a little bit under the mean for Savoie 2 (0.93), a little bit over the mean in Paris 2 (1.06) and even more so for Paris 1 (1.12). The amount of socioeconomic diversity observed in the three first schools thus appears to be typical of what can be found in most French middle schools. Paris 1, on the other hand, seems to be among the most diverse schools in the country, with only 11% of the national sample exhibiting a higher standard deviation. The reason for this is that there are relatively few middle class occupations in Paris 1, such that the socioeconomic distribution tends to be polarized between working-class and upper-class occupations. Given that the measure used is rather imprecise, though, this should be taken as a rough estimate – in particular, it may be the case that the standard deviation of the schools in the national sample tends to be over-evaluated, in which case our schools may in fact be more diverse, relative to the national population, than they appear to be here[68].

67 The measures given here are only meant to give a rough idea of social segregation among schools, in order to better understand the position of the four studied schools. They are very "naive" measures, as they do not consider the spatial distribution of social groups across France, the size of schools, the distinction between the public and private sector, etc. See Oberti, Préteceille, and Rivière (2012) or Merle (2012) for a more detailed examination of socioeconomic segregation at school.

68 The data on parental occupation in the Central School Base comes from the occupation that parents themselves declared to schools upon registration of their children, and which is then transferred to the Ministry of Education. By contrast, data for my four schools comes from me asking students about their parents' occupation. On the one hand, parents can be expected to know their own occupation better than their children do – which would mean the Central School Base data is more precise. On the other, I collected data in a thorough fashion, repeatedly asking students about further details and using various tricks to improve the coding process (cf. above), when parents' declaration of their occupation to schools is not mandatory and is completely inconsequential for them (the schools do not verify the information and it does not affect their relation with the parents), so many of them probably give imprecise answers – which would mean my own data is more precise. In any case, introducing random noise into a statistical distribution can be expected to generally increase its standard deviation, so whichever measure is noisier may have slightly inflated standard deviations. I tend to believe that this is more likely the case for the Central School Base data, which would mean that the four schools may in fact sit a bit more toward the right of the distribution of the standard deviations of SAS (i.e. they are all a bit more diverse than the French average).

Figure 3-3a: Distribution of the Mean SA Score (Socioeconomic Background) in French Middle Schools

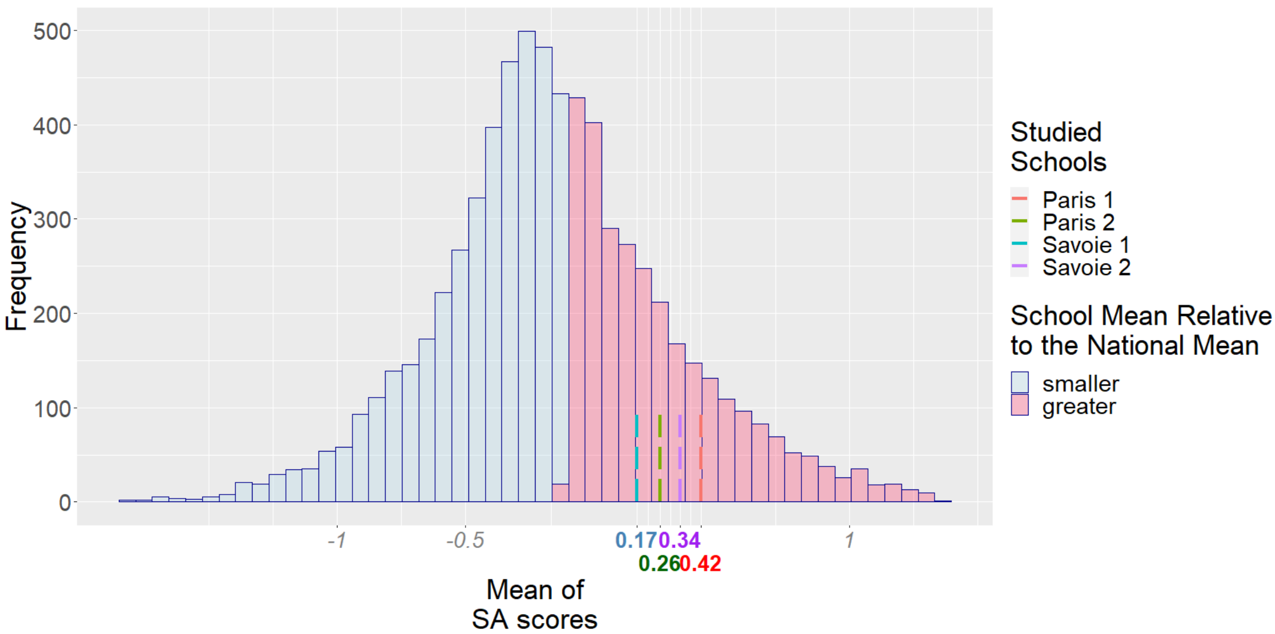
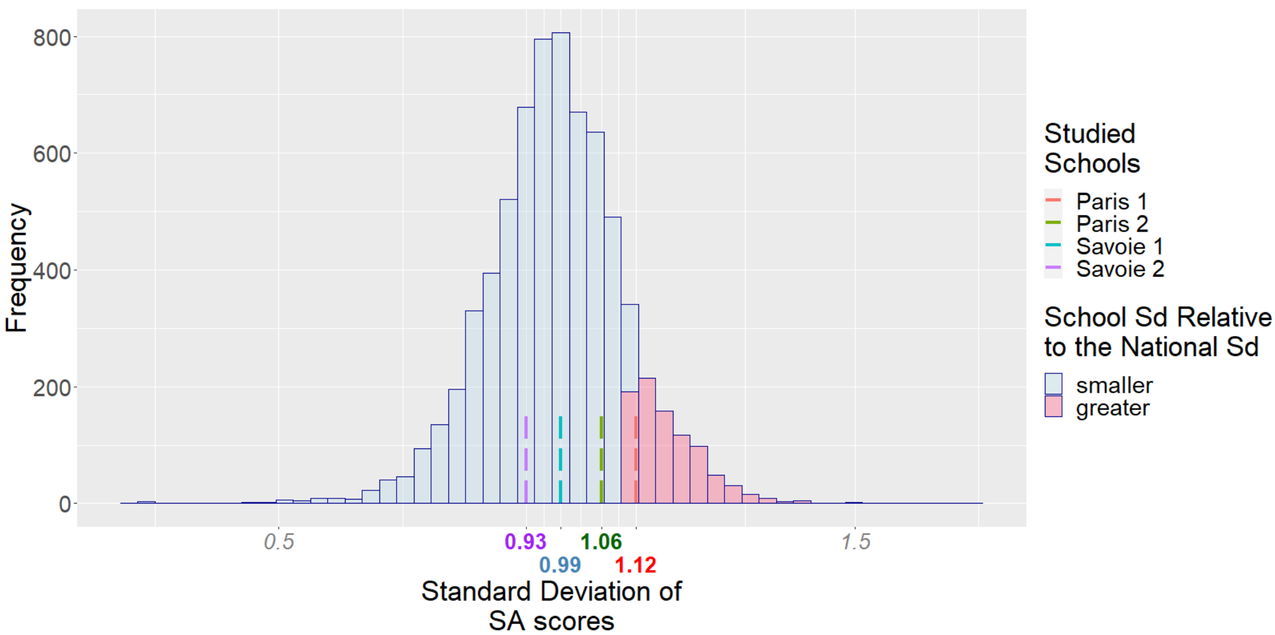


Figure 3-3b: Distribution of the Standard Deviation of SA Scores (Socioeconomic Background) in French Middle Schools



To get a better idea of the socioeconomic composition of the studied schools, Table 3-3 gives descriptives for the SAS distribution in each school. Moreover, Table 3-4 indicates the highest household occupation of students in the sample. Altogether, the four schools can be attributed a “middle-upper-mixed” profile, following the typology of Oberti et al. (2012), with a significant number of upper-class students but a reasonably high level of socioeconomic diversity, nevertheless.

It is interesting to remark that our four schools, intentionally selected to be diverse, do not appear to be exceptional in that regard at the national level, except perhaps Paris 1. It is true that the schools – with the exception, again, of Paris 1 – are less socioeconomically diverse than they should be if students were distributed completely randomly at a national level; nevertheless, they do present a substantial level of diversity, in the sense that students from very different backgrounds are present in substantial proportions. By extension then, this seems to be the case in a large number of French middle schools. Paradoxically, while the socioeconomic segregation of middle schools is largely denounced in public discourses in France, both in the academic and political fields, socioeconomic diversity in fact appears to be relatively frequent – probably more so than in most other social contexts, including primary and high schools, higher education, and the labor market. In any case, the study of socioeconomic diversity among middle school students should not been as that of an exceptional or peculiar situation, as it concerns, to various extents, a significant number of schools.

Table 3-3: Descriptives on the Distribution of SA Scores (Socioeconomic Background), per School (all waves)

	Paris 1	**Paris 2**	**Savoie 1**	**Savoie 2**
Minimum	-1.48	-1.74	-1.72	-1.72
1st Quartile	-0.44	-0.59	-0.70	-0.46
Médiane	0.33	0.08	-0.09	0.10
Moyenne	0.47	0.24	0.18	0.27
3rd Quartile	1.51	0.89	0.94	1.00
Maximum	2.43	2.43	2.14	2.43
Écart-type	1.11	0.97	0.98	0.95
Missing	7 / 104	10 / 300	19 / 214	9 / 243

Table 3-4: Number of Students per PCS of the Reference Parent (level 1), per School (all waves)

	Paris 1	Paris 2	Savoie 1	Savoie 2	Mean SAS in the sample (all schools)
PCS 1 (farmers)	0	0	2 (0.01)	3 (0.01)	0.34
PCS 2 (independent craftsmen, shop owners, company owners)	8 (0.08)	27 (0.9)	15 (0.07)	32 (0.10)	-0.06
PCS 3 (managers, executives, professionals)	33 (0.32)	76 (0.25)	35 (0.16)	58 (0.24)	1.48
PCS 4 (intermediate occupations)	12 (0.13)	57 (0.19)	58 (0.27)	54 (0.21)	0.59
PCS 5 (employees)	30 (0.28)	91 (0.30)	54 (0.25)	59 (0.27)	-0.46
PCS 6 (workers)	6 (0.06)	23 (0.08)	26 (0.12)	19 (0.09)	-0.96
PCS 8 (inactive)	2 (0.02)	9 (0.03)	3 (0.01)	2 (0.02)	-1.16
No information	13 (0.13)	17 (0.06)	21 (0.10)	16 (0.08)	-
Total	104 (1)	300 (1)	214 (1)	243 (1)	-

Note: column percentages are between parentheses. The parent of reference is the one that has custody of the child or, when both do, that has the highest PCS. When the PCS of one parent was ambiguous, the other was used.

2.2. Characteristics of the Chosen Schools: Sector and Location

The schools in the sample were selected on a criterion of socioeconomic diversity. Given this common point, I then looked for dissimilarity among the schools on two dimensions: whether the school was public or private, and whether it was in an urban or rural zone. Each of these aspects entails a number of related characteristics, regarding, in particular, selection biases in the student population; pedagogical methods and school-parent relations (public/private); and ethnicity and residential configurations (urban/rural). This means we can generally distinguish between two opposite pairs for each of those characteristics (e.g. two schools with a large immigrant population and two schools with mainly native families; or two schools where children live close to one another and two schools where a car is required for transportation). There is one school for each of the four possible combinations: urban-public (Paris 1), urban-private (Paris 2), rural-public (Savoie 1) and rural-private (Savoie 2).

2.2.1. The Public / Private Distinction

In France, two main schooling sectors coexist. The public sector, funded and directed by the State, has an obligation to welcome any student, regardless of her resources or academic level. Schooling is entirely free, and up to middle school, no selection of students is allowed (except for certain special tracks). Allocation to a school is based on the place of residence. However, it is possible for families to ask the public administration for special authorization to allow their child to attend a different school that is not the local one, if they can find an appropriate justification; most of the time, this would be because certain schools propose rare optional disciplines (e.g. Chinese language). These special authorizations are sometimes used by upper-class families to avoid schools with a bad reputation; knowing this, many headmasters intentionally open some rare disciplines so as to attract these families, proceeding *de facto* to a hidden selection (Delvaux and Van Zanten 2006; Oberti et al. 2012). Still, this selection remains relatively limited in most cases, and the social population of a school is tied to that of its surrounding neighborhood.

Private schools, on the contrary, are not bound by spatial sectors. They are free to fix their schooling fees and, when they have more demands than available places, they can choose which children to accept. This sector welcomes about 20% of middle school students in the country (Da Costa and van Zanten 2011). In the vast majority of cases, these schools are under contract with the state, which means that despite having greater autonomy, they are bound to state programs regarding the content of the instruction and have to prepare students for the same national test as public schools (*brevet des collèges*, at the end of middle school). In return, teachers are paid by the state, and the financial autonomy of the school pertains only to the material environment (buildings, pedagogical material, etc.). Nevertheless, headmasters in the private sector generally have more freedom to conceive and implement original pedagogical projects. This is also one of the selling points of private schooling, which, in a free-market logic, tries to offer parents some instructional niches, compared with a public sector often accused of being too monolithic. As a consequence, there is more pedagogical variance among private schools, from “modern”, liberal pedagogies (such as Montessori) to traditional catholic educations.

A key point to keep in mind is that private schooling is often chosen by middle- and upper-class parents to avoid the working-class environments of their local public school. At the aggregated level, private schools contribute to social segregation more than public ones (Da Costa and van Zanten 2011). Yet, since my study is on socioeconomically diverse contexts, I purposely chose private schools that welcome students from working-class backgrounds. Thus, the private schools in this study are not representative of the broader population of private schools. Indeed, they were selected for their exceptionality in that their headmasters uphold the policy of favoring socioeconomic diversity, each for their own reasons (see below). This also probably means that there are selection biases at work on the type of families that attend these schools, which is something that we will explore further in chapter 7.

Another point worth noting is that the two private schools chosen are not just middle schools, but also include a primary and a high school. In both cases, these are relatively separated from the middle school (e.g. having distinct playgrounds), and it is not necessary for students to follow the entire curricula – both the primary and the high school being smaller than the middle school anyway. Still, many students in the sample attended primary school in the same institution (30% in Paris 2 and 24% in Savoie 2), and some of them will also attend the corresponding high school. As a result, the familiarity of these students with the school might be greater than what is typically found in public schools, and the relationship between families and the school might also be stronger. By contrast, neither of the two public schools include a primary school; and, while Paris 1 does have a high school as well, only a small minority of its middle school students will eventually go there (due to the fact that the allocation of students in the transition from middle school to high school is managed at the town level for the public sector).

2.2.2. The Urban / Rural Distinction

The second key distinction between the schools is their location: two are in the Parisian agglomeration (one in Paris *intra muros*, one in the close by suburbs), and two in Savoie (Savoie 1 is located in a very small town, while Savoie 2 is in a small to medium-sized town; however, in both cases, they welcome students coming from several small villages of the region). The first crucial consequence is that the residential and spatial configurations are very different. For Paris 1 and 2, most of the students live close to the school, and get there on foot. Most also attended primary schools in the same neighborhood. Many students live in walking distance from one another and, even when they do not, there is a dense network of public transportation connecting all parts of towns. On the contrary, for Savoie 1 and 2, many students live in small villages and need to take the school bus or go car to get to school. There is, however, a notable asymmetry between students that come from villages, who generally only have a few classmates who live close to them and that attended the same primary school; and the students who live in the town in which the middle school is located. Not only do these students live in walking distance from school, but they are also the most numerous group of people from the same town – although they remain a minority in the overall student population.

Another important characteristic related to the Paris/Savoie distinction is the ethnic composition of the schools. Roughly 70% of students have no identifiable immigrant background in Savoyard schools, and 15% have a European immigrant background (figures are similar in Savoie 1 and 2). By contrast, only 35% of the students have no immigrant background in Paris 1 and 16% in Paris 2, while the students with a non-European background make up for respectively 55% and 70% of the student population. These features reflect the composition of the local population: Paris 1 and, even more so, Paris 2 are in neighborhoods with a high proportion of immigrants. Paris 2, in particular, is located in a town in the Parisian suburbs which is known at the national level for its immigrant population: it is sometimes mentioned on the news or in some (right-wing) politicians' speeches as an illustration of the issues supposedly related to immigration. Thus, beyond the effective composition of the population, there is a symbolic label attached to Parisian suburbs in general, and Paris 2's town in particular, which one may expect to have some kind of impact on students' perception of their own ethnicity and that of their peers. By comparison, the minority students of Savoie 1 and Savoie 2 are not only less numerous but are also likely living in a context where ethnicity is less politicized.

Finally, the rural/urban distinction also matters as per the exact socioeconomic composition of the schools. While the general proportion of working-class and upper-class students is comparable across the four cases, the specific types of parental occupations found differ. For upper-class groups, executives in the private sector (PCS 37) and engineers (PCS 38) are more present in Savoie 1 and 2 than in Paris 1 and 2. They represent 61% of upper-class occupations in Savoie 1 and 54% in Savoie 2, against 26% in Paris 1 and 40% in Paris 2. Conversely, teachers and scientific occupations (PCS 34) and executives from the public sector (PCS 33) are heavily represented in Parisian schools: they represent about 38% of upper-class occupations in both Paris 1 and 2, against 19% and 22% in Savoie 1 and 2. In other words, Parisian schools host more of a public/cultural upper-class, and Savoyard schools a private/economic one (though we do find some families from both types of background in all schools). For working-class occupations, industrial workers (PCS 62 and 67) are more frequent in Savoyard schools, particularly among men. They represent 26% of father occupations in Savoie 1 and 2, against 3% in Paris 1 and 2. This is because they are located in industrialized area, and many parents work as general laborers, but also as technicians and executives in large factories (especially in Savoie 2). Most working-class occupations in Paris 1 and 2 pertain to craftsmanship and construction (for men) and care work (for women), although these are also very frequent in Savoyard schools – there is not really a category of working-class occupations that is strongly over-represented in Parisian schools compared to Savoyard ones.

2.3. A Generic Description of the Schools

I now briefly present the four schools in turn. Beyond the Paris/Savoie and Public/Private aspects, for each one I try to answer the following question: why is this school socially mixed? This is by no mean obvious, as the sociological literature insists on the fact that social segregation is an important feature of the French school system (cf. chapter 1 section 1.2.1; however, as we saw in section 2.1. of this chapter, socioeconomic diversity remains relatively frequent in middle school, at least to a certain extent).

2.3.1. Paris 1

Paris 1 is a small middle school of the public sector. Each grade-level is composed of three classrooms of about 25 students each. 83 students entered the school in the fall of 2017, 76 of them answered the questionnaires of the first wave.

The school is located at the border of socially different areas. It is within Paris strictly speaking (*intra muros*), but only a few hundred meters from the administrative separation with the suburbs, where housing prices tend to be lower than in Paris itself. In this bordering neighborhood, still inside Paris, there is a large park of social housing, which provides the school with most of its working- or lower-middle class families. A few students also come from bordering suburb towns; these are also relatively modest households. Upper-class students, on the other hand, live in flats rented or bought on the private market, generally located slightly deeper within Paris. Importantly, this large housing and social diversity is concentrated in a small zone: social buildings and wealthy flats are not separated by more than a few hundred meters, and there is no obvious demarcation line or sudden change in the urban configuration that would instantly reveal the underlying social divides. Students could easily walk from one zone to the other, although it is an empirical question as to whether they actually do.

The fact that the school is socially mixed, though, does not only come from this neighborhood composition. Paris 1 has the peculiarity of offering an international option. Students choosing this option have a reinforced English program: they are expected to be fluent, attend some of their classes in English, and prepare for the International Baccalaureate at the end of high school. The recruitment procedure for this option is separated from the main recruitment of the school: students need not live in the neighborhood and are selected on their English level as well as their overall academic level (whereas there is in principle no academic selection at the entrance of middle school for regular curricula). Out of the 81 students of wave 1, 16 were part of the international program, all in the same classroom. There are also non-international students in this classroom; these students attend most classes together with the international students, except for additional courses in English for the international section – about 8 hours per week. Most of the international students are French but have lived abroad at some point in their life; a few of them are also foreigners who now live in Paris. Most of them come from different neighborhoods in Paris specifically to attend the international program. Among the students, this group of "internationals" is well identified, and they form a fairly homogeneous cluster of friends, relatively distinct from the rest of the network.

As one might expect, this international program is almost exclusively made up of upper-class students. Out of 16, the parent of reference for 14 of them is either a a manager or professional (PCS 3), while the remaining 2 hold intermediate occupations (PCS 4). By comparison, out of the 60 non-international students for which occupational background is known, 11 (~18%) have a manager or professional as their reference parent, and 31 (~52%) a working-class occupation (employees and workers, PCS 5 and 6).

2.3.2. Paris 2

Paris 2 is a large catholic school from the private sector[69]. There are eight classrooms per grade-level, each with about 33 students. 264 students entered the school in the fall of 2017, 262 of which answered the questionnaires of the first wave.

69 Note that students need not be catholic to enter the school – in fact, most of them are not, Muslims being probably the main religion (based on discussions with teachers).

Paris 2 is located in the center of a large town in the Parisian suburbs. The town is known – again, to the level of a national *cliché* – for its immigrant and working-class population. As I said before, private schooling is often used as a distinction strategy for middle and upper-class families. The peculiarity of Paris 2 is that it draws families from a relatively high social status compared to the local population; yet these families still live in a working-class neighborhood. Paradoxically, in a segregated lower-class environment, the "elitist" school happens to be the mixed one. Still, working-class households can access the school: the fees are not outside the reach of low-income families, though it is certainly a significant financial expense (several hundred euros per year). Furthermore – and although discourse should not be mistaken for practice – the headmaster of the school as well as certain teachers I talked with seemed quite attached to promoting socioeconomic diversity in the school, for ideological and religious reasons (the school belongs to a catholic educational confederation with a tradition of caring for the poor). Still, they recognize that the school is primarily sought out by wealthy families, their greater mastery of the administrative game giving them a significant advantage in the admission procedure (building a proper admission file and sending it in time). In the end, upper-class households make up about a quarter of the population (PCS 3, 76 out of 300), middle-class ones a little more than a quarter (PCS 2 and 4, 84 out of 300) and working-class ones almost 40% (PCS 5 and 6, 114 out of 300) (the rest is made of inactive or households that I have no information for, cf. Table 3-3 above).

A striking feature of the school is that occupational background is an extremely poor predictor of students' academic performance. In the other three schools, socioeconomic background accounts for 12% to 29% of the variance in grades; in Paris 2, it accounts for less than 4%[70]. Part of this is likely due to a selection bias specific to the private sector: working-class families that are willing to cut on their budget to send their child to a "good" school are probably more motivated, more ambitious or simply more confident in the child's ability to succeed. This is further confirmed by the fact that Savoie 2, the other private school, also has a lower figure compared to Paris 1 and Savoie 1 (12% of explained variance).

However, another plausible hypothesis pertains to the fact that many students are first- or second-generation immigrants. Upon arrival in France, workers that were highly qualified in their native country may experience a downward mobility, because of discrimination, lack of local networks, or no international recognition of their degrees and work experience. As such, certain immigrant parents with a high level of education occupy unskilled positions in

70 R2 of a bivariate linear regression, grades regressed on SA scores. The figures per school are 29.2% (Paris 1), 3.7% (Paris 2), 23.4% (Savoie 1) and 12.3% (Savoie 2).

France, creating a discrepancy between their classification as working-class households, and the cultural and academic capital they can transmit to their children (Ichou 2018). This explanation also fits well with the particular situation of Paris 2, a selective private school in a segregated working-class town: it might attract downgraded immigrant families who want to maintain distinctive schooling practices[71]. This remains a hypothesis, though, which cannot be confirmed nor rejected with the data at hand (in particular, I have no information about parental education, only occupation).

2.3.3. Savoie 1

Savoie 1 is a public school located in a small town in the South-East of France, in the Alps. There are seven to eight classrooms per grade level, depending on the year (the eighth was removed in the beginning of the second year), with 25 to 30 students in each one. The sampled cohort is made of 183 students on the first wave, 158 of which answered the questionnaires.

A small number of students live in that same town, but most of the students come from neighboring villages, most of which have their own primary school but no middle school. Here, socioeconomic diversity primarily comes from the fact that the recruitment sector of the school covers a wide geographical area, spanning over different residential zones, with various types of housing.

Moreover, Savoie 1 is much less embedded in its town than the other schools are. By this, I mean that students typically only come to Savoie 1, precisely, for school; most of them take the school bus or come by car with their parents. For 11 years old students, this might not be such a crucial difference, but looking at older grade-levels, the contrast is striking with Paris 1 and 2, or even Savoie 2: there, adolescents frequently hang around the schools' neighborhood after class or during their breaks, going to parks, for fast food or just walking the streets. In Savoie 1, the school surroundings are hardly a place of sociability. This does not mean that students see their friends less outside of school time – by my measures, they do so about as much as those in the other samples[72]. However, I expect that parents have greater control over their children's leisure time in Savoie 1, as out-of-school sociability is strongly dependent on access to parental transportation.

71 Ichou (2018) showed that students with an immigrant background perform slightly better at school than natives when their socioeconomic background is accounted for. He mentions downward mobility upon migration as a potential explanation.

72 To the question "who have you already seen outside of school?", students have nominated on average 3.8, 3.4, 4.2 and 4.1 peers, respectively in Paris 1, Paris 2, Savoie 1 and Savoie 2 (wave 1).

There is, however, an important exception: the students that live within the schools' town. Most of them have a working-class background, as the town has an important block of social housing. They are also more frequently from an ethnic minority; in particular, there is a Roma community, clearly identified by the teachers. For these local students, the situation feels much closer to what is observed in urban schools: they have an outdoor sociability located around the school and on the streets, which also relates to the fact that they live in flats, and not in individual houses like most of the students who come from villages. They are also the biggest group of students coming from the same primary school (that of the town), and the interviews suggest their family ties are also denser (there are more siblings and cousins among them than other students). Descriptive network measures also suggest they form a relatively cohesive group[73].

Finally, another particularity of Savoie 1 is that the school hosts a SEGPA program (*Section d'Enseignement Général et Professionnel Adapté)*: that is, a classroom dedicated to students with severe learning difficulties, or sometimes light intellectual disabilities, which effectively act as a precocious orientation toward vocational training[74]. There is a classroom fully dedicated to the SEGPA, although it has few students (8 for the first wave of data collection, 13 a year later). Most of their classes take place in a different building than the other classrooms, and their group is almost completely separated from the rest of the student body in the friendship networks. All the SEGPA students come from working-class households.

2.3.4. Savoie 2

Savoie 2 is a private catholic school located in a middle-sized town. It has seven to eight classrooms per grade depending on the year (eighth classroom deleted in the beginning of the third year), with about 30 students in each. 219 students entered the cohort on the first wave, 218 of which answered the questionnaires.

While part of the student population comes from that town, others live in the surrounding towns or villages – remember that private schools are not bound by geographic

73 In the friendship network, the subsection of the matrix made of the nodes that come from the town's primary school has a density of 0.25, compared to 0.05 for the whole matrix. In other words, 5% of all possible friendship nominations on average are realized in the school, and 25% among these students.

74 In principle, all French students attend the same track in middle school, and the general/vocational separation only begins in high school. In practice, however, students in great difficulty can be placed into so-called "atypical" classrooms, which are integrated in regular middle schools but with a different instructional content, adapted to these students. In the vast majority of cases, they are then oriented toward vocational training at the end of middle school (Palheta 2012).

sectors for recruitment. In that regard, Savoie 2 is something of an intermediary case between Paris 2 and Savoie 1 (large central geographical cluster in Paris 2, with some students coming from further away, vs a small central cluster and many students coming from further away in Savoie 1).

According to the teachers and administration, there is a strong local tradition of attendance to this school. The school has existed for several decades, and many teachers happen to be former students. Teachers also tend to keep their jobs for a long time – indeed, some of the younger teachers were even students of some of the older ones. Moreover, the middle school shares some of its buildings with a high school, and is also connected to a primary school not far away; certain students thus spend all their instructional years in this broad school community. According to the headmaster, it is also frequently the case that the parents of current students were former students themselves. Thus, the presence of working-class families in the school might in part pertain to the school's strong local history and to family traditions (probably backed up by catholic beliefs for some)[75]. One should be careful about this interpretation, however: this community aspect may tend to be exaggerated by the adult staff, who want to give a good image of their school (this was clear in the case of the headmaster that I talked with). Still, since this was mentioned by several different persons in informal discussions, it is probably at least true in part.

In any case, there is also a clear element of distinction in the choice made by some families to come to Savoie 2. Several students mentioned during an interview that they had come to this school because the public school in the town had a bad reputation – it also seems to be a widespread view among Savoie 2's students that peers from this local public school are trouble-makers or aggressive (students from the two schools meet on the school bus, or simply when hanging out in town). One teacher also told me that the public school had a large working-class and immigrant population (which is something I could not verify myself), and who believed that this was a strong reason for Savoie 2's attractiveness. In that regard, there may be an ethnic dimension to the avoidance of the public school – perhaps including native working-class households that are wary of immigrant working-class ones. There are also

75 One interesting feature of catholic private schooling in France is that the socioeconomic composition of its population, and the motivation for parents to attend, strongly depend on the regions. E.g., most catholic schools in Paris *intra muros* are elitist schools, used by upper-class families to avoid the public sector, regardless of their actual religion. In some very catholic regions, however, a significant part of the population goes to private schools, including working-class households, to the point where private schooling feels like the "normal option" locally (the proportion of middle school students attending the private sector is superior to 50% in a few departments) (these figures come from a discussion with a professional of the sector, so I do not have references to offer).

immigrant and working-class students in the school, though, whose parents were nevertheless worried of them going to a “bad” school (based on interviews with these students).

There is a waiting list for the school, as there are more families applying than can be accepted; unfortunately, I have no information as per the selection process. However, it is clear that students can be expelled from one year to the next if they are seen as having poor behavior and/or not working enough. I interviewed some students that were about to get expelled; and I also had access to the reports that students receive at the end of each trimester, where it was sometimes explicitly mentioned that the student would be expelled if they did not start working harder. This only concerns a handful of students per year, but it shows that getting expelled from this school is easier than in the public system.

A pedagogical particularity of Savoie 2 is that the grading system is different from most French middle schools. Students are not given numerical grades. Instead, teachers in each discipline have to check a list of “competences”, and a synthesis is provided at the end of the trimester on youths’ advancement on the competence list (a competence would be, for example, “knowing how to extract information from a text” or being “able to read and understand the instructions of an exercise”). While this remains a form of evaluation, it is symbolically very different from the usual French ways, in which all exercises or tests are graded from 0 to 20, and the trimester average is given to students with a precision of 2 decimals. This might make academic hierarchies and competition less salient here than in other schools, though it is clear from the interviews that students are still aware of who has good results or not. Finally, the school also had a SEGPA class, just like Savoie 1, but only in wave 1; it was suppressed the next year, with students either leaving the school or being integrated into regular classrooms.

Table 3-5 offers a summary of the features discussed thus far for each school.

Table 3-5: A Summary of Important School Features

	Paris 1	**Paris 2**	**Savoie 1**	**Savoie 2**
Sector	Public	Private	Public	Private
Location	Paris *intra muros*	Parisian agglomeration	Small town in Savoie (~ 4000 inhabitants)	Middle-sized town in Savoie (~15 000 inhabitants)
Ethnic Composition	Majority of students come from Extra-European Immigration	Majority of students come from Extra-European Immigration	Majority of students are natives	Majority of students are natives
Cohort Size (wave 1)	83	264	183	219
Occupations over-represented compared to other schools – upper-classes	Higher intellectual occupations (teachers, academics, legal occupations...)	Higher intellectual occupations (teachers, academics, legal occupations...)	Managers from the private sector and company owners	Managers from the private sector and company owners
Occupations over-represented compared to other schools – working-classes	Construction and Service care workers	Construction and Service care workers	Factory workers	Factory workers
Polarization of the Social Distribution (Presence of Middle-Classes)	Polarized (few middle-class occupations)	Intermediate	Weakly polarized (many middle-class occupations)	Intermediate
Main Hypothesized Factor of Socioeconomic Diversity	Bordering zone between wealthy and social housing neighborhoods + International Program	Private Schooling in a dominantly working-class town	Extended recruitment sector	Local tradition for catholic private schooling + Local public school has a bad reputation
Pedagogical Particularities	International Program	-	SEGPA section	SEGPA section (wave 1 only) + Non-numeric evaluation system + higher requirements toward students in terms of discipline and work (threat of not being accepted again next year)

2.4. Descriptives about Friendship Networks

To conclude this chapter, I present simple descriptives about the friendship networks of the four schools. Since these networks are the basic object that will be manipulated throughout most of the thesis, it is useful to get a sense of their general properties.

Among the various relational items that were in the questionnaires (cf. section 1.2.1. and 1.2.2. above), the ones that I will be relying on the most are the “friend” and “very good friend” nominations. The corresponding name generators were present at each wave, and there was no limit to how many persons each respondent could nominate. From there, I build two distinct networks: the “very good friend” and “friend” networks. Note that, by convention, the “friend” networks are made of both friend and very good friend nominations – this is because very good friends are a sub-category of friends, rather than a completely different type of relation[76]. Table 3-6 presents the size, density, reciprocity, transitivity and centralization of these networks in all four schools. Only figures for wave 1 are presented here, because the situation remains essentially the same across waves. Figures for all waves are in Appendix 3B.

Unsurprisingly, friend networks are denser than very good friend ones. The variation in density across schools is largely explained by their difference in size[77], except for Paris 2, whose “very good friend” network is denser than the ones of Savoie 1 and 2, even though it is much bigger. As I mentioned before (see section 1.2.1.), students declared many more very good friends on average in this school than in the other ones, for unknown reasons.

76 This is important because, if we were to only consider “friend” nominations, some network measures would be highly distorted. For instance, transitivity estimates would be biased. Consider a triplet $i \rightarrow k \rightarrow j$ of “friend” nominations, with $i \rightarrow j$ being a “very good friend” nominations. If we only looked at “friend” nominations, this would appear as an intransitive triad (an “open” triangle), whereas in reality it clearly is a transitive one.

77 For a given average outdegree, a larger network is automatically sparser, because the total number of emitted ties is proportional to the number of nodes but the total number of possible ties is proportional to the squared number of nodes.

Table 3-6: Descriptives about Friendship Networks in Wave 1

Very Good Friends

	Paris 1	Paris 2	Savoie 1	Savoie 2
Size	83	264	169	219
Density	0.093	0.069	0.051	0.039
Reciprocity	0.58	0.49	0.49	0.50
Transitivity	0.42	0.33	0.29	0.26
Indegree Centralization	0.11	0.14	0.15	0.08
Outdegree Centralization	0.42	0.34	0.15	0.17

Friends + Very Good Friends

	Paris 1	Paris 2	Savoie 1	Savoie 2
Size	83	264	169	219
Density	0.29	0.15	0.17	0.13
Reciprocity	0.71	0.61	0.61	0.59
Transitivity	0.52	0.39	0.38	0.34
Indegree Centralization	0.24	0.23	0.24	0.16
Outdegree Centralization	0.55	0.36	0.53	0.31

Size: number of nodes in the network.
Density: number of realized ties among all possibles ones; that is, number of ties / [n*(n-1)] with n the number of nodes.
Reciprocity: proportion of all the ties $i \rightarrow j$ for whom $j \rightarrow i$ exists.
Transitivity: proportion of all the triplets $i \rightarrow k \rightarrow j$ for whom $i \rightarrow j$ exists ("outgoing transitive partner" definition of transitivity)
Centralization: Freeman's centrality index (Freeman1978). The index goes from 0 to 1, with 0 being a perfectly decentralized network (all nodes have the exact same degree) and 1 a perfectly centralized network (a single node receives/emits all ties in the network).

Reciprocity appears relatively high, which is to be expected for friendship networks. On average, about 60% of all nominations are reciprocated in the "friend" networks (70% in Paris 1, most likely due to it being smaller). The values are lower in the "very good friend" ones (around 50%), but this actually indicates more reciprocal nominations once the difference in density between the two types of network is accounted for[78]. Though high

78 By construction, a denser network will have a higher reciprocity value (if there are more ties altogether, then $i \rightarrow j$ is more likely to be observed, regardless of whether $j \rightarrow i$ exists or not). Therefore, it is useful to consider the reciprocity/density ratio, which indicates how much more likely reciprocated ties are relative to any tie (i.e. the probability change in $i \rightarrow j$ given that $j \rightarrow i$ exists). For example, in Savoie 1, reciprocity and density are of respectively 0.49 and 0.051 in the "very good friend" networks, and 0.61 and 0.17 in the "friend" networks. If we compute the corresponding odd-ratios, we find that a "very good friend" nomination $i \rightarrow j$ is 17 times more likely to happen, rather than not, when $j \rightarrow i$ exists; for any friendship nomination (either "friend" or "very good friend"), this is only 7.67. Thus "very good friend" nominations are in fact more reciprocal

compared to any random network, this is a pretty moderate amount of reciprocity for friendship networks specifically: more than a third of ‘friend’ nominations are not reciprocated, meaning that it is often the case that two students have a different perception of their relationship with one another. With that said, unreciprocated nominations are strongly concentrated around a small number of students, which is something I will come back to in a moment.

Transitivity is a measure that captures the tendency toward the closure of triplets in the network ($i \rightarrow k \rightarrow j$ implies $i \rightarrow j$). It is fairly high for both types of networks, between 30% and 50% (which can be interpreted as the probability for $i \rightarrow j$ to exist given $i \rightarrow k \rightarrow j$[79]). The prevalence of closed triplets is a good measurement of the propensity of ties to cluster into higher-order structures in the networks. This means that students’ friendships tend to aggregate into relatively coherent groups – or, to put it otherwise, that the friends of one’s friends tend to be one’s friends as well. This is coherent with the literature on adolescent sociability, which points out to the importance of the broader peer group in determining dyadic relationships (e.g. Bidart 2010; Coleman 1961).

Finally, centralization is a measurement of the concentration of ties around a few nodes; that is, it measures sociometric inequalities in the networks. Starting with the indegree centralization, that is, the inequalities in terms of received nominations, the Freeman centralization index appears to be very small: around 0.10 for “very good friend” networks, and around 0.20 for “friend” ones (with minor variations across schools, Savoie 2 being less centralized than the others). This means that *received* nominations are rather evenly split between students. Certain students do have more friends than others, of course, but we are far from a situation of strong inequalities (another way to put it would be that the networks do not have a very pronounced core-periphery structure). This seems to suggest that popularity processes, which are often described in the literature on adolescent friendships as particularly strong (cf. the references presented in chapter 1, section 1.1.2.), are relatively moderate in these schools; this is something that we shall come back to later on (in chapter 5 and in the conclusion of the dissertation). It is also worth noting that “friend” networks are more centralized than “very good friend” ones. This is probably due to the fact that being close friends with someone demands a stronger investment in terms of time and attention, compared to weaker ties that can easily be maintained through occasional contact and/or group

than “friend” ones once density is accounted for.

79 Just like reciprocity, the measure should be compared to network density, i.e. the baseline probability for any tie to occur.

interactions. Therefore, it is easier for a student to keep a large number of friends than of very good friends.

Compared to indegree centralization, outdegree centralization appears much higher (for all schools in the “friend” networks, and for Parisian schools only in the “very good friend” networks). This is an interesting, perhaps counter-intuitive result: students do not differ from one another by how much friendship nominations they receive, so much as by how many they emit. This shows that they can have widely different understandings of the terms “friend” and “very good friend” (cf. section 1.2.1. above). This also contributes to explaining the relatively high number of unreciprocated nominations that I mentioned before: the fact that emitted nominations are much more concentrated than received ones implies that there are a few students who send much more ties than they receive. These “super-nominators”, that adhere to a very extensive definition of friendship[80], thus drag reciprocity estimates down; but this does not necessarily mean that the corresponding relations are truly asymmetrical (there may be balanced, weak ties, that one student calls a friendship and not the other because they understand the word differently, bur not because their attachment or emotional investment into their relationship is different).

All the measures presented so far only concern the networks of the first wave. As I mentioned, these structural features do not change much over time: friendship networks remain more or less equally dense, reciprocal, transitive and centralized at least up to wave 4. There seems to be a slight drop in friendship density in wave 6, which could be related to the COVID pandemic and the related lock-down; however, this is very difficult to say for sure with the data at hand, and I will not engage with this hypothesis here (see Appendix 3A for a discussion). What should be retained from these results is that there is no clear temporal trend as per the structure of friendship networks. However, this very much does not mean that there is no relational change whatsoever. On the contrary, the renewal rate of ties is pretty high. This is shown in Table 3-7, which gives the proportion of nominations at a given wave that was repeated at the next wave.

80 In some cases, they may also simply over-state the number of their social contacts, either because they want to brag to the interviewer or, more subtly, because they try to improve their own self-image (since the number of friends is such an important status criteria for teenagers). In particular, I observed a few cases of students that were objectively marginalized among their peers, ut that nevertheless declared many friends in the questionnaires. However, I tend to believe that this remains relatively rare, and cannot be enough to explain the strong centralization of emitted ties.

Table 3-7: Proportion of Friendship Ties Maintained from One Wave to the Next

Very Good Friends

	Paris 1	Paris 2	Savoie 1	Savoie 2
wave 1 → 2	0.70	0.55	0.54	0.55
wave 2 → 3	0.57	0.59	0.52	0.50
wave 3 → 4	0.61	0.61	0.58	0.59
wave 4 → 6	0.56	0.46	0.49	0.41
wave 1 → 6	**0.39**	**0.28**	**0.30**	**0.23**

Friends + Very Good Friends

	Paris 1	Paris 2	Savoie 1	Savoie 2
wave 1 → 2	0.70	0.61	0.63	0.59
wave 2 → 3	0.72	0.72	0.66	0.66
wave 3 → 4	0.72	0.66	0.64	0.64
wave 4 → 6	0.66	0.55	0.56	0.56
wave 1 → 6	**0.53**	**0.38**	**0.37**	**0.32**

We can see that, depending on the wave and school, between 55% and 72% of all friendship nominations get repeated to the next wave. For "very good friend" nominations, this is even a bit less, from 41% to 70% (with an average around 60%). Altogether, this means that less than two thirds of the nominations that students emit at one wave are repeated to the next. If we consider the first and last waves, things are even more striking: about a third of the nominations emitted on wave 1 are also emitted 2 years and a half later, in wave 6[81]. Therefore, there is ample change in the friendship networks over time, which shows that students exhibit a high level of relational activity.

Given what we saw before about the stability of aggregate network features in time (transitivity, centralization etc.), this means that this renewal occurs while leaving the overall structure of the network mostly untouched. This seems to suggest that the network is in a state of *equilibrium*, concerning its main structural features at least (we shall see in chapter 7 that this may not be true for other aggregate features, notably socioeconomic homophily).

81 Of course, these measures are computed while accounting for the fact that certain students left the school in the meantime (the percentages given are expressed in proportion of the total number of ties that *could have been* maintained).

Chapter 4: Socioeconomic Mixing: A Descriptive Approach

Having reviewed the necessary theoretical and methodological tools, we can now turn toward the empirical analysis of socioeconomic homophily among students. To begin with, it is necessary to establish the extent to which their sociability is impacted by their socioeconomic background, in a descriptive fashion. In doing so, particular attention should be paid to the fact that this impact may depend on the type of ties that are considered, and may also vary across schools or evolve over time. This chapter is therefore meant to answer the three first research questions that were raised in the introduction:

> *RQ1: Given a situation of socioeconomic diversity, to what extent is there socioeconomic homophily in teenagers' friendships?*
>
> *RQ2: Does the strength of socioeconomic homophily depend on the type of relationship that is considered?*
>
> *RQ3: Does the strength of socioeconomic homophily vary over time and/or across schools?*

Additionally, it will also provide some exploratory insights regarding the link between socioeconomic background and peer status, following the discussion from chapter 1.

Homophily is defined as the fact that relationships occur at a higher rate among individuals that are similar in regard to a particular attribute. Similarity can refer to exact matching for categorical attributes (e.g. being of the same gender) as well as to relative proximity for continuous ones. For socioeconomic background, which is measured through a continuous variable (Socio-Academic Scores), the question is therefore that of the socioeconomic distance between students.

There are several ways in which this broad concept can be operationalized. The first section therefore introduces different measures revolving around the concept of homophily. The objective is to provide a cross-sectional picture of the state of socioeconomic homophily at the beginning of the study (1). Then, the second section looks at temporal trends from wave 1 onward, in order to see whether the homophilic patterns in students' relations evolved over the course of middle school (2). The third section compares the size of socioeconomic homophily with that of other attributes: gender, ethnicity and grades (3). Finally, the last section discusses patterns of hierarchy and centralization in friendship networks (4).

1. Socioeconomic Homophily in the First Year of Middle-School

1.1. Permutation Tests

To measure the socioeconomic distance between two students, the simplest measure is to consider the absolute difference of their Socio-Academic Scores (SAS)[82]. By extension, socioeconomic homogeneity in friendship ties can be operationalized as the average value of this difference over all ties in a given network. One would say that there is socioeconomic homophily in the network if this value is *lower* than what would be expected by chance (lower values indicate smaller socioeconomic differences, i.e. more similarity). However, how we define this reference scenario of "what would be expected by chance" is not straightforward. Indeed, if we were to re-allocate ties randomly among nodes in the network, compute the corresponding value of the average SAS difference, and compare it to the observed value, we would nevertheless lack a notion of statistical significance. This is because the usual assumption of inferential statistics that observations are independent does not hold in the case of a network. As of now, there is no generic statistical law that permits inference under these conditions (Lazega 2014).

To circumvent this issue, it is possible to simulate a fictive distribution of networks, and to use it as an empirical reference distribution to assess statistical significance. Most of the models that will be used throughout the dissertation rely on this basic idea. One of the simplest ways to do so is to resort to *permutation tests*. The idea is to randomly permute the nodes' labels in the sociomatrix of the observed network, and to recompute the statistic of interest in these permuted networks (here the average absolute difference in SAS among ties). In practice, this means that the network structure remains the exact same, but that students are allocated random positions in that network: Claire "becomes" Mohamed, Mohamed "becomes" Pedro, etc. However, they keep the same individual attributes after they have been permuted – Claire keeps the same SA score in all the permutations and in the observed data. This procedure preserves the structure of the original network, which means that the dependency of ties remains untouched (for e.g., there is the same number of triangles in all the permuted networks); but it breaks its relationship with the tested statistic (since attributes are

82 As a quick reminder: socio-academic scores are a continuous measure of socioeconomic background, which are built from the expected academic performance of occupational groups (cf. chapter 3 section 1.2.3).

not permuted), thus providing a null distribution to assess the statistical significance of the observed value. Empirical p-values are computed as the proportion of permuted networks for which this statistic is smaller than in the observed network[83].

This is done for four networks in each school: the “friend”, “very good friend”, “out-of-school friend” and “parental friend” networks. The “friend” network is made of both “friend” and “very good friend” nominations (a good friend is also a friend), and the “very good friend” network is made of the “very good friend” nominations exclusively. “Out-of-school friends” are the very good friends that the respondent also declared to sometimes see outside of school; and “parental friends” are the very good friends whose parents know the respondent’s parents (according to the respondent). Each of these networks is less dense than the previous one, as each is a strict subset of the previous one (all parental friends are out-of-school friends, which are very good friends, which are friends). Thus, the corresponding nominations can be seen as increasingly demanding. For now, only the first wave of data collection is considered (end of the first year of middle school). Table 4-1 gives the observed values of the socioeconomic homophily statistic, its mean values in the simulated networks, and the empirical p-values[84]. Moreover, Figure 4-1 shows the density curves of SAS differences among ties, again in observed and permuted networks (only for “very good friend” networks; the pattern is similar for the other networks). If there is socioeconomic homophily, then there should be more low-difference ties and less high-difference ties in the observed networks compared to the permuted ones.

This is indeed what we observe in three schools: Paris 1, Paris 2 and Savoie 1. Moreover, in all three cases, the homophily statistic is significantly lower in the observed networks than in the permuted ones. Visually, Figure 4-1 suggests that socioeconomic homophily is particularly strong in Paris 1: the solid curve, representing the observed values, is further away from the shaded array, which represents the simulated ones. Table 4-1 confirms this, but with a nuance: the observed value of the average SAS difference is in fact very close in the three schools, but the mean value in permutations is higher in Paris 1. In other words, students in the three schools have on average the same socioeconomic distance from their friends; but, because the socioeconomic diversity of the school is higher in Paris 1, this represents a larger gap compared to what would be expected by chance. Thus, we can

83 This is similar to the *Quadratic Assignment Procedure* (QAP) first proposed by Krackhardt (1988)

84 Appendix 4A further shows the same test for the networks made of the nominations of the five friends students spend the most time with. Homophily has more or less the same strength here as it does in the “very good friend” networks.

indeed say that socioeconomic homophily appears higher in this school – but not socioeconomic homogeneity. There is also a slight difference between Paris 2 and Savoie 1, but it is too small to be interpreted with confidence.

By contrast, in Savoie 2, the observed values of the statistic are not statistically different from the mean of the simulated values. The observed value is slightly inferior to the mean of the simulated ones for 'very good friend' nominations. For 'friend' nominations however, the observed value is slightly greater than the mean in simulations, meaning that friendships are *more* diverse than what would be expected by chance (this is not statistically significant though). Thus, there does not appear to be any socioeconomic homophily in friendship ties in Savoie 2.

Finally, Table 4-1 reveals that socioeconomic homophily is more pronounced for “very good friend” nominations than for “friend” ones: the expected values of the statistic under the permutations are virtually the same, but observed values are lower for the “very good friend” networks[85]. With the exception of Savoie 2, the observed values are also lower in the “out-of-school friend” and/or “parental friend” networks, compared to the “very good friend” one. Therefore, it seems that homophily is stronger for more demanding relationships. The comparison is not easy with these figures, though, because the strength of socioeconomic homophily remains hard to grasp intuitively. The next sections will attempt to remedy this issue.

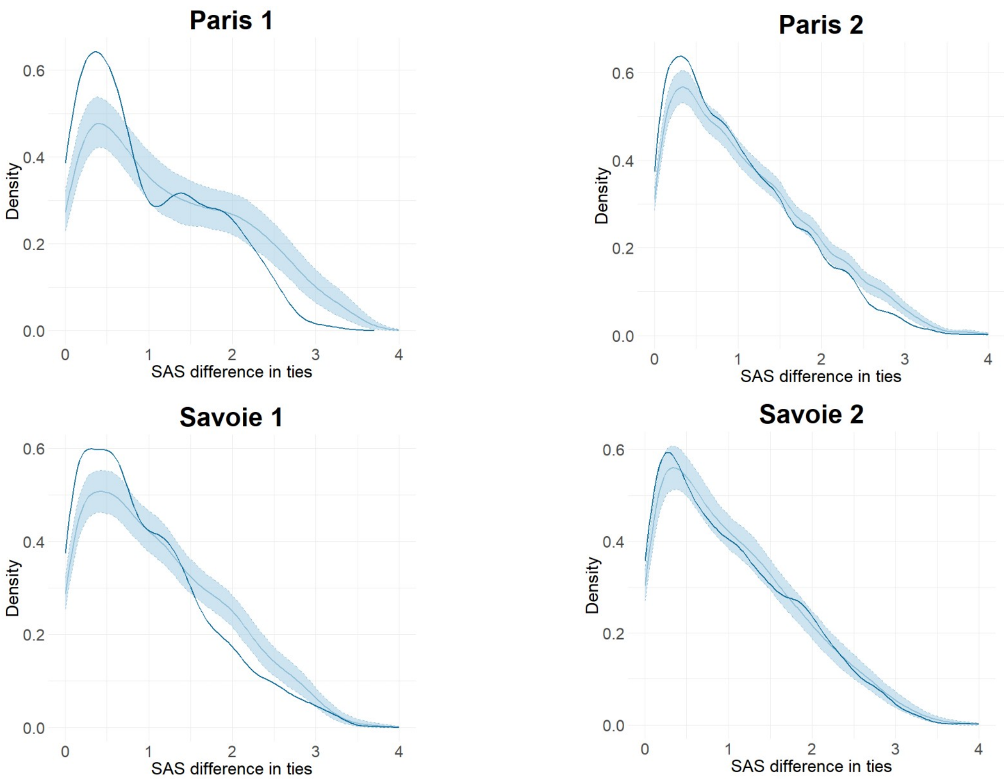
85 Note that the “friend” networks also include the “very good friend” nominations (more or less one third of all ties), so homophily would appear even less pronounced if we only considered “friend” nominations alone. The reason I include both types of nominations in the “friend” networks is because a network made of the sole “friend” nominations would have a somehow misleading structure (e.g. a friendship tie between two students may appear as isolated, whereas it is part of a dense cluster of “very good friend” ties, which would be invisible when only considering “friend” nominations).

Table 4-1: Average SAS Absolute difference in Friendship Ties – “Friend”, “Very Good Friend”, “Out-of-School Friend” and “Parental Friend” Networks, Wave 1

	Friend			Very Good Friend		
School	Observed	Mean in Permutations	P-Value	Observed	Mean in Permutations	P-Value
Paris 1	1.10	1.24	0.00	0.99	1.24	0.00
Paris 2	1.02	1.09	0.00	0.98	1.09	0.00
Savoie 1	1.05	1.13	0.00	0.98	1.13	0.00
Savoie 2	1.12	1.08	0.99	1.05	1.08	0.22

	Out-of-School Friends			Parental Friends		
School	Observed	Mean in Permutations	P-Value	Observed	Mean in Permutations	P-Value
Paris 1	1.02	1.24	0.00	0.95	1.24	0.00
Paris 2	0.92	1.08	0.00	0.89	1.09	0.00
Savoie 1	0.89	1.13	0.00	0.88	1.12	0.00
Savoie 2	1.03	1.08	0.20	1.05	1.08	0.25

Figure 4-1: Density Curves of SAS Absolute Differences in Friendship Ties – “Very Good friend” Networks, Wave 1



Note: Observed networks = solid dark blue line. Average in permutations = light blue line. 95% confidence interval in permutations = light blue shaded array.

1.2. Probability of Friendship Nominations: Introducing ERGM

Permutation tests are a valuable tool: they account for the structure of the network and offer a non-parametric test (density curves do not assume any shape for the distribution of SAS absolute differences). However, they have important limitations. First, their assessment of the strength of socioeconomic homophily is not easy to interpret, nor to compare between different networks.

Second, they only consider the "raw" value of the tested statistic – in that case, the average absolute difference in SA scores among ties. However, this is not necessarily the best indicator of socioeconomic homophily, as it does not control for the centrality of students from different backgrounds. Imagine that upper-background students are both more active (they send more ties, regardless of the receiver) and more popular (they receive more ties, regardless of the sender) than lower-background ones. This would mechanically result in ties between upper-background students being more likely than among lower-background ones. Conceptually, however, this would not signal a greater proximity between same-background students: the increased activity or popularity of certain groups would be "blind" to the socioeconomic origin of the receiver or sender of the tie. This is particularly important if one wishes to separate socioeconomic homophily and hierarchies, which, as we saw in chapter 1, is an important question raised by the literature.

Therefore, we need a different way to look at homophily. The one most commonly used in the literature, which is also the one that will be used throughout most of the dissertation, is to consider the odds of friendship nominations as a function of dyadic similarity. In the case of socioeconomic homophily, we therefore ask whether the chances to observe a tie $i \rightarrow j$ (student i nominated j as a friend) depends on the socioeconomic distance between i and j. This opens the way to multivariate modeling, where one controls for specific attributes or network features when computing homophily estimates. In that regard, the permutation tests presented above can be seen as analogous to univariate tests of correlation in "classical" statistical analysis.

Network inferential models are essentially logit regressions performed at the level of the dyad. Each directed dyad – i.e. ordered pair of nodes (i;j) – is considered as an observation, which can take the value 0 or 1 depending on whether i named j as a friend. Independent variables, usually called "covariates", are also located at the level of the dyad, although they can be a combination of individual-level attributes. For example, gender similarity would be defined as the two nodes *i* and *j* having the same gender. Similar to a "classical" logit model, the estimated coefficients then give the conditional change in the log-odds of observing a tie associated to the value of the covariate, net of the other modeled effects. In this chapter, I will resort to a class of cross-sectional models called the Exponential Random-Graph Model (ERGM) (Lusher, Koskinen, and Robins 2013). There is more to these models than a tie-level logit regression; but, for now, I only use simple models, which do not attempt to model dependency across network ties.

The models below include four parameters. The "edge" parameter captures the baseline odds of any one tie occurring – that is to say, it controls for network density. It is an equivalent of the intercept term in classical regression models. The second parameter is a homophily parameter. It captures the change in the odds of observing a tie as a function of the socioeconomic distance between the two nodes, measured as the absolute difference between their Socio-Academic Score (SAS). The estimated coefficient is the log-odd change in tie probability for each unit change in this absolute difference. If there is socioeconomic homophily in the network, this coefficient should be negative: indeed, a one unit increase in the absolute difference means that the nodes are more distant, which should translate as lower odds of tie occurrence. Finally, two parameters capture the propensity of nodes to emit and receive ties as a function of their SA scores. The first gives the log-odd change in tie probability for each unit change in the SAS of the node that emits a nomination, and the second for each unit change in the SAS of the node that receives a nomination.

This same model is fitted separately to the four friendship networks introduced above: the "friend", "very good friend", "out-of-school friend" and "parental friend" networks. Again, only wave 1 of data collection is considered. Results are presented in Table 4-2.

Table 4-2: Socioeconomic Homophily, Activity and Popularity in Friendship Networks (ERGM) – Wave 1

		Friends	Very Good Friends	Out-of-School Friends	Parental Friends
Paris 1	Density	-0.559*** (0.049)	-1.803*** (0.073)	-2.664*** (0.103)	-3.387*** (0.151)
	Absdiff SAS	-0.242*** (0.035)	-0.389*** (0.059)	-0.360*** (0.081)	-0.529*** (0.117)
	SAS receiver	-0.036 (0.029)	-0.071 (0.047)	-0.087 (0.064)	0.206** (0.091)
	SAS sender	-0.161*** (0.029)	-0.068 (0.047)	0.118* (0.065)	0.219** (0.091)
Paris 2	Density	-1.579*** (0.018)	-2.396*** (0.025)	-3.996*** (0.053)	-4.212*** (0.060)
	Absdiff SAS	-0.142*** (0.014)	-0.196*** (0.021)	-0.310*** (0.047)	-0.377*** (0.055)
	SAS receiver	0.026** (0.012)	-0.006 (0.017)	0.016 (0.037)	0.027 (0.044)
	SAS sender	-0.006 (0.012)	-0.044*** (0.017)	0.031 (0.037)	0.003 (0.044)
Savoie 1	Density	-1.367*** (0.029)	-2.613*** (0.048)	-3.298*** (0.069)	-3.380*** (0.073)
	Absdiff SAS	-0.167*** (0.022)	-0.272*** (0.040)	-0.447*** (0.061)	-0.483*** (0.064)
	SAS receiver	-0.048*** (0.018)	-0.036 (0.032)	-0.063 (0.048)	0.021 (0.050)
	SAS sender	-0.085*** (0.018)	-0.122*** (0.032)	-0.0002 (0.048)	0.082* (0.050)
Savoie 2	Density	-2.000*** (0.024)	-3.207*** (0.041)	-4.055*** (0.062)	-4.249*** (0.068)
	Absdiff SAS	0.011 (0.017)	-0.125*** (0.031)	-0.131*** (0.047)	-0.127** (0.049)
	SAS receiver	0.133*** (0.014)	0.157*** (0.025)	0.162*** (0.038)	0.249*** (0.040)
	SAS sender	0.163*** (0.014)	0.224*** (0.025)	0.198*** (0.038)	0.288*** (0.040)

***: p<0.01; **: p<0.05; *: p<0.1

Note: standard deviations are indicated between parentheses.

I will only comment on the coefficients that are indicative of socioeconomic homophily (*absdiff* SAS) for now; centralization will be tackled in section 4 below. In all four schools, these coefficients are negative and statistically significant, as they indicate that there is socioeconomic homophily in students' friendships. The one exception is the "friend" network in Savoie 2, where the effect is close to 0; meaning that weak friendship ties are not homophilic in this school..

Starting with Savoie 2, it is important to ask why there is a slight effect of socioeconomic homophily now, when it did not show in the permutations. This comes from controlling for the sender and receiver effects: differences in centrality masked the homophilic patterns in the permutations. Additionally, permutation tests tend to be more conservative than ERGM in the estimation of the standard error of the parameters.

In any case, homophily estimates tend to be higher for more demanding ties – very good, out-of-school, and parental friends, in that order (note that comparing effect sizes across logit models can be tricky; see the methodological note in the boxed section below). This confirms the pattern observed in permutation tests. There are two ways to understand it. First, socioeconomic homophily appears to be stronger for more intimate or important relationships. This seems to invalidate the idea that socioeconomic boundaries weaken with intimacy. One could imagine, for example, that homophily would be less important for strong relationships, where people know each other well and rely on shared experience and history not directly related to socioeconomic origin; and that this origin would act as a quick, efficient category of judgment for more distantly related peers. In other words, the more you know someone, the less ascribed characteristics such as socioeconomic background would matter. Yet this does not appear to be the case[86].

Second, students' sociability outside of school might be more heavily impacted by a number of factors that induce socioeconomic homophily. These will be discussed in greater details in chapter 5, but we can already speculate that residential segregation, out-of-school activities or parental control tend to reinforce homophilic patterns. This is particularly true for networks of inter-generational closure – i.e. relationships between parents – as we would expect adults to be more sensitive to socioeconomic distance than children are. In that regard, within-school friendships appear more mixed than out-of-school ones; which in turn suggest that socioeconomic diversity at school may act as a heterophilic force relative to non-school contexts.

This gradation of homophily with the strength of ties brings layers of depth to the concept of "socioeconomic mixing". Indeed, the binary opposition between mixing and segregation may not be sufficient to describe the impact of social class on relational

86 The same pattern is observed for gender and ethnic homophily; the magnitude of homophily grows with the strength of ties (models not shown).

structures: mixing on a relatively superficial level – 'pals', acquaintances, playmates, persons that you find 'nice' – seems to coexist with (some) segregation at a deeper level – best friends, confidants, people that you spend the most time with or that are part of several of your social circles (school, activities, family, etc.).

Methodological Note: Comparing Log-Odd Coefficients across Models

An issue that we will encounter several times throughout the dissertation is that log-odd coefficients cannot be directly compared across logit models. This is because the size of the coefficients is sensitive to the total amount of variance explained by the model (Mood 2010). In general, researchers can resort to marginal effects instead, but this is impossible for network models (or to be more accurate, it has not been thoroughly investigated yet whether marginal effects are a reliable measure in the case of network models).

I will use various techniques to get around this issue throughout the dissertation, such as using one model coefficient as a benchmark to evaluate the effect size of other coefficients (intra-model comparison is still allowed) or restraining the interpretation to the sign and statistical significance of coefficients. However, in some cases, there are good reasons to still compare coefficients across models. Again, the reason that the size of coefficients cannot be compared is that they depend on how much total variance each model explains. Therefore, the bias induced strongly depends on the difference in goodness-of-fit across models; if one model is a poor fit and the other a good fit, the inter-model comparison of coefficients will be strongly biased. From there, it has been argued (controversially so) that the bias induced is small enough to be ignored in most applications (Kuha and Mills 2018).

I will not take sides in the broader discussion around this issue. What seems clear, however, is that comparing coefficients across models can be reasonable in specific cases, when (a) the goodness-of-fit of both models is similar (meaning we expect a small bias due to unobserved variance) and (b) the difference in size between the considered coefficients is large enough that it cannot be explained by this bias alone. Of course, this implies that the comparison cannot be numerically precise, but should only pertain to ordinal considerations (greater/smaller). In general, these conditions are met for models with a small number of explanatory variables, like the ones presented here.

Additionally, the permutation tests presented in Appendix 4A can be compared across different types of ties, since they do not rely on logged measures.

Now, how substantially important should these effect sizes be considered? Since SA scores are somewhat abstract, it is good to bear in mind the following benchmarks: the average score of the students whose head of the household is a manual worker or laborer (category 6 of the PCS) is -0.96, whereas it is 1.48 for managers and intellectual occupations (category 3)[87]. Therefore, one can consider a typical shift of 1.48+0.96 = 2.44 SAS points between two students from working- and upper-class backgrounds.

Log-odd coefficients can be turned into odd-ratios for easier interpretation. So for 2.44 points of social distance, the odd-ratio of a friendship occurring in Paris 1 is exp(-0.242*2.44) = 0.55: the nomination is 1.80 times less likely to happen rather than not, compared to a nomination among students with the exact same background[88]. Looking at very good friend nominations, this is 2.58; and for parental friends, 3.64[89].

Figures are lower in Paris 2 and Savoie 1: respectively 1.61 and 1.94 less chances for very good friends[90]. Interestingly, the homophily effect appears particularly strong for out-of-school friends in Savoie 1 (log-odd of -0.447, meaning an odd-ratio of 2.98), which likely has to do with the residential structure there[91]. Finally, the effect sizes are lowest in Savoie 2: for a 2.44 points difference, the odds ratio is of 1.36 for very good friends[92]. Moreover, unlike in other schools, it does not increase for out-of-school or parental friends. As for mere "friend" nominations, as we mentioned already, they are not homophilic in this school.

Overall, these effect sizes are of a moderate magnitude in Paris 2 and Savoie 1, at least for "friend" and "very good friend" nominations; and they are particularly small in Savoie 2. Figures are higher in Paris 1, but they remain somehow moderate. Surely, these levels of homophily are not to be neglected, especially considering that there has to be some noise in the measure of social background, which should reduce the size of estimates. Nevertheless, these are not the figures one would expect if friendship networks were completely segregated by social class; cross-background ties, though less likely than others, are by no means uncommon.

87 The head of the household is defined as the parent with whom the child lives for divorced couples, and as the highest occupation out of the two parents otherwise.

88 0.55 *more* chances is equivalent to 1.80 *less* chances (1/0.56=1.80).

89 Respectively: 1/exp(-0.389*2.44) and 1/exp(-0.529*2.44).

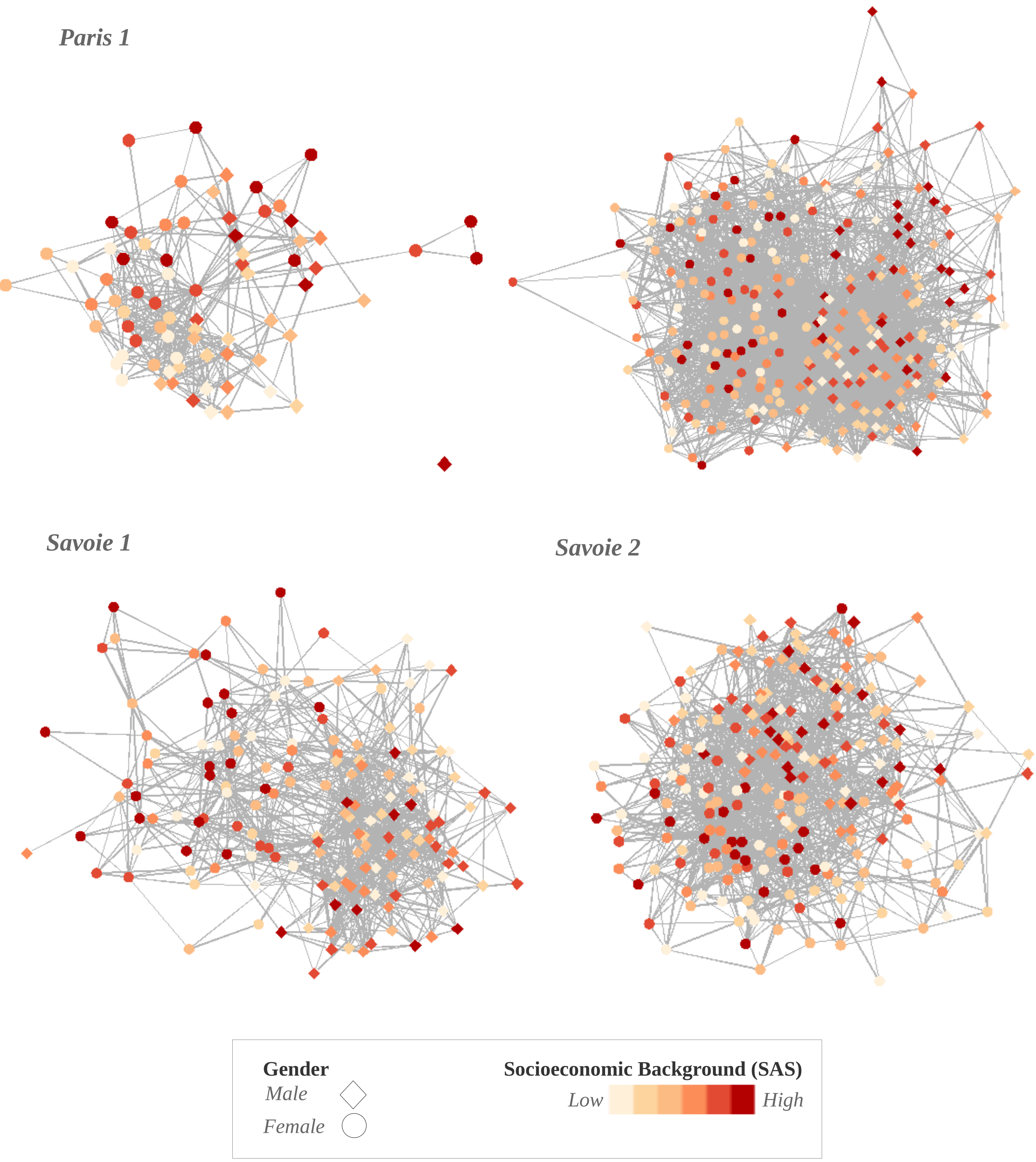
90 1/exp(-0.196*2.44) and 1/exp(-0.272*2.44).

91 Savoie 1 is the school with the most spatial dispersion among students' homes (rural area), which could imply that social boundaries are particularly salient when transportation issues are involved (which is the case for out-of-school friends, much more than for very good ones).

92 1/exp(-0.125*2.44).

A quick look at the networks confirms this intuition visually. In Figure 4-2, “very good friend” networks are plotted for each school. Individuals are represented as dots, and nominations as straight lines (note that symmetric and asymmetric ties are not differentiated; while one could add arrows to indicate the direction of nominations, this would make the graph harder to read and is not essential here). The nodes’ arrangement in space is meant to minimize the apparent length of ties. Nodes are colored depending on their SA score (darker shades mean higher scores), and gender is indicated by the shape of the node (diamond for boys, circles for girls). As can be seen, students from different backgrounds are distributed relatively evenly in the relational space, except perhaps in Paris 1 – where there are nevertheless many cross-background ties. By contrast, the split between boys and girls is much more apparent. Therefore, these networks do not appear to be socioeconomically segregated, in the strongest sense, though statistical analysis revealed that they are however, homophilic.

Figure 4-2: Networks of 'Very Good Friend' Nominations – Wave 1



1.3. Exposure

Both permutation tests and ERGMs consider socioeconomic homophily from the perspective of the dyads: observations are made of friendship ties, not of the individual among which these ties are established. However, to get a better sense of what socioeconomic homophily implies for students, it might be more appropriate to look at things from an individual perspective: for example, how many students are "exposed" to others that are socially different from them? Crucially, the theoretical relevance of exposure is not necessarily linear: there may be a bigger difference between having no distant-background friend, or having a single one, than between having seven and eight of them. In other words, instead of treating all ties as equally important, the objective here is to reason in terms of threshold: do students have *at least one* friend that is socially far-away from them? Even though, say, an upper-class student might have a majority of upper-class friends, it remains important to know whether they have some lower-class friends as well: this could matter a great deal in terms of positive attitudes toward other social groups, or in terms of "openness" to socially different others.

In order to do this, I start by dividing the student population into four quartiles of Socio-Academic Scores (SAS), for each school separately[93]. Then, for each student in one of the extreme quartiles (i.e. part of the 25% of students in the school with the lowest/highest SAS), I simply count how many friends they declared from the opposite extreme quartile. This is done once again, for friends, very good friends, out-of-school friends, and parental friends' networks. These measures are clearly biased, in the sense that they do not account for the density of the networks or the size of the population; yet, for this very reason, they come closer to the effective amount of socioeconomic diversity in students' networks than the other network measures presented so far.

Note that I focus here on relationships among students that stand at both extremes of the social spectrum, which makes for a conservative measure. Indeed, students from the middle-classes are structurally closer to all other students: thus, homophily will mechanically appears to be higher when excluding them from the treatments. However, when discussing socioeconomic mixing at school, the relationships between students from working-class and upper-class backgrounds are generally the ones that are expected to be the most difficult, so it makes sense to pay particular attention to them.

93 The cut-off values change slightly per school, but one can consider the following benchmarks: the first quartile roughly corresponds to SA scores from -1.7 to -0.55, the second quartile from -0.55 to 0.20, the third quartile from 0.20 to 1 and the fourth quartile from 1 to 2.5.

Results are shown in Table 4-3. Since these are individual-level measures, the four schools are pooled together. There is some variation across schools, which follow to varying degrees the orders of magnitude observed for log-odds in the previous section; but not to the point where it would change the conclusions. School-specific figures are in Appendix 4B.

Table 4-3: Exposure to Socioeconomically Distant Others – Wave 1

		Nb of Emitted Nominations toward Students from the Opposite Quartile			
		0 friend	1 friend	2 friends	3+ friends
Friends	students from lowest quartile (%)	2	1	3	94
	students from highest quartile (%)	1	3	5	91
Very Good Friends	students from lowest quartile (%)	12	18	16	54
	students from highest quartile (%)	15	19	16	50
Out-of-School Friends	students from lowest quartile (%)	44	34	12	10
	students from highest quartile (%)	55	21	13	11
Parental Friends	students from lowest quartile (%)	61	24	9	6
	students from highest quartile (%)	60	23	13	4

Example of Interpretation: out of all the students that are in the lowest SAS quartile of their school, 12% declared no very good friend from the upper SAS quartile; 54% declared 3 or more very good friends from the upper SAS quartile; 44% have no out-of-school friend from the upper SAS quartile.

To put these figures into perspective, one can keep in mind the median number of emitted nominations for each type of tie: 27 (friends), 8 (very good friends), 3 (out-of-school friends) and 2 (parental friends) (the means are higher, respectively 31.4, 11.7, 3.5 and 2.9). Furthermore, by construction, each quartile of students makes up for 25% of the total population.

We can see that 94% of students from the lowest SAS quartile declared 3 or more friends from the highest quartile. Conversely, 91% of students in that high quartile nominated 3 friends or more from the lowest quartile. It is therefore clear that an overwhelming majority of students have at least some positive relationships with different-background peers, even though these can be somewhat weak ties. Moving on to "very good friend" nominations – indicative of stronger relationships – only about half of the students have 3 or more friends from the other extreme quartile. Still, less than 15% of students have no friends at all in the other extreme quartile (15% for those in the high quartile and 12% for those in the low quartile)[94].

Finally, looking at out-of-school and parental friends, about half of the students made no nomination at all toward the other group (up to 60% for parental friends). This is of course linked to the fact that there are less out-of-school ties in general; still, this means a large share of students have little contact with peers of an opposed background outside of school. In other words, whereas inter-group contact seems to be the dominant pattern within school, the lack of contact is much more ordinary outside of it.

1.4. Heterophobia

Strictly speaking, socioeconomic homophily is about positive contact (or lack thereof) between individuals from different backgrounds. However, it stands to question whether negative contact also occurs in relation to social distance; in other words, whether homophily – liking similar others – goes alongside *heterophobia* – disliking dissimilar others.

In the questionnaires, along with the "friend" and "very good friend" nomination rosters, students could also indicate which peers they disliked. The number of nominations was not limited. The resulting networks are slightly less dense than the "very good friend" ones, except in Savoie 2 where they are denser (see chapter 3, section 1.2.2). Furthermore, another question let students name people that bullied them, with a maximum of five nominations. These networks are applied the same treatment as in section 1.1, i.e. using simple ERGMs to get a measure of the odds for a nomination to be observed depending on the social distance among students. Results are shown in Table 4-4.

94 Working-class students tend to nominate more peers overall as friends, which is why they appear as slightly more "open" here.

Table 4-4: Socioeconomic Homophily, Activity and Popularity in Dislike and Bullying Networks (ERGM) – Wave 1

		Dislike	Bullying
Paris 1	Density	-2.170*** (0.081)	-4.780*** (0.256)
	Absdiff SAS	-0.257*** (0.061)	-0.099 (0.180)
	SAS receiver	0.008 (0.049)	-0.172 (0.147)
	SAS sender	-0.056 (0.049)	0.272* (0.147)
Paris 2	Density	-3.318*** (0.036)	-5.577*** (0.113)
	Absdiff SAS	-0.082*** (0.028)	-0.208** (0.097)
	SAS receiver	0.065*** (0.023)	-0.132* (0.079)
	SAS sender	-0.015 (0.023)	-0.002 (0.079)
Savoie 1	Density	-2.921*** (0.051)	-5.644*** (0.209)
	Absdiff SAS	-0.020 (0.037)	-0.208 (0.177)
	SAS receiver	-0.119*** (0.031)	-0.379*** (0.143)
	SAS sender	0.072** (0.031)	-0.265* (0.143)
Savoie 2	Density	-2.972*** (0.036)	-5.973*** (0.158)
	Absdiff SAS	0.043 (0.028)	0.035 (0.127)
	SAS receiver	-0.127*** (0.023)	-0.290*** (0.106)
	SAS sender	0.040* (0.023)	-0.128 (0.106)

***: p<0.01; **: p<0.05; *: p<0.1

In Paris 1 and 2, the coefficient is statistically significant and negative for disliking networks, which means that nominations become *less* likely as socioeconomic distance increases. This is also the case for bullying nominations in Paris 2. This is a somehow counter-intuitive result: disliking primarily occurs among socioeconomically similar peers. A plausible explanation might be that conflict occurs primarily among former friends or rivals within the same social circles, which would imply that homophily tends to produce "homophobia" as well (i.e. increased conflicts among similar individuals[95]). As for Savoie 1 and 2, there is simply no evidence of a relationship between socioeconomic distance and disliking or bullying (coefficients are non-significant and close to 0).

Therefore, it is clear that disliking and bullying nominations are not more frequent among distant-background peers; for Paris 1 and 2, they are even less frequent. This should not be misinterpreted as proving that no conflict emerges in relation to socioeconomic distance – as a matter of fact, some do, as we shall see in chapter 6. Rather, we can say that

95 To be clear, the term "homophobia" is used here to mirror that of "homophily". It has no relation to the usual sense of the term, as referring to discrimination toward homosexuality.

such conflicts do not appear to be more frequent than those among same-background peers, at least inasmuch as they crystallize into disliking or bullying nominations. There may be lighter forms of tensions that are not captured by these measures, though. At the very least, students from different backgrounds, although less likely to befriend one another, do not appear to be in open and systematic conflict.

2. Evolution of Socioeconomic Homophily throughout Middle-School

So far, we have focused on the first wave of data collection, when students were still in their first year of middle school; it stands to question whether the situation remains identical as they grow older. Once again, the objective is descriptive: to see whether the broad features observed in wave 1 still hold in later waves, regardless (for now) of the relational mechanisms that underlie this evolution.

To summarize, four main patterns emerged in the first wave. First, socioeconomic homophily appears relatively high in Paris 1, moderate in Paris 2 and Savoie 1, and very weak in Savoie 2. In Paris 1, students' friendships are more homophilic than in Paris 2 and Savoie 1, but not more homogeneous; the school has the highest amount of diversity in its population, so the same amount of socioeconomic homogeneity in friendship ties correspond to a greater gap compared to that expected under a random allocation of ties. Second, socioeconomic homophily gets stronger when looking at more demanding friendships. Third, this homophily remains of a moderate magnitude for within-school networks, where it does not preclude positive contacts among different students. Outside of school, however, segregation appears much more prevalent. Fourth, social distance does not increase the probability for disliking or bullying nominations, and even reduces it in some schools, which suggests that potential conflicts among socioeconomic groups are mostly kept in check. Now how do these fare two years and a half after wave 1[96]? The next subsections are organized just like the ones above, except for the permutation tests, which are not presented here because their results are redundant with those obtained from the ERGMs. They are presented in Appendix 4A. The next sections thus tackles friendship odds (2.1), exposure (2.2) and heterophobia (2.3).

96 In wave 1, students were nearing the end of their first year of middle school (age 11-12). Wave 6 took place in the first trimester of the third year (age 13-14).

2.1. Probability of Friendship Nominations

Friendship networks (both “friend” and “very good friend”) are observed at every wave of data collection. Out-of-school nominations, however, are only available for waves 1 to 3; the latest observation is therefore only a year after wave 1, not even halfway through the cohort’s time in middle school (end of second year). For this reason, I only focus here on friendship networks. Up to wave 3 at least, the trends are similar for out-of-school friends and parental friends as to what is observed in in-school friendships (Appendix 4A).

Similar to section 1.2. above, a simple ERGM including four terms (density, SAS homophily, SAS emission and SAS reception) is fitted to each school, and for each wave separately. Figure 4-3 shows the estimated homophily coefficients: the value of the coefficients is indicated by dotted lines, and the 95% confidence intervals by the shaded arrays in-between solid lines. Note that the y-axis scale is reversed: negative coefficients are positioned upward, so that a positive slope indicates an increase in homophily over time. The complete model outputs with all parameters are in Appendix 4C.

Starting with Paris 2 and Savoie 1, homophily estimates appear stable over time, for both “friend” and “very good friend” networks. Some minor oscillations can be observed at certain waves, notably for Savoie 1’s very good friends, but this does not make for any clear trend.

In Savoie 2’s “very good friend” networks, socioeconomic homophily, which was statistically significant in wave 1, disappears entirely in waves 2 and 3 (dots are under the 0 mark), but appears again in waves 4 and 6. Given that effect sizes are relatively small in general, these variations are not necessarily indicative of a major shift in homophilic patterns: they may simply illustrate minor variations under a stable mean, which is to be expected for most time series. However, it is also possible that they indicate an actual curved trend, with socioeconomic homophily decreasing toward the 5th and 4th grades then increasing again by 3rd grade. In either case, the evidence remains too limited to conclude with certainty. As for simple friendship networks, the trend is clearer: while there is no homophily whatsoever from wave 1 to 3, a negative (albeit statistically non-significant) coefficient is observed in wave 4, which grows again (this time reaching significance) in wave 6. As such, by the end of the period, there is a small homophily effect among weak friendships in Savoie 2, which was not the case in the early waves.

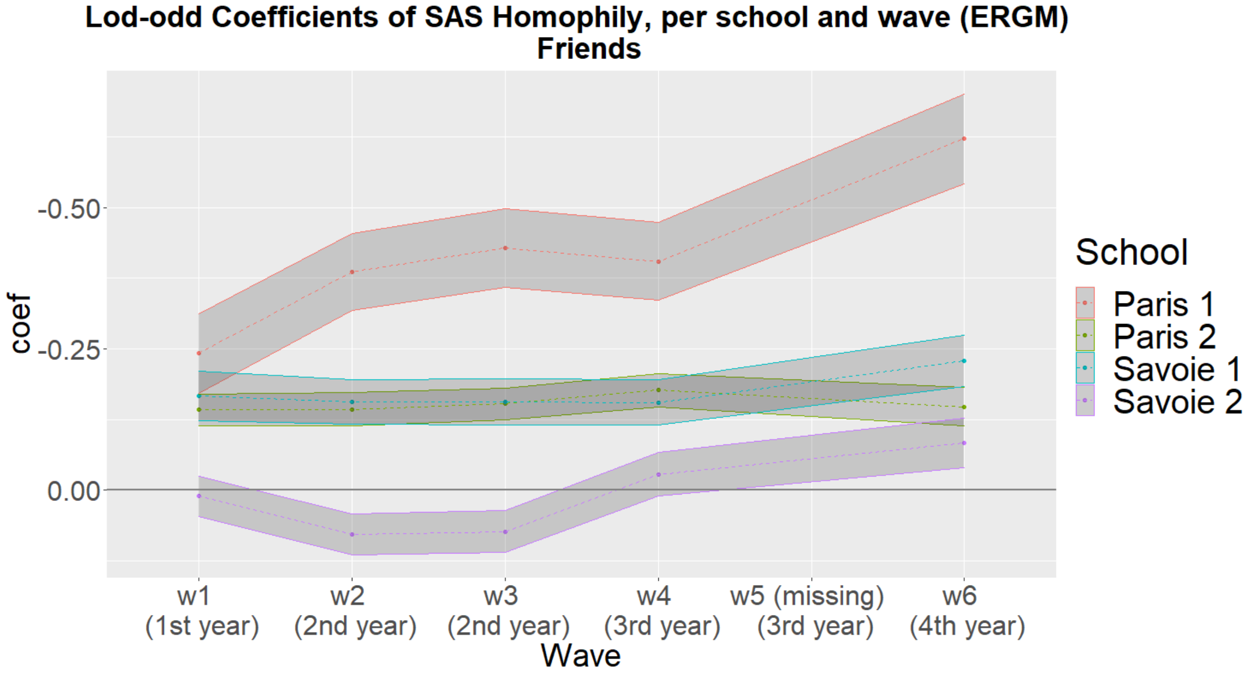
Finally, Paris 1 is the only school where a clear and consistent trend is observed across all waves. Indeed, there is a strong increase in socioeconomic homophily over time, both for friendship and good friendship networks. In wave 1, log-odd coefficients were equal to -0.242 and -0.389; in wave 6, they are of -0.622 and -0.666. This means a tie between an upper- and lower-background students (2.44 points of SAS difference) is now 4.56 (friends) and 5.08 (very good friends) less likely to occur, rather than not, than between same-background peers[97].

This is a striking evolution, and one which sets Paris 1 apart from the other three schools. While it was already the school with the highest levels of homophily in wave 1, the gap remained modest, and could potentially be explained by differences in network size or density (remember that exact numerical comparison of coefficients across schools is at risk of being biased). By wave 6, however, the school's networks became markedly more homophilic than in the other schools. In fact, these effect sizes are now strong enough that the friendship networks can legitimately be termed as socially segregated. Cross-class ties appear very rare, and the homophily of socioeconomic background is even higher than that of gender (see section 3 below). This can also be seen visually: Figure 4-4 shows Paris 1's "very good friend" networks at each wave. In wave 1 already, there is a rather apparent split between upper-background nodes (top right part of the plot) and middle- and lower-background ones (bottom left part). This divide becomes more pronounced over time, with the number of ties connecting the two zones also decreasing.

97 Corresponding odd-ratios, computed as 1/exp(-0.622*2.44) and 1/exp(-0.666*2.44).

Figure 4-3: Evolution of Socioeconomic Homophily over Time (waves 1 to 6) – Networks of friends (top) and very good friends (bottom)

Note: coefficients are indicated by dots and connected by dashed lines. 95% confidence intervals are marked by a shaded area and bounded by straight lines. The y-axis scale is not the same on the two plots. Negative coefficients are placed upward, as they indicate more homophily.
Numerically precise comparison of coefficients should not be made between different schools due to substantial differences in network size and density, but is reasonable for within-school time trends. There is approximately 6 months between each wave. Wave 5 was not observed (covid-related lockdown).
Other model parameters not shown here: density, SAS emission and SAS reception. See Appendix 4C.



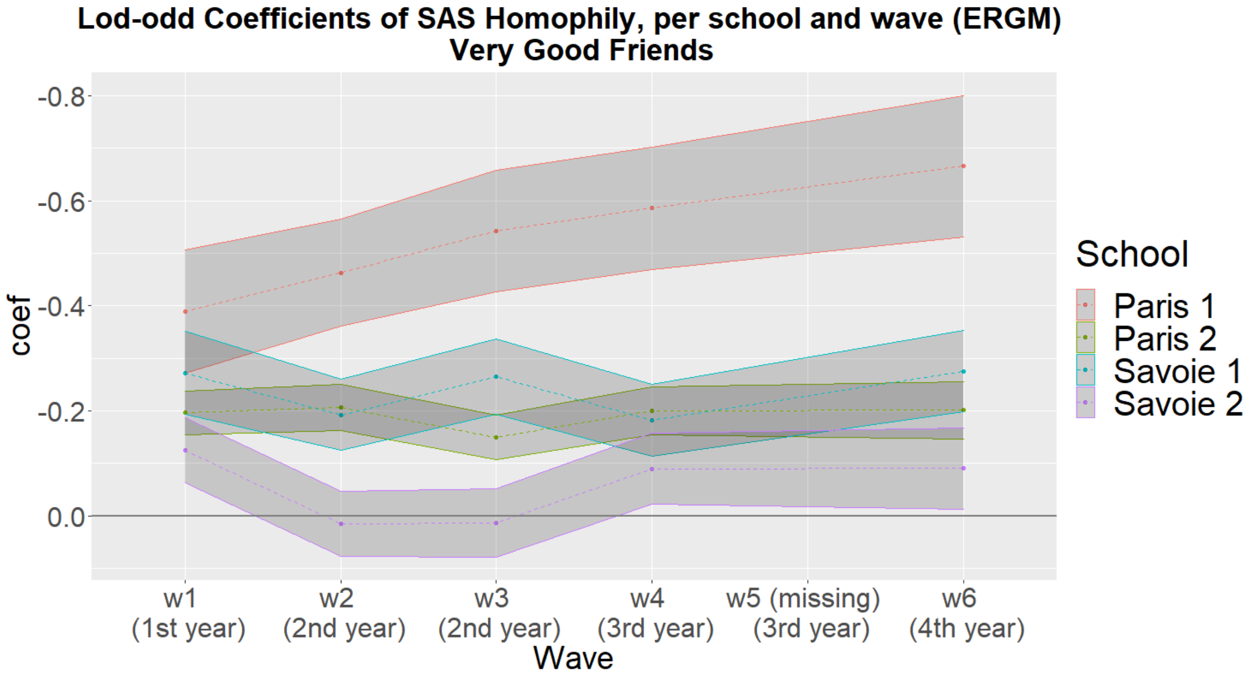


Figure 4-4: Networks of 'Very Good Friend' Nominations – Paris 1, Waves 1 to 6

2.2 Exposure

Let us now look at time trends in terms of exposure, which I define as the fact of having at least some friends from a distant background. This section replicates the analyses of section 1.2, but with wave 6 in addition to wave 1 in order to compare the first and last points of data collection. Only “friend” and “very good friend” nominations are considered, for reasons explained above (no out-of-school data after wave 3). Results are shown in Table 4-5: figures for waves 1 and 6 are presented jointly in each cell (wave 6 is at the bottom).

Table 4-5: Exposure to Socioeconomically Distant Others (wave 1 / wave 6)

		Nb of Emitted Nominations toward Students from the Opposite Quartile			
		0 friend	1 friend	2 friends	3+ friends
Friends	students from lowest quartile (%)	2 / 5	1 / 6	3 / 5	94 / 84
	students from highest quartile (%)	1 / 5	3 / 5	5 / 7	91 / 83
Very Good Friends	students from lowest quartile (%)	12 / 21	18 / 24	16 / 21	54 / 34
	students from highest quartile (%)	15 / 21	19 / 20	16 / 23	50 / 36

Example of Interpretation: out of all the students that are in the lowest SAS quartile of their school, 12% declared no very good friend from the upper SAS quartile in wave 1, and 21% in wave 6.

Overall, we see a reduction in the proportion of students that have 3 friends or more from a distant background: figures go up in each column, but down in the last one. Therefore, exposure to socially different others has been reduced between waves 1 and 6.

However, strictly speaking this is not primarily due to an increase in homophily. As we saw previously from ERGM coefficients, socioeconomic homophily has only increased over time in Paris 1, which happens to be the smallest out of the four schools (it makes up for about 11% of the total sample). As such, its impact on these pooled measures remains limited, and cannot account for this global reduction in exposure. Instead, the driving factor behind this reduction is a drop in network density on wave 6. Indeed, the average number of friendship nominations emitted by students remains relatively stable from waves 1 to 4, but abruptly decreases in wave 6. In all likelihood, this is a consequence of the national lock-down imposed in France during the COVID pandemic and of the subsequent sanitary restrictions, although it remains difficult to separate this from the effect of students getting older (see Appendix 3A for details). In other words, students simply have less friends in wave 6, regardless of social background: the fact that fewer respondents declared 3 or more friends from a distant background is due to the fact that their personal networks became smaller, not that they became more homophilic. This result constitutes a noteworthy reminder of the fact that to promote positive contact between distant-background peers, decreasing socioeconomic homophily is not the only efficient course of action: larger networks will also entail more contact with distant-background peers, through the sheer effect of having more friends altogether.

Comparing exposure between waves 1 and 4 instead – which have similar densities –, a slight decrease in exposure can also be observed, but much less pronounced than the one observed here (see Appendix 4B for the corresponding tables).

In any case, despite this reduction, positive exposure to dissimilar others remains relatively high in wave 6 for in-school friendships. Most students still have at least a few ties with peers from a distant background in the three biggest schools, namely Paris 2, Savoie 1 and Savoie 2. However, in Paris 1 specifically, exposure markedly decreased, in coherence with the increase in homophily observed in the previous section. This is shown in Table 4-6, which is an equivalent of Table 4-5 for Paris 1 alone (tables for the other 3 schools are in Appendix 4B).

Table 4-6: Exposure to Socioeconomically Distant Others – Paris 1 (wave 1 / wave 6)

		Nb of Emitted Nominations toward Students from the Opposite Quartile			
		0 friend	1 friend	2 friends	3+ friends
Friends	students from lowest quartile (%)	0 / 5	0 / 32	11 / 21	89 / 42
	students from highest quartile (%)	0 / 30	16 / 15	32 / 20	52 / 35
Very Good Friends	students from lowest quartile (%)	16 / 37	47 / 42	21 / 16	16 / 5
	students from highest quartile (%)	58 / 65	16 / 5	0 / 20	26 / 10

Example of Interpretation: out of all the students that are in the lowest SAS quartile of their school, 17% declared no very good friend from the upper SAS quartile in wave 1, and 37% in wave 6.

We can see in Table 4-7 that 37% of students in the lowest SAS quartile declared no friend at all from the highest quartile, and a mere 5% declared more than three. Conversely, 65% of high-quartile respondents declared no low-quartile friends, and 10% declared more than 3. Therefore, by the end of middle school, it seems that friendship ties with socioeconomically distant others have become the exception rather than the rule in Paris 1, even though a few ties still exist between dissimilar peers (also, remember that we compare the two extreme quartiles here: there is more mixing when considering students from a middle-class background).

2.3 Heterophobia

Finally, what of time trends in terms of socioeconomic heterophobia? Does the relationship between social background and the odds of negative ties change over time?

Similar to what was done for friendship networks, an ERGM with four terms is fitted to the disliking and bullying networks of each school and wave (density, SAS distance, SAS emission and SAS reception). Figure 4-5 presents the estimated log-odd coefficients for the homophily parameter (i.e. the absolute difference in SA scores between nodes). Additionally, these coefficients are also presented in Table 4-7, as the plots are difficult to read due to the different curves overlapping.

By construction, positive values indicate heterophobia (nominations are more likely as the social distance between students increases) and negative values homophobia (nominations are less likely as the social distance increases). The other coefficients (density, emission, and reception) are not shown here but in Appendix 4C.

Table 4-7: Log-odd Coefficients for SAS Heterophobia in Dislike and Bullying Networks (ERGM), per School and Wave

	Paris 1		**Paris 2**		**Savoie 1**		**Savoie 2**	
	Dislike	Bullying	Dislike	Bullying	Dislike	Bullying	Dislike	Bullying
Wave 1	-0.257*** (0.061)	-0.099 (0.180)	-0.082*** (0.028)	-0.208** (0.097)	-0.020 (0.037)	-0.208 (0.177)	0.043 (0.028)	0.035 (0.127)
Wave 2	0.022 (0.064)	-0.644* (0.339)	-0.130*** (0.033)	-0.324** (0.151)	-0.034 (0.038)	-0.113 (0.202)	0.071** (0.031)	0.176 (0.135)
Wave 3	0.097 (0.067)	0.043 (0.148)	-0.026 (0.031)	-0.076 (0.118)	-0.086** (0.038)	-0.490** (0.209)	0.117*** (0.031)	-0.228 (0.168)
Wave 4	-0.125* (0.074)	-0.260 (0.241)	-0.031 (0.035)	0.188 (0.208)	-0.019 (0.035)	-0.221 (0.224)	-0.071* (0.038)	-0.127 (0.261)
Wave 5 (missing)	-	-	-	-	-	-	-	-
Wave 6	0.131* (0.068)	0.119 (0.208)	-0.126*** (0.042)	-0.032 (0.199)	-0.063 (0.041)	-0.616** (0.309)	0.040 (0.045)	0.396 (0.259)

***: p<0.01; **: p<0.05; *: p<0.1

Note: The covariate whose effect is captured by the coefficients is the absolute difference in SA scores between nodes (“SAS distance” or “SAS heterophobia”).

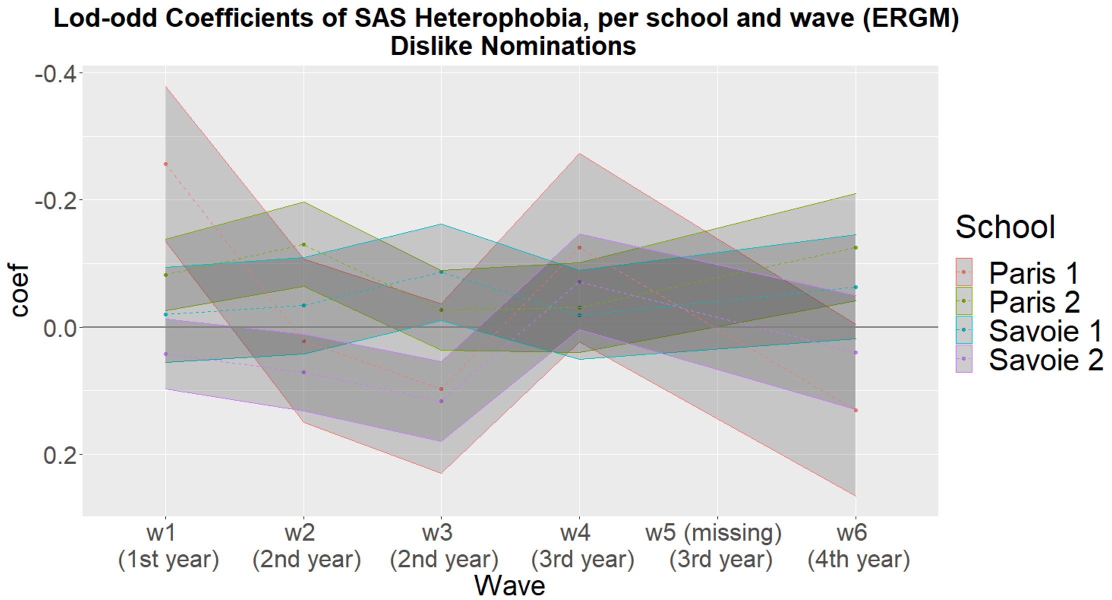
Parameters for density, emission and reception not shown. See Appendix 4C.

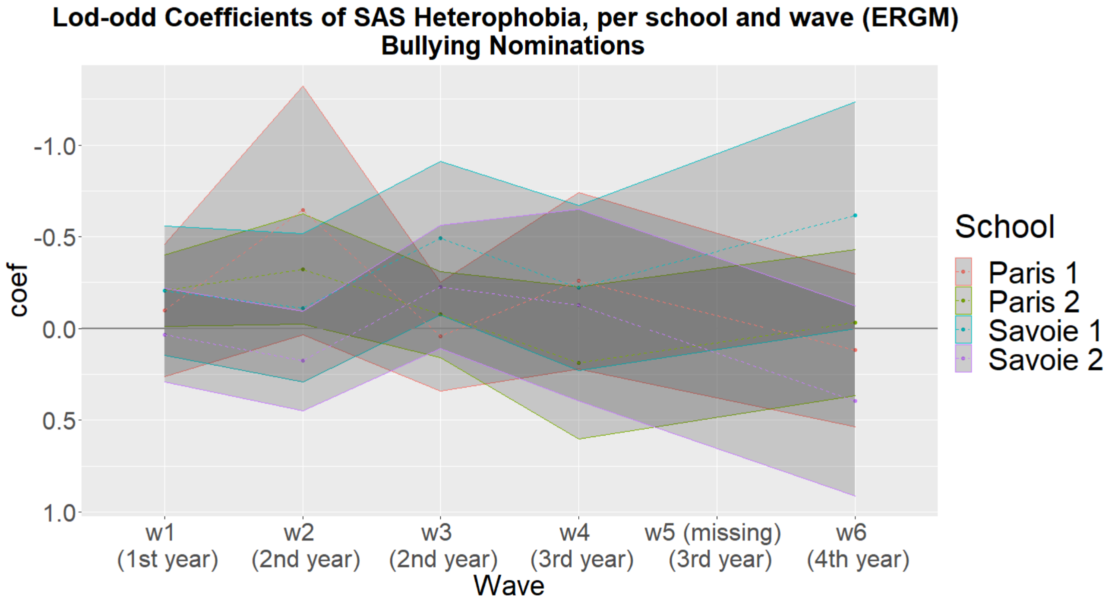
Figure 4-5: Evolution of Socioeconomic Heterophobia over Time (waves 1 to 6) – Dislike (top) and Bullying (bottom)

Note: Coefficients are indicated by dots and connected by dashed lines. 95% confidence intervals are marked by a shaded area and bounded by straight lines.

The y-axis scale is not the same on the two plots. Negative coefficients are placed upwsard. Numerically precise comparison of coefficients should not be made between different schools due to substantial differences in network size and density, but is reasonable for within-school time trends.

There is approximately 6 months between each wave. Wave 5 was not observed (covid-related lockdown).





In Paris 2 and Savoie 1, coefficients are either statistically insignificant and close to 0, or significant and negative, depending on the type of nomination and wave (e.g. bullying in Savoie 1 is negative and significant for waves 3 and 6 only). Therefore, there is no evidence that social distance increases the odds of conflict; on the contrary, there is moderate evidence that it reduces them, for certain waves at least.

In Savoie 2, there are significant and positive coefficients for disliking nominations at waves 2 and 3, which might indicate increased conflict as socioeconomic distance grows. However, the effect sizes are relatively small (at most 0.117, meaning 1.33 more chances of a nomination between socially far-away students[98]); there is no such effect in earlier nor later waves; nor is there one for bullying networks. Altogether then, this makes for rather inconclusive evidence.

Finally, Paris 1 shows an interesting pattern. In wave 1, there is a significant negative effect of a moderate size for the disliking network (1.87 less chance of a nomination among socially far-away nodes[99]). However, the effect disappears in waves 2 to 3, where it is non-significant and close to 0, and appears again in wave 4 but largely reduced, and significant at the 90% threshold only. Finally, in wave 6, a positive effect is observed which is statistically significant at the 90% threshold, but not the 95% one, and with a small effect size (1.38 more chance of a nomination among far-away nodes[100]). This may be taken as hinting to an initial pattern of socioeconomic homophobia that disappears after wave 1, and perhaps even of a slight heterophobic pattern that emerges by the end of the period. These results should be considered with care: the effect sizes are small bordering on significant, and the time trend is not clear as the coefficient for wave 4 does not really align with this alleged trend. Nevertheless, these results resonate with the previous finding that, in Paris 1, socioeconomic homophily in friendship networks strongly increased over time. This gradual segregation of the relational structure might have elicited increased tensions among students from different backgrounds, although the evidence at hand remains inconclusive. This calls for a deeper investigation into the processes at work in Paris 1, which we shall come back to in chapter 7.

Altogether, socioeconomic distance does not seem to matter nearly as much for negative ties as it does for positive ones. Unlike friendship networks, where socioeconomic homophily was consistently observed in all schools, disliking and bullying networks show no

98 For 2.44 points of SAS difference: exp(0.117*2.44) = 1.33.

99 For 2.44 points of SAS difference: 1/exp(-0.257*2.44) = 1.87.

100 For 2.44 points of SAS difference: exp(0.131*2.44) = 1.38.

clear nor consistent pattern across schools and over time. In Paris 2 and Savoie 1, socioeconomic distance actually seems to reduce the odds of conflict, though only at certain time points. As for Savoie 2 and Paris 1, some form of social heterophobia might occur in certain waves, but the evidence remains unclear and the estimated effect sizes are small[101]. Therefore, while socioeconomic homophily appears as a clear pattern of students' social life, the same cannot be said of socioeconomic heterophobia.

3. Comparison with other Homophilies

In the previous sections, the strength of socioeconomic homophily has been measured in two ways: using odd ratios, which provide a measure of relative chance that is quite straightforward to interpret, and by looking at the raw number of friends from distant backgrounds that a student has (exposure). However, these do not tell us whether the considered effect sizes are important in relative terms; that is, whether socioeconomic background matters more or less than other factors in shaping students' relations. Is socioeconomic homophily one of the strongest regularities that is to be observed in teenagers' relationships? Or is it merely a second-order pattern, occurring on top of deeper and stronger determinations? To answer this, it is necessary to compare the magnitude of socioeconomic homophily to that of other relational patterns. This section will consider homophily along three individual attributes besides socioeconomic background: gender, ethnicity, and academic results (grades). Additionally, the next section will investigate centrality and peer status, in order to see whether socioeconomic origin mostly shapes student friendship through homophily or through social hierarchies.

Gender and ethnicity, similar to social class, are ascribed characteristics that relate to macro-level dimensions of social stratification. There is an abundant sociological literature studying their impact on teenagers' friendship networks, which consistently found both of them to elicit homophily (see the review of chapter 1, section 1.1.3). In fact, if scholarly interest for a phenomenon tells anything about its substantial importance, then one would expect the homophilies of gender and ethnic background to outsize that of socioeconomic background, as the latter has been far less studied. This is indeed what previous works suggest: although not primarily interested in socioeconomic variables, most of them found

101 It is also possible to pool all the waves together and to estimate mean effects over the entire period, which masks time heterogeneity but increases statistical power. When doing so, a statistically significant negative effect (meaning homophobia) appears in Paris 2 for disliking and bullying and in Savoie 1 for bullying, all other estimates being non-significant (models not shown).

parental occupation to be a (much) worse predictor of friendships than both gender and ethnicity.

Academic results are also well-known as a source of homophily. Compared with other attributes, they are more directly related to the school context specifically, being the product of students' academic activity, as well as one of the few principles of differentiation among them that is explicitly recognized and enforced by the school institution (along with perhaps, discipline and curricular options). Furthermore, they are partly endogenous to students' sociability: friends can socialize each other into certain academic attitudes and thus results, although the strength of this influence effect is debated. This tends to foster homophily, as students not only befriend peers with similar grades, but also become similar to their friends in terms of grades.

In order to compare homophily on these attributes, I will rely on tie odds in friendship networks estimated with ERGMs. Additionally, measures of heterophobia are presented in Appendix 4D. For the most part, these attributes are not associated with increased conflict across groups, except for academic distance which sometimes entails higher odds of disliking (depending on the school and wave)[102].

3.1. Comparing Homophilies on Different Attributes

When comparing different homophilies in a given network, there are a number of issues to be considered. First, the corresponding attributes may be related to one another: in the present case, socioeconomic background, ethnicity, and grades are correlated in most schools. Likewise, gender is correlated to grades (girls have better results). Second, cross-model comparison of log-odd coefficients is at risk of being biased due to unobserved variance, as was discussed previously. For both these reasons, it is preferable to estimate each effect net of the others, which allows us to separate the different homophily effects and to make the estimated coefficients directly comparable. Note that this tends to "favor" gender and ethnic homophilies: they are generally less correlated to other modeled attributes, such that their estimates are less reduced when controlling for the other homophilies than those of grades and socioeconomic background are[103]. However, this does not affect my substantial

102 Moreover, for gender, disliking is usually more frequent among same-gender peers, especially starting in wave 2 (gender homophobia), coherent with the notion of "proximity dislikes" mentioned earlier.
Also, note that I do not consider measures of exposure for these variables: it would be too difficult to establish classes of equivalence between categorical variables and continuous ones, so the comparison would not be very informative (see the discussion of classes of equivalence below).

103 To the extent that academic results are partly caused by students' socioeconomic background, measuring socioeconomic homophily net of grades homophily can be seen as under-estimating its actual effect (grades

conclusions, as univariate and multivariate estimates are fairly close anyway in the present case (additional models with univariate estimates not shown).

Another problem is that the different homophilies are expressed through different metrics, which means that ERGM coefficients are not directly comparable. Gender and ethnicity are categorical variables, meaning that homophily is defined by exact matching: log-odd coefficients give the change in tie odds when the two nodes belong to the same gender or ethnic category. On the contrary, grades and SA scores are continuous variables, which, as we saw, means that log-odd coefficients express the change in tie odds for each additional point of absolute difference between nodes. Furthermore, in the case of ethnic homophily specifically, there is heterogeneity across ethnic groups: the propensity of, say, native students to over-nominate one another is not necessarily the same as that of students with an Asian or African background. This further implies that homophily estimates are sensitive to the exact categorization that is used (e.g. should immigrants from North-Africa be grouped together with those from Sub-Saharan Africa?).

The issue boils down to finding, for each individual attribute, dyadic configurations that can be deemed equivalent to one another. What difference in grades is the difference between a boy and a girl “worth”? What amount of socioeconomic distance is equivalent to the difference between two ethnic backgrounds? These questions do not have a single, true answer, but are instrumental to performing a reasonable comparison.

For SA scores, similar to what was done in previous sections, I will consider the fictive case of two students whose absolute difference is equal to 2.44, which is the average distance between the child of a worker and of a manager[104]. Grades are standardized per school and wave, meaning that the measure has a standard deviation equal to 1: this is practically equal to the standard deviation of SA scores (0.98 in the sample), such that the two measures can reasonably be compared. Therefore, I also consider a 2.44 difference for grades. For gender, there is only one coefficient that can be considered, which is that of being from the same gender. Finally, for ethnic background, the individual attribute has 13 categories (cf. chapter 3 section 1.2.5). I consider the dyads that belong to the same category, relative to those that do not. This masks the fact that there is some heterogeneity in the homophily of these different groups, mostly because small ones have structurally little opportunities for

homophily is a mediator through which socioeconomic background impacts relationships). The same could be said for gender and ethnicity, although the effect of the latter on grades is much smaller (Ichou 2018).

104 To be clear, the model is estimated with a continuous effect, similar to what has been done throughout the section: it is only when presenting results that I will give the predicted effect for a 2.44 difference.

homophilic bonding. However, in practice, ERGM coefficients are mostly impacted by the three largest groups, which are also the most homophilic ones (again, for structural reasons): natives, North-African backgrounds, and Sub-Saharan African backgrounds. There may also be different social distances among ethnic groups (e.g. different immigrant origins may be closer from one another than from natives), but this is difficult to model in ERGM, as it cannot be captured by matching covariates nor continuous ones. Therefore, I only consider exact ethnic matching[105].

A final issue that arises when comparing homophilies on different attributes is that the corresponding measures can be unequally noisy. Gender and academic performance are almost exact measures, as they correspond to institutionalized attributes that are provided by the schools, and which are known to students in almost the same way as they are to the observer. On the contrary, ethnic and socioeconomic backgrounds are measured through proxy variables (family name and parental occupation), thus necessarily noisy. As a consequence, it is likely the case that ethnic and socioeconomic estimates are slightly under-estimated relative to gender and academic ones.

3.2. Results

For each school and wave, for the networks of friends and very good friends separately, an ERGM that includes a homophily parameter for each of the four attributes is fitted, along with a density parameter, and emission and reception parameters for each attribute. The estimated homophily coefficients are presented in Figure 4-6. The complete model outputs are in Appendix 4C.

Remember that the log-odd coefficients for grades and SA scores are multiplied by 2.44 in the plots and their sign is inverted, such that homophily is indicated by positive coefficients for all attributes. Moreover, in the case of Paris 2, grades are only available at wave 1 (cf. chapter 3 section 1.2.4).

105 In theory, one may introduce parameters that consider the log-odd change associated to specific dyadic configurations; for example, native to North-Africa, or Sub-Saharan Africa to North Africa. This could be done with the “nodemix” term in ERGM terminology. However, this raises identification issues, because such mixed-ethnic dyads are used by the model to estimate emission and reception effects for the different ethnic groups. This is a generic issue with how categorical attributes are modeled in ERGM: cross-category ties can be seen as indicating different levels of social activity and popularity of the groups (which is equivalent to assuming that there are no latent distances underlying the group structure, i.e. the attribute is truly categorical and not ordinal), but also as indicating relative proximity among the groups. The latter is probably more realistic for ethnic groups, but it is also less elegant and harder to interpret from a modeling perspective, which is why the former option is often preferred in the literature on ethnic homophily. Given that my aim here is not to investigate ethnic homophily *per se*, but simply to put socioeconomic homophily in perspective, I chose to stick with the traditional approach, which means treating mixed dyads as indicative of social activity and popularity rather than of differential proximity across groups.

The main finding of Figure 4-6 is that gender homophily generally appears as the strongest out of the four, with academic, ethnic and socioeconomic homophilies having smaller and somehow close effect sizes. The one exception to this pattern is Paris 1, where the drastic rise in socioeconomic homophily over time makes it the strongest homophily by the end of the period.

In Paris 2, Savoie 1 and Savoie 2, gender homophily exhibits log-odd coefficients from approximately 0.8 to 1.2 for friends, and from 1.2 to 1.8 for very good friends (there is a noticeable decrease over time for Savoie 1 and Savoie 2, resulting in lower figures in the latest waves). This means nominations between same-gender peers are at least 2, and up to 6 times more likely than among different-gender nominations[106]. By comparison, most other coefficients are around 0.5 at most (odd-ratio = 1.65). Notable exceptions to this are ethnic homophily in Savoie 1's networks of very good friendships, which peaks at 0.731 in wave 3 (odd-ratio = 2.08), as well as academic homophily in that same school which strongly increases in waves 4 and 6, reaching a coefficient of almost 1 in this later wave (odd-ratio=2.70). This however remains far less than the corresponding effect of gender at the same waves (1.41 in wave 3, odd ratio = 4.10).

As for the relative ordering of academic, ethnic, and socioeconomic homophilies, this differs across schools. In Paris 2, the effect sizes are extremely close for ethnicity and SA scores; that is, the odds of a friendship between two students from the same ethnic background, relative to different-background ones, are about the same as those of students with the same socioeconomic background, relative to two students from working-class and upper-class backgrounds (between 1.38 and 1.70 more chances, depending on the wave[107]). In Savoie 1 and Savoie 2, socioeconomic homophily appears to be lower than ethnic homophily over most of the period, but almost catches-up in wave 6. As for academic homophily, it is roughly equivalent to ethnic homophily (thus higher than socioeconomic homophily) in the early waves in Savoie 1; however, it strongly increases in waves 4 and 6 for unknown reasons, reaching a much higher level than ethnic and socioeconomic homophily (though still lower than gender homophily). In Savoie 2, academic homophily is of approximately the same size as ethnic homophily – perhaps slightly larger as well. On the contrary, looking only at the first wave in Paris 2 (the only wave where grades were available), academic homophily appears substantially lower than in the other schools, and is not significantly different from 0.

106 exp(0,8) = 2.23 ; exp(1,2) = 3.32 ; exp(1.8) = 6.05.

107 Meaning log-odds from 0.32 to 0.53.

In Paris 1, the situation in the first wave is similar to that of the other schools (ethnicity, grades and SAS homophily at about the same level, gender homophily above the rest). However, socioeconomic homophily strongly rises over waves, coherent with what we saw in previous sections, while the other homophilies remain approximately constant (there is a decrease of gender and academic homophily in friendship networks, but not really in good friendship ones). Thus, in wave 6, socioeconomic background matters roughly as much as gender for “very good friend” nominations, and considerably more for “friend” ones (2.26 more chances for a same-gender tie against 3.97 for a same-background one[108]).

108 Log-odds of 0.817 and 1.38 respectively.

Figure 4-6a: Evolution of Socioeconomic, Ethnic, Grades and Gender Homophilies over Time (waves 1 to 6) – Networks of Friends

Note: Coefficients are indicated by dots and connected by dashed lines. 95% confidence intervals are marked by a shaded area. The coefficients of grades and SAS have been multiplied by -2.44.

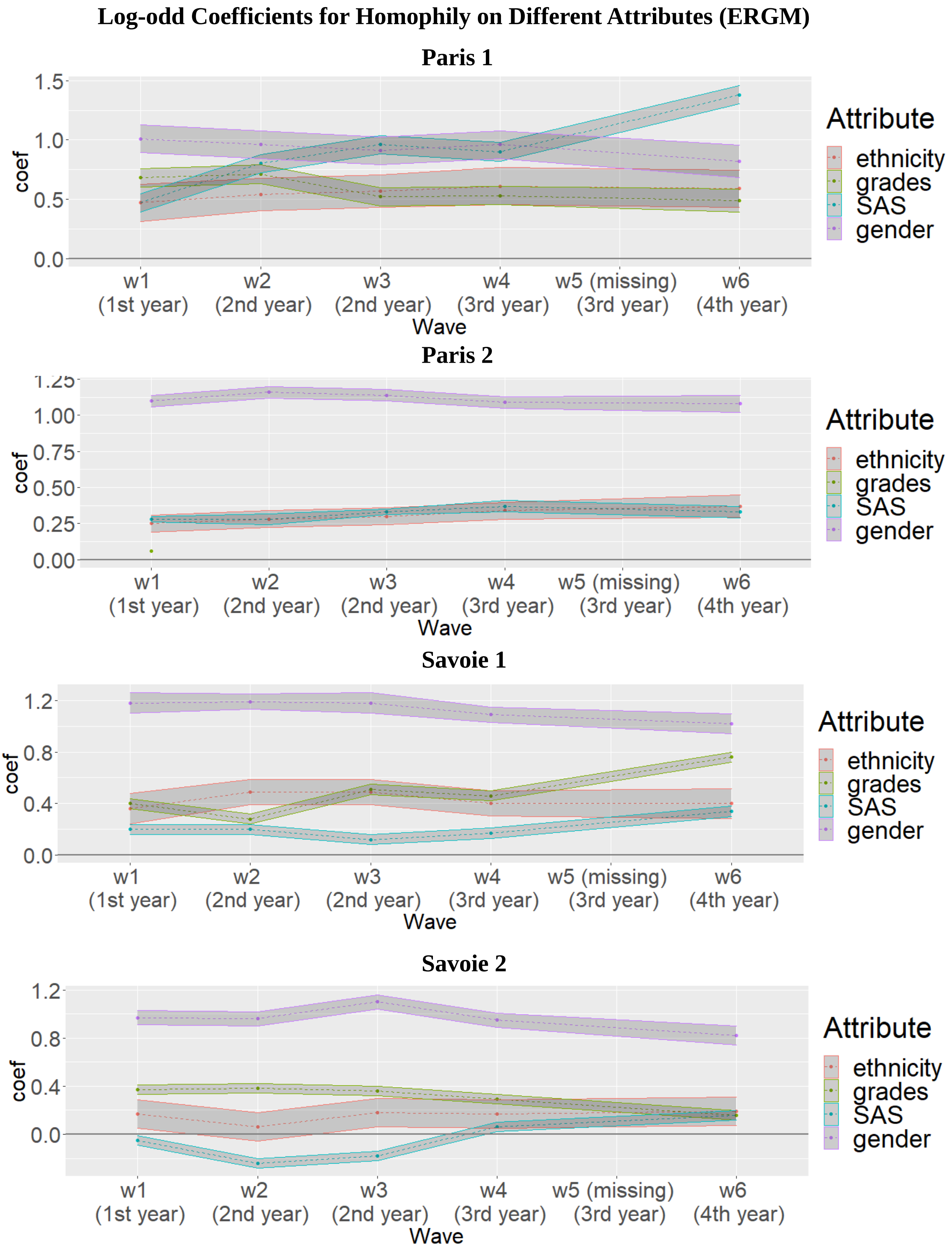
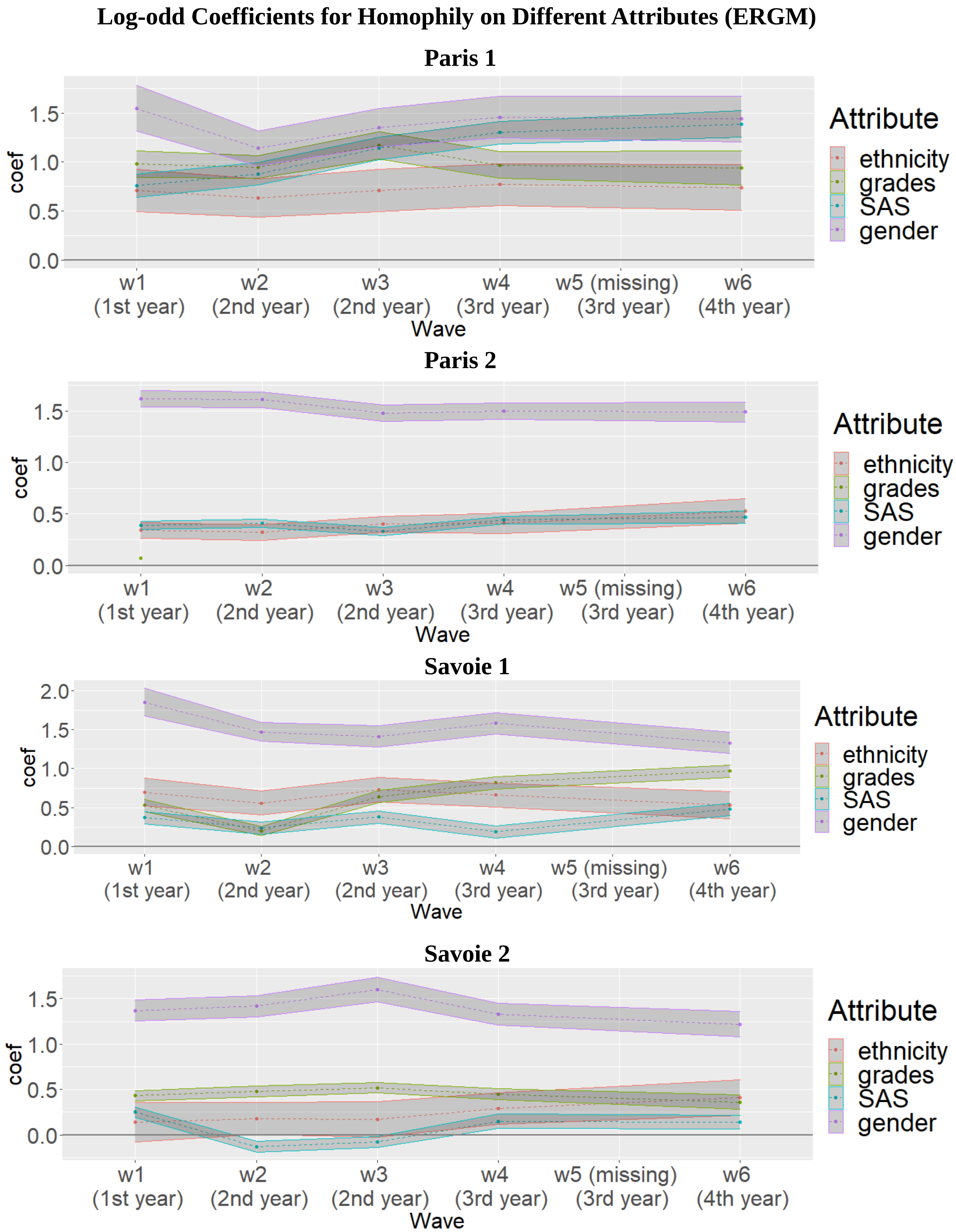


Figure 4-6b: Evolution of Socioeconomic, Ethnic, Grades and Gender Homophilies over Time (waves 1 to 6) – Networks of Very Good Friends

Note: Coefficients are indicated by dots and connected by dashed lines. 95% confidence intervals are marked by a shaded area. The coefficients of grades and SAS have been multiplied by -2.44.



4. Network Centrality and Peer Status

To conclude this chapter, let us briefly consider the impact of socioeconomic origin on friendship networks along their vertical dimension. Are peers from certain backgrounds more central than others? Are they regarded as more prestigious, popular or well-liked than their peers? As we saw in the literature review of chapter 1, a few studies consider that, when it comes to socioeconomic background, hierarchy prevails over homophily, and that upper-background students are more prestigious whereas lower-background students are at a higher risk of marginalization (section 1.1.3).

To see if this is the case in our four schools, we can refer to Table 4-2 above. The estimated ERGMs include four parameters: density, socioeconomic homophily, socioeconomic emission and socioeconomic reception. This latter term captures the propensity of students to receive more or fewer ties as their socioeconomic background increases, which is a straightforward way to conceptualize centrality in the friendship networks. The corresponding coefficient can be compared to that of socioeconomic homophily to compare the two effects. Looking back on Table 4-3, it can be seen that the receiver effect is generally small and statistically insignificant in three of the schools, but not in Savoie 2. In this school only, high background students are both on average more active (emission) and more popular (reception).

One way to represent this in a clearer fashion than Table 4-3 is to use the log-odd coefficients estimated by ERGM to predict the probability of different types of ties. This is shown in Figure 4-7 for the networks of very good friends (wave 1). In each school, three fictive students are considered, with Socio-Academic Scores of -1.8, 0.3 and 2.4 (which are the minimum, mean and maximum of the SAS range in the sample). For each of them, a curve indicates the probability that the student has of receiving a friendship nomination (y-axis) depending on the SAS of the student that would send the nomination (x-axis). These predictions use all the coefficients from the ERGMs of Table 4-3: the probability of the ties depend jointly on homophily, emission and reception patterns, depending on the SAS of both the sender and receiver[109].

109 For a sender i and a receiver j, and an ERGM with four estimated coefficients (d,h,e,r) for density, SAS homophily, SAS emission and SAS reception, the log of a tie occurring is given by:

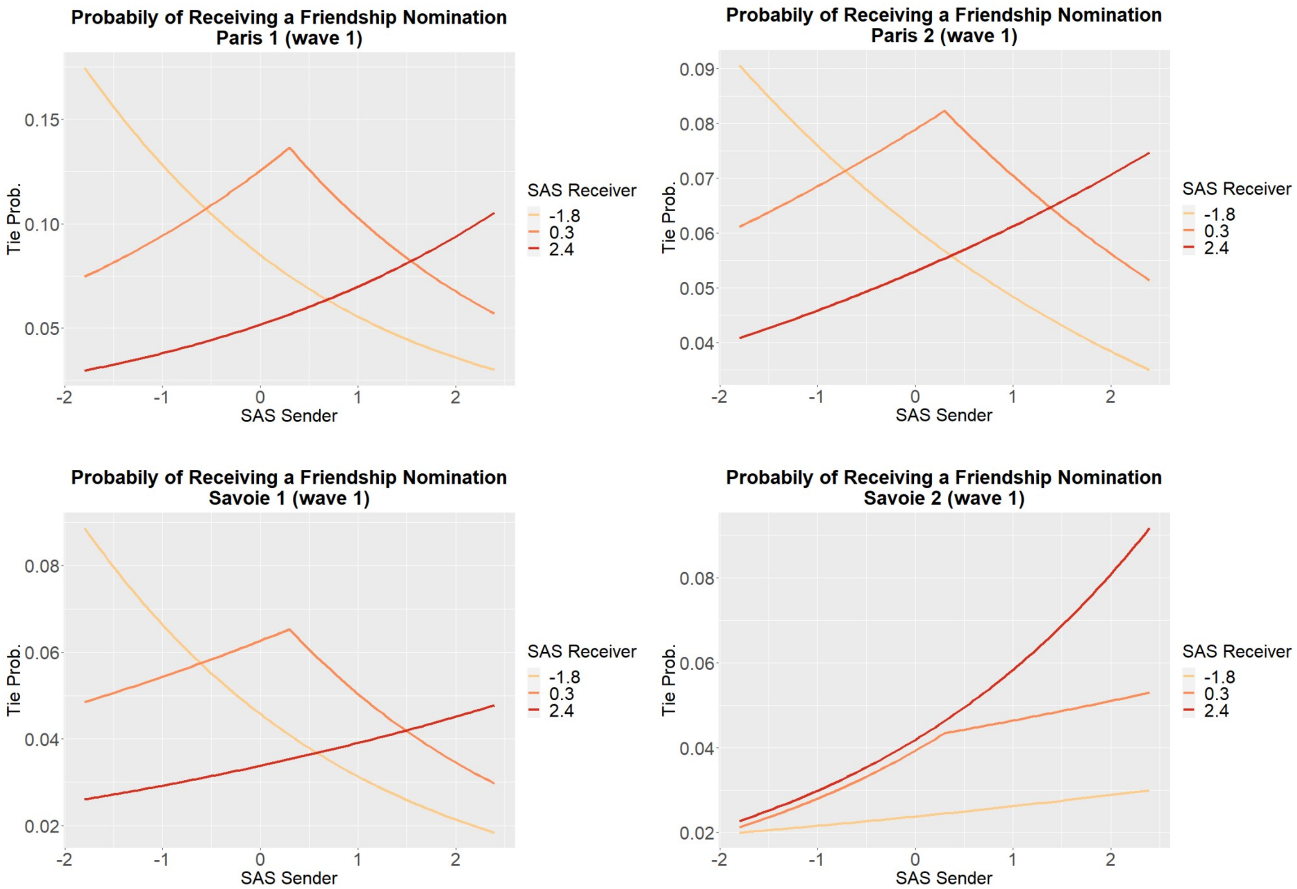
log_tie = d + e*SAS(i) + r*SAS(j) + h*| SAS(i) - SAS(j) |

The probability of the tie is then computed as log_tie / (1+log_tie).

Figure 4-7: Probability of Receiving a nomination as "Very Good Friend" Depending on the Socioeconomic Background of the Sender and Receiver (ERGM) – Wave 1

Note: predictions derived from the ERGM of Table 4-3 (four parameters included: density, SAS homophily, SAS sender and SAS receiver).

The x-axis represents the Socio-Academic Score of the sender of a nomination, and the 3 lines correspond to 3 fictive receivers with different SA Scores (-1.8 and 2.4 are the sample range of SAS, and 0.3 the middle of that range).



Whether homophily or hierarchy prevails can be seen by looking at the shape of the three curves. In a primarily homophilic network, the probability of a tie depends more on the similarity between the sender and the receiver than on their individual attributes. Therefore, the probability of receiving a tie should peak when the sender and receiver have the same value of the attribute, and the curves of upper- and lower-background students should intersect as the background of the sender changes. This is typically the pattern observed for Paris 1, Paris 2 and Savoie 1. On the contrary, in a primarily hierarchical network, the probability of a tie mostly depends on the attribute of the receiver. In Savoie 2, the curves never intersect: high-background students are more likely to receive a tie no matter what the background of the sender is (though the gap grows larger as this background increases).

Therefore, the one school where socioeconomic homophily is weak, Savoie 2, is also the one where upper-background students have more friends than lower-background ones. In the other three schools, there seems to be little difference in the centrality of socioeconomic groups on friendship networks, and homophily is clearly the dominant pattern[110]. Note that Figure 4-7 only shows results for the networks of very good friends in wave 1. Results for later waves and other friendship ties are in Appendix 4E: the conclusions remain the same.

A second question that was raised in the literature review is that of the relation between socioeconomic origin, academic achievement, and peer status. Many studies found high academic performances to be associated with lower status among classmates. Given the association of grades and socioeconomic background, this may be expected to contribute to the popularity of lower-background students on average – and contrary to the expectations of another part of the literature, for whom lower-background is associated with lower status. Alternatively, peer status may be essentially dependent on conformity to group norms, in which case students may be rejected if they are scholarly deviant, but also if they are overly zealous regarding academic rules.

Since this is more distantly related to the objectives of this chapter, I simply summarize the results here: complete analyses are in Appendix 4F. In short, academic results shape students' friendships in both horizontal and vertical terms: there is academic homophily and academic centrality at the same time. In Paris 1 and Paris 2, higher grades are associated with less friendship nominations, both received and emitted. This does not really entail a statutory penalty for upper-background students, however, even though they tend to have better grades: first, because the effect sizes are not strong enough relative to the correlation of socioeconomic origin and grades, and second, because the negative association of grades and peer status is weaker among upper-background students than among lower-background ones (i.e. there is a negative interaction between the popularity effects of grades and of socioeconomic background). This is something that we will come back to in chapter 6, so we can put it aside for now. In Savoie 1, things are slightly more complex: it seems that well-graded students emit fewer nominations, but do not necessarily receive fewer. As a result, there is not much of an impact on the indegree centrality of students depending on their socioeconomic origin. Finally, in Savoie 2, the pattern is reversed: high grades are associated

110 Interestingly, this is also the case for the nominations of popularity. In the questionnaires, certain items asked students about the peers that they thought were "popular", "well-liked by many people" or "known by many people". For the most part, these items were not used in the present dissertation (which is why I have not introduced them in chapter 3), but these nominations also exhibit patterns of socioeconomic homophily, much more so than socioeconomic centralization. See Appendix 3C.

with more central positions, both in terms of received and emitted nominations. This is coherent with the pattern observed for socioeconomic origin, and in fact, partly explains it. Upper-background students tend to receive more friendship nominations because they have higher grades (i.e. a large part of the popularity effect of socioeconomic background is "explained away" by academic popularity in a multivariate ERGM).

As for the possibility of a curvilinear pattern, with either academic achievement or socioeconomic background being associated to more friendship ties in the middle of the distribution and fewer at the extremes, there is no clear evidence to support this idea[111]. This does not necessarily mean that the effect does not exist, because the models that I use remain relatively rough, and various suppressing effects could blur the relationship of interest; but, at the very least, this pattern, if it exists, is neither obvious nor strikingly strong. In any case, remember that all the analyses of this section pertain to the aggregated state of friendship networks, not to the relational processes that underlie it. The fact that students with certain attributes are more central in friendship networks does not imply that these attributes are a source of prestige among peers, similar to the fact that homophily does not imply homophilic selection.

Conclusion

In this section, I have proposed a broad description of the schools' relational structures in regard to socioeconomic distance among students. The first thing that should be noted is that the situation is not identical in the four schools: although certain features are to be found in most of them, they also strikingly differ on essential aspects. The conclusion is therefore organized in two steps: I first try to draw a global picture that is relatively coherent across schools, and then I discuss the deviations from this picture that are observed locally.

Altogether, three main transversal features emerged. First, socioeconomic homophily goes up with the strength of ties. Weak friendship ties exhibit small amounts of homophily, while very good and out-of-school friendships are much more clearly impacted by socioeconomic origin. Second, the overall level of homophily appears to be of a moderate magnitude. While socioeconomic homophily certainly exists, it does not result in complete segregation, and cross-group ties are very much possible. In particular, almost all students have at least a few weak ties to peers from different backgrounds to their own, and most of them also have some strong ties toward these peers. In that sense, most students experience

111 Adding covariates for either squared SAS or squared academic achievement in the ERGMs does not reveal any consistent pattern (models not shown).

socioeconomic mixing in their sociability circles. Moreover, as the analysis of section 3 has shown, this level of homophily is close to that of academic performance and ethnic background – perhaps even a bit lower in some cases – and clearly inferior to that of gender. Finally, socioeconomic distance does not increase the probability of disliking ties and, in some cases, even reduces it. While this does not mean that there are no tensions among students that arise as a result of social distance (as a matter of fact, there are; see chapter 6), it does show that socioeconomic background is not the main force shaping students' animosities. The level of antagonism among social groups, while not necessarily null, seems neither strong nor systematic enough to induce a polarization of the relational structure (which is also reflected in the fair amount of mixing observed in weak friendship ties).

Condensing these traits into an ideal-type, I propose qualifying the situation of the schools as being characterized by *positive contact*, *homophilic bonding* and *peaceful cohabitation* among socioeconomic groups. *Positive contact* refers to the fact that most students do enter into positive relationships with others from a different background. However, it should be clear that many of these relationships are relatively superficial; at a deeper level, considering stronger relations and, perhaps more importantly, relations that stretch out of the school onto exterior sociability circles, social class boundaries appear significantly stronger. Following this logic of homophily growing larger among strong ties, it is likely the case that socioeconomic origin would matter even more if we could look at yet more demanding relationships: in particular, one might expect the friendships that survive middle school, and that will be carried on to later stages of students' lives, to be particularly homophilic. It is in that sense that the relational structure is marked by *homophilic bonding*: as much as *exposure* largely overcomes social boundaries, intimate *bounding* among students remains impacted by socioeconomic origin. Finally, the expression *peaceful cohabitation* is meant to underline the fact that socioeconomic distances, although clearly active in shaping students' sociability, do not induce systematic open conflict among students.

This ideal-type applies well to Paris 2 and Savoie 1, arguably the two intermediate cases of the sample. Paris 1 and Savoie 2, however, deviate from it for different reasons.

In Paris 1, the three traits can be observed in wave 1. Even though socioeconomic homophily is already higher than in other schools, this is due to the greater socioeconomic diversity of the population, not to the fact that student friendships are more homogeneous; and many weak ties still connect the different socioeconomic groups. However, over time, this homophily sharply increases, resulting in a situation of strong relational segregation by the end of middle school. Socioeconomic background thus becomes the strongest factor of homophily in the school's friendship networks, well above the homophilies of academic performance, ethnic background and even gender. The impact of socioeconomic origin on negative ties still appears to be low, but there are hints of increasing tensions as well, although this should be considered with care. In general, I do believe that the situation in later waves is still marked by *peaceful cohabitation*, but the notion of *positive contact* does not seem appropriate anymore.

Savoie 2, on the other hand, is singled out by remarkably low levels of socioeconomic homophily, that almost disappear at certain waves. Moreover, the gradation of this homophily with the strength of ties is not as pronounced as it is in other schools: in-school very good friends are virtually as homophilic as out-of-school ones (which does not mean much). Moreover, Savoie 2 is the only school where a pronounced popularity pattern can be observed in relation to socioeconomic origin, with higher-background students being more central in friendship networks.

This panorama leaves us with many questions. How can we explain the common patterns observed across schools? And can this explanation also account for the differences across schools, particularly in the two extremes that Paris 1 and Savoie 2 seem to represent? This will be the subject of the next three chapters.

Chapter 5: Where Does Homophily Come From? Testing for the Impact of Low-Order Relational Processes

In the previous chapter, we saw that there is socioeconomic homophily in friendship networks, in a descriptive sense: ties are more likely to occur among students from similar backgrounds. The next chapters will try to explain this pattern; that is, to understand which social processes led to the homophily that we observe.

As was discussed in chapter 2, there are a number of relational processes that might lead to homophilic network structures. I proposed classifying these in four categories: imposed foci, chosen foci, expressed selective dispositions and network endogenous processes. Imposed and chosen foci impact friendship formation through propinquity, that is, differential contact frequencies among students; expressed selective dispositions through the homophilic selection of similar peers (also referred to as 'assortative mixing'); and network mechanisms through a combination of both. Moreover, each of these processes can be understood as the joint outcome of contextual and dispositional factors, that is, properties of the context and properties of students themselves. However, in empirical settings, we never get to observe these processes directly, at least not through quantitative network measures (the contribution of qualitative methods will be discussed in the next chapter). Instead, we ought to reconstitute their impact based on the outcome that they produce, reasoning backward from the observed friendship networks. That is the aim of the present chapter.

In order to do this, it is necessary to proceed following two steps. First, from the observed networks, we can try to infer the underlying relational processes that guide friendship formation among students. This is done by means of dedicated statistical models, which I will present in a moment. However, these models take the entire network structure as their dependent variable; in other words, they are meant to explain friendship formation in general, not homophily specifically. Therefore, a second step is needed, which focuses on the impact of the identified processes on socioeconomic homophily specifically.

The chapter is organized into two sections, following these two steps. Section 1 starts with a discussion of model-based approaches in network analysis (1.1.), then moves on to presenting the model specifications (1.2.) and the results of the estimation (1.3.). In section 2, I introduce *contribution scores*, a simulation-based method meant to assess the impact of relational processes on homophily specifically (2.1.), which I then apply to the previously estimated models (2.2.).

This entire chapter focuses exclusively on the networks of "very good friend" nominations, though I will speak of "friendship networks" for simplicity. The reason is that this is the type of nomination that contains the most relevant information in terms of relational structure. Mere "friend" nominations are too numerous, which means the measure captures relatively weak ties and is rather noisy in general. The high density of the networks also makes it harder to estimate complex models. On the other hand, the nominations of the five friends that the respondent spent the most time with also raise modeling issues, due to the upper-bound placed on the number of nominations. As for out-of-school ties, they were not observed over a sufficiently long time-period.

1. Modeling the Evolution of Friendship Networks over Time

1.1. Generative Models in Network Analysis

1.1.1. Generic Rationale of Network Modeling

There are two types of network models that are commonly used for the study of friendship networks: the Exponential Random Graph Model (ERGM, Robins et al. 2007) and the Stochastic Actor-Oriented Model (SAOM, Snijders 2017). Although there is a number of differences between those (see Block, Hollway, et al. 2019; Block, Stadtfeld, and Snijders 2019; Leifeld and Cranmer 2016), the basic rationale behind them is the same. Therefore, the following discussion applies to both types of models equally.

Network analysts are often interested in how an observed macro-structure (i.e. a property of the network as a whole) can emerge out of different micro-level processes (i.e. rules of tie formation at the node or tie level). This is particularly true for friendship networks, as they are typically assumed to be emergent from low-order social processes. Indeed, when forming friendship relations, individuals mostly consider their immediate environment; and strategic action, though it may exist, is probably over-shadowed by affective attachment or compliance to social norms. Furthermore, in the case of school students, there should not be any selective pressure exerted at the macro-level on the shape of the network[112]. As a result, it

112 As a counter example, Lazega (1999) showed how friendship ties among business lawyers were essential to their social capital and professional resources. In such cases, one may consider that the overall structure of the network (e.g. how sparse or how centralized it is) is itself important in understanding its emergence, from a functionalist perspective. On the other hand, it seems unlikely that this would be the case for the

seems reasonable to model friendship formation as occurring within low-order structures: nodes form ties within dyads, triads, small groups or any other local configurations, but they do not "*cast their gaze across the entire network*" (Robins, Pattison, and Woolcock 2005, p.2). The aggregate network structure is thus seen as being the contingent outcome of local processes, rather than a functional necessity.

This view stands at the core of both the ERGM and SAOM frameworks[113]. Indeed, identifying the social processes that underlie an observed network requires the use of multivariate quantitative modeling. In one of the articles that introduced ERGM, Robins et al. (2007) clearly laid out the issue:

> "*Sometimes, different social processes may make similar qualitative predictions about network structures and it is only through careful quantitative modeling that the differences in predictions can be evaluated. For instance, clustering in networks might emerge from endogenous (self-organizing) structural effects (e.g., structural balance), or through node-level effects (e.g., homophily*[114]*). To decide between the two alternatives requires a model that incorporates both effects and then assesses the relative contribution of each.*"

Both ERGM and SAOM are an answer to this problem. They are simulation-based models that aim to find the combination of micro-level processes most likely to generate the observed macro-structure. Still in the words of Robins et al., ERGM and SAOM try to explore "*how localized social processes and structures combine to form global network patterns, and of whether such localized processes are sufficient to explain global network properties*".

In order to do this, the basic rationale of the models is the following. The network's aggregated structure is described by a number of statistics, referred to as *sufficient* or *target* statistics. It is the user that chooses what statistics to specify, knowing that each statistic will have one corresponding effect on the model. Crucially, a statistic is always a count of low-order configurations: for example, the number of closed triplets in the network, the number of reciprocated ties, or the number of same-gender ties. The estimation procedure will then try to find the conditional probabilities of each of these low-order configurations, i.e. how likely we

friendship networks of school students.

113 Leifeld and Cranmer (2016) discuss how other networks may have different properties, such as networks of international diplomatic alliances, where actors are highly rational, forward looking and consider the entire network structure before performing an action. They suggest that ERGM may be more appropriate than SAOM for these types of networks. Their argument, however, has been strongly criticized, and there is currently no clear consensus on the matter (Block, Hollway, et al. 2019).

114 Here the authors use "homophily" in the sense of "homophilic selection" (i.e. the low-order process, not the aggregate outcome).

are to observe each one net of all the others. For example, if a network is described in terms of its total number of edges (density), number of same-gender edges (gender homophily) and number of closed triplets (transitivity), the model will try to estimate how likely any edge, same-gender edges, and transitive edges are, such that these probabilities *jointly* predict the correct aggregate values of the statistics.

More precisely, each statistic has a corresponding log-odd *parameter*, that gives the change in the odds of observing a tie when it contributes to forming the low-order configuration of interest (e.g. closing a triangle). The value of these parameters is unknown at first and is estimated through a Markov Chain Monte Carlo (MCMC) algorithm. At each step, the algorithm draws certain parameter values and simulates fictive networks from these, compares the statistics in these networks to the observed ones, adjusts the parameters accordingly, and re-iterates until convergence is reached, i.e. the simulated statistics are similar enough to the observed ones. It is in that sense that the model can be called “generative”, as it relies on simulations to find the set of parameters that best predict the observed structure of the network. This is key to accounting for the dependency of network observations, as the state of the network at a given point during the simulation process is taken into account when assessing tie probabilities and simulating further networks (granted that some of the model parameters capture tie dependencies, for e.g. reciprocity or transitivity terms).

As we saw, each effect specified into the model by the user has a macro and a micro side: the statistic (observed) and the parameter (estimated). This is very valuable theoretically, because this micro-macro form of the model fits well with substantive interpretations of micro processes and macro outcomes: *“one can interpret the structural features of the network as resulting from biases towards particular types of tie-configurations in a hypothetical network formation process”* (Block et al. 2018). However, it is not necessarily the case that parameters represent realistic social processes: model-wise, they are simply indicating the over- or under-representation of certain network structures, relative to what would be expected by chance, and controlling for other modeled effects. It is therefore an interpretative choice to treat those as indicative of actual social processes (*ibid*)[115]. In the empirical sections below, I will try to indicate when that interpretation is reasonable, and when it should be questioned.

1.1.2. Classes of Models: ERGM and SAOM

Everything mentioned above roughly applies to ERGM and SAOM equally. However, there are important differences between these two types of models. SAOM tends to fit the objectives of this chapter more than ERGM does: this leads me to prioritize SAOM in the present chapter, with ERGM only being used in the appendices for robustness checks. In this section, I explain the reasons for this choice, through a short discussion of the differences between the two classes of models. Detailed presentations of the models, including their mathematical specifications, can be found in Robins et al. (2007) for ERGM and Snijders (2017) for SAOM. The *Rsiena* manual, which is the manual for the R package implementing SAOM, also offers a very pedagogical presentation of this class of models (Ripley and Snijders 2020).

The Exponential Random-Graph Model (ERGM) is a cross-sectional model that aims to model the state of a network observed at a single point in time. By contrast, the Stochastic Actor-Oriented Model (SAOM) is a longitudinal model that takes the first observation for granted and models the evolution of the network from that first point onward. Note that it is also possible to extend ERGM longitudinally, resulting in what is called a TERGM (Temporal ERGM); however, I will only resort to SAOM for longitudinal modeling, for the same reasons as I favor it over ERGM in general.

ERGM can be seen as a logit model performed at the level of the dyad (either directed or undirected). As explained above, the model is defined by a number of target statistics, which are count of tie configurations in the network (e.g. number of same-gender ties or number of closed triangles). Therefore, any dyad can be attributed, for each target statistic, a *change statistic*, which is defined as the change in the value of the target statistic when that dyad is toggled from 0 to 1, conditional on the rest of the network. These change statistics are effectively the model's independent variable; and the log-odd parameters give the odds of observing a tie as a function of its change statistics. For example, let us imagine that a gender homophily term is added to the model. The sufficient statistic is the number of same-gender

115 This typically demands additional assumptions, notably that the network is in an equilibrium state in the case of ERGM (*ibid*). Most effects are also assumed to be homogeneous over all nodes, though this can be mitigated by using interactions effects. Moreover, relevant interpretations are different between cross-sectional ERGMs, temporal ERGMs, and SAOMs. Block et al. (2018) argued that cross-sectional ERGMs and SAOMs can be interpreted in terms of micro-level processes, but not temporal ERGMs (which is still debated among network researchers).

ties in the network, such that any same-gender dyad will have a change statistic of 1, and any different-gender dyad a change statistic of 0. The log-odd parameter will thus indicate how much more likely a tie is when its change statistic is 1 (thus a same-gender dyad) rather than 0 (different-gender dyad), all other things being equal. The simulation process described above then adjusts the value of all parameters so that networks drawn from the corresponding probabilities have the right values of the target statistics on average[116].

Compared with ERGM, SAOM has two important particularities. First, it is actor-oriented (thus its name), which means that tie configurations are always considered from the perspective of a focal node. Second, it is time-continuous, meaning that network evolution is modeled as a series of “mini-steps” in-between the discrete observations. In practice, the simulation process for SAOM works as follows. Between two observations, the model simulates a large number of mini-steps. At each mini-step, a node is selected at random, and it is asked to perform an action. The node – given its current position in the network – can “decide” to send a tie, to delete one of its outgoing ties, or to do nothing. The probabilities for these different actions are determined by a multinomial logit model that uses the parameters’ value (e.g. log-odd change when a tie is part of a triangle, when it is directed toward a same-gender node, etc.). A rate parameter determines the number of the mini-steps that are performed, which is also one of the parameters that the model estimates, so as to reach the correct amount of tie change in-between observations. The estimation process is then similar to that of ERGM, with the value of the parameters being updated iteratively until the simulation yields fictive networks with the correct values of the target statistics.

Therefore, there are important differences between ERGM and SAOM. These are discussed in detail in Block et al. (2018); Block, Hollway, et al. (2019); and Leifeld and Cranmer (2016). Here I will only mention a few points of particular interest. First, the actor-oriented dimension of the SAOM is not necessarily as decisive as one may think. It is primarily a way to “organize” the changes that operate in the network for the simulation algorithm but is not fundamentally different from ERGM’s focus on ties rather than nodes (Leifeld and Cranmer 2016). Consequently, the fact that the nodes “decide” to send ties in the simulation does not imply that this is the way in which real-life ties are formed; even though SAOM aligns well with individualistic or agency-based theories, it is not a formal

116 Note that in cases where all the target statistics are *dyad-independent*, i.e. the change statistic of a tie does not depend on other ties in the network, then this is equivalent to running a classical logistic regression (this was the case for the simple ERGMs of chapter 4). However, for statistics such as the number of closed triplets, the simulation approach becomes necessary.

requirement of the model to adhere to one of these (Ripley and Snijders 2020, p.10). Second, an important property of the SAOM is that it explicitly reconstitutes a transition between observed network states, through the simulated mini-steps. As a result, the model clearly posits a certain type of data-generation process that drives the evolution of the network; and the parameters reflect biases toward certain network structures within that process. By contrast, ERGM is more theoretically agnostic, in the sense that the parameters can be seen as simply indicating an over- or under-representation of certain tie configurations. While they *can* be interpreted as elements of the data generation process, this is not necessary, and is not "built-in" within the estimation procedure the way it is in SAOM. Finally, dependence assumptions tend to be stronger in SAOM, in the sense that the current network state, as well as past network observations, weight more on other model parameters. This is of course due to its longitudinal form but remains true even when comparing SAOM to a Temporal ERGM (Block, Stadtfeld, et al. 2019).

Altogether then, SAOM offers a clearer framework for how parameters can be interpreted in terms of social processes, though this probably comes at the cost of stronger assumptions. As we shall see below, this micro-level interpretation of parameters is pivotal to the method of contribution scores, meant to assess the impact of specific social processes to socioeconomic homophily specifically. Indeed, contribution scores rely on a counter-factual reasoning (what would happen *if* certain social processes *were* different), which is only meaningful if one assumes that model parameters approximate real-life processes. Moreover, having stronger dependence assumptions means that we reduce the chances of wrongly attributing to an attribute-based process (e.g. same-ethnic selection) the effects of unobserved tie-dependencies (e.g. delayed transitivity over time). Thus, it seems worth taking the risk of more demanding assumptions, especially considering that these remain relatively reasonable in the present case[117].

Still, there are two important limits to the use of SAOM. First, as I just mentioned, stronger assumptions about the data-generation process mean that the model could be wide off the mark if these assumptions are unreasonable (though I do not believe that this is the case). Second, SAOM takes the first observation as granted, and only models the networks from the second observation onward. This can be problematic in the present case, because the social processes that drive the evolution of students' friendships throughout middle school may

117 In-school friendship networks among teenagers are probably the type of networks for which SAOM is most frequently used and, to an extent, for which it has been originally developed. This can increase our confidence that the model assumptions are well-fitted to this particular type of network.

differ from those that led to the formation of this early network: the first months of middle school are a particularly intensive period of friendship formation, as most students did not know one another at the beginning of the year. In order to mitigate these issues, I also estimate cross-sectional ERGMs on the first wave of data collection only. These are meant to hint at what may have happened in the early stages of middle school, as well as to offer some form of robustness check for the SAOM. These ERGMs are presented in Appendix 5E, along with a short discussion of how they compare to the SAOM in section 1.3.6 below. For the most part, results are robust against this alternative specification: although some minor differences exist, they do not fundamentally change the conclusions, and they are small enough that they could be attributed to technical details of how the two classes of models function, thus not necessarily indicating actual differences in the relational processes prior to wave 1[118].

We can now turn toward the specification of the SAOMs that will be used to infer the relational processes that guide friendship formation among students.

1.2. Model Specification

1.2.1. Time Periods for SAOM

As mentioned in chapter 3, the data collection for this thesis was disrupted by the lock-down adopted in France during the COVID pandemic. As a result, wave 5 was not collected, such that there is a one-year gap between waves 4 and 6, instead of six months as is the case for the other periods. Moreover, during wave 6, various sanitary restrictions were still in place: for instance, students had to wear surgical masks, and a curfew forbade them from leaving their homes after 6 P.M. This makes it difficult to model the evolution of friendship networks from wave 1 to 6 as a uniform process: not only were friendships not observed at regular intervals, the situation in wave 6 may imply different relational processes compared to waves 1 to 4. Therefore, I chose to focus on the evolution of friendship networks from wave 1 to 4, which makes for three periods of six months to be modeled by SAOM (wave 1 to 2, 2 to 3 and 3 to 4). There is enough information during these periods – i.e. enough changes in the networks – to meaningfully infer the social processes of interest without risking over-fitting issues.

118 Since the models are different and consider different types of dependencies, there is no reason to expect similar parameters to have exactly the same values in the ERGM and SAOM. Nevertheless, qualitatively at least, the two models should lead to roughly similar conclusions if it is the case that the social processes driving friendship formation are similar during the two periods (and assuming both models capture these processes reasonably well).

However, this limits the possibility of studying changes across time in these social processes themselves, since the models now cover only a year and a half, out of the four years of middle school. For this reason, I specified models on even waves only; that is, waves 2, 4 and 6, which makes for two periods of one year each. This is a way of using the information of wave 6 and investigating potential changes in network formation processes by the end of the students' time in middle-school, while avoiding issues of uneven time periods. Time dummies (see 2.6 below) are used to account for the heterogeneity of effects between period 2-to-4 and period 4-to-6. These "extra" models are not presented in the body of the text, but in Appendix 5D, since they pertain to a rather exceptional period. When interpreting the results from the main models, I will signal relevant differences with these alternative specifications.

1.2.2. Choice of Effects

When choosing the effects to add to the models, three things need to be considered. First, given the focus on socioeconomic homophily, it is important to include the processes most likely to favor friendships among same-background students. Second, following the theoretical discussion of chapter 2, we want, where possible, to use statistical effects that correspond to the different classes of tie-formation processes that were identified. Finally, the model should provide a good fit for the data. Indeed, if we are to treat the parameters as capturing real-life processes, rather than merely accounting for the over- or under-representation of certain network features, then relevant confounders should be accounted for, such that the value of the parameters may be deemed (relatively) unbiased. In that regard, a well-fitted model is less likely to omit essential covariates that are key in explaining friendship formation, although the possibility still exists.

The theoretical typology of friendship formation processes is summarized in Figure 2-4 and Table 2-1, back in chapter 2 (section 1.6). To reiterate, four classes of process are identified: propinquity through imposed foci, propinquity through chosen foci, disposition-driven selection and network endogenous processes. In the next sections, I discuss how these different categories can be operationalized through SAOM effects, starting with network endogenous processes (1.2.3), then considering imposed foci (1.2.4) and finally expressed dispositions (1.2.5).

A word of caution as per the terminology that I use. Any term added to the model has a macro and a micro side, respectively the target statistic and the parameter. The estimated value of a parameter is also called a coefficient. Parameters pertain to micro-level processes from the perspective of the model, but they are still defined in regard to a target statistic. For example, the “SAS homophily parameter” is the parameter that guarantees that the model reaches the right target statistic of SAS (Socio-Academic Scores) homophily. Remember that homophily is a descriptive concept that pertains to a macro-level structure: thus, micro-level parameters are named after macro-level outcomes - that of the statistic they are associated to. By contrast, both “selection” and “propinquity” refer to real-world social processes and should not be confused with the behavior of nodes in the model’s simulations. It is only through interpretation that we can consider that the SAS homophily parameter represents or captures socioeconomic selection, or that the classroom homophily parameter captures classroom propinquity (typically, if we assume that relevant confounders are accounted for by other model effects).

1.2.3. Network Endogenous Processes

Network endogenous processes can be modeled by structural effects that represent dependencies among ties. These are also called dyad-dependent terms, because they imply that the state of a (directed) dyad depends on the state of other dyads in the network. For the present application, three types of structural effects are considered: mutuality, triadic closure, and centrality.

1. The **mutuality** effect captures the prevalence of reciprocal ties. The parameter considers the log-odd change for the tie $i \rightarrow j$ to be observed when $j \rightarrow i$ exists.

2. Triadic closure effects capture the prevalence of certain triangle configurations. In directed networks, there are several different types of triangles, based on the direction of ties between the three nodes. Two are considered here (see Figure 5-1). First, **transitivity**: the parameter considers the log-odd change for $i \rightarrow j$ when both $i \rightarrow k$ and $k \rightarrow j$ exist. Conceptually, this relates to the idea that “the friends of my friends are my friends”: i wants to befriend j because his friend k is friends with j. Second, a **3-cycle** effect: the parameter considers the log-odd change for $i \rightarrow j$ when $i \leftarrow k \leftarrow j$ exists. This is often assumed to represent the aversion toward hierarchy (i, k and j all share equal status because each wants to befriend one of the others and is also befriended by exactly one other), but this remains a subject of debate in network theory (Block 2015). It is included in certain schools only, where it strongly improved the model’s fit (see Table 5-2).

All triadic effects are based on the definition of geometrically-weighted-edgewise shared partner (gwesp) effects (Hunter 2007; Ripley and Snijders 2020). What this means is that, in cases where a dyad $i \rightarrow j$ is nested within several triangles $i \rightarrow k1 \rightarrow j$, $i \rightarrow k2 \rightarrow j$, etc., each additional shared partner is attributed a decreasing weight (cf. Figure 5-1). In substantial terms, this means that there is a greater difference between having zero and one friend in common, than between having seven and eight. A decay parameter controls the marginal decrease in shared partners' importance, which takes a fixed value (chosen per school and model so as to improve the model's fit).

Finally, an **interaction effect between mutuality and transitivity** is included, following Block's (2015) recommendation. This is relatively close to a 3-cycle effect in terms of what it means for the network, but the typical interpretation differs: it is meant to represent the decreasing marginal relevance of shared foci, which means that we would generally expect the interaction to be negative: if *i* and *j* already met, as shown by the fact that $j \rightarrow i$ exists, then having some friends in common becomes less impacting, relative to two nodes who would have no prior connection and, thus, who might meet for the first time through these shared friends. In most models, this effect completely replaces the 3-cycle effect, but in some schools, both are included, as it improves the fit (they still capture slightly different configurations, cf. Figure 5-1).

3 **Centrality** terms are meant to account for the concentration of ties around high-degree nodes. Substantially, the idea is that it is easier for well-connected individuals to receive and/or emit further ties. For incoming ties, this can capture the social process of *popularity* (people with many friends are perceived as high-status, thus more peers want to befriend them); for outgoing ties, it can capture *activity* (having more friends provides more opportunities to make even more friends). Additionally, these terms can account, to an extent, for unobserved properties of individuals that impact the degree distribution (for example, the fact that some students are intrinsically more likable than others, or the fact that some of them have a looser definition of friendship, thus emitting more nominations altogether).

The terms of **indegree popularity** and **outdegree activity** capture the odds of the focal node sending a tie as a function of, respectively, the indegree of the receiver node and this focal node's own outdegree. One can choose between a linear function of the in/outdegree distribution (each additional tie that is emitted/received weighs the same) or a square root version (see Ripley and Snijders 2020, pp. 44 and 128). The decision to use either of those depends on their contribution to the model's fit, and I will use different ones for different schools. Moreover, it is generally advised to add a third term that models the interaction of in- and outdegrees; in this case, an **outdegree popularity** effect (odds of receiving a tie as a function of the receiving node's outdegree).

The different structural effects included in the model specifications are summarized in Table 5-2. Can we assume that these structural effects will appropriately reflect realistic network endogenous processes? The substantive meaning to attribute to **outdegree activity**, **outdegree popularity** and **3-cycle** effects is rather unclear. These terms control for key structural features of the networks that are related to many other effects, thus improving the model's fit as well as the interpretability of other parameters. However, there is no consensus on how their own parameters should be interpreted in terms of real-life social processes. In other words, these are primarily control terms.

On the other hand, the parameters of **mutuality**, **transitivity** and **indegree popularity** are meant to directly mirror the social processes of reciprocity (friendships are highly reciprocal relationships), transitive closure ("the friends of my friends are my friends") and popularity (high-status individuals attract more friends). These are important network endogenous processes, whose potential contribution to homophily was underlined in chapter 2. Transitivity in particular is expected to aggravate homophilic tendencies (cf. chapter 2 section 2.3); by contrast, we lack clear theoretical predictions for either reciprocity or popularity.

Nevertheless, these structural terms might also capture other processes beyond network endogenous ones. Mutuality and transitivity are not independent from homophily patterns: if individuals befriend similar peers, then there will be more reciprocated ties and closed triangles than expected by chance. Even if we control for various individual characteristics, these terms will capture, to an extent, homophily on unobserved characteristics, as well as the sharing of unobserved foci. As for centrality effects, they may similarly capture popularity processes that are not endogenous, but associated to unobserved characteristics (e.g. students with older siblings may be more popular). This is not necessarily a bad thing: it means that structural terms offer a certain amount of protection against omitted covariates, thus facilitating the interpretation of the parameters that pertain to observed ones (Ripley and Snijders 2020). However, this comes at the cost of a more complex interpretation of structural parameters themselves.

The size of the bias due to unobserved covariates will of course vary by networks and models. Still, Foster et al. (2011) found, based on stylized simulations, that the sensitivity of transitivity (as a network outcome) to homophilic selection (as a low-order process) was relatively low – lower, at least, than the opposite (the sensitivity of homophily to transitive closure). Moreover, the models presented here include several important covariates (gender, social background, academic results, place of residence, etc.) as well as the longitudinal form of SAOM which should indirectly capture part of the effect of unobserved covariates as well (the 1st network observation, which is taken for granted, contains a lot of information in that regard). Therefore, it seems reasonable to assume that the estimated coefficients for structural effects are not exceedingly inflated due to the omission of unobserved covariates, and that they approximate the corresponding social processes relatively well, especially in the case of SAOM. This applies to reciprocity, transitivity and indegree centrality, which are well-understood effects with clear theoretical foundations in terms of students' behaviors[119]. I will refrain, however, from making a similar interpretation for other structural effects.

119 Crucially, we also have qualitative evidence for the corresponding social processes. Ethnographic research has established that the notion of reciprocity is pivotal to the definition of friendship in our societies (Bidart 1997); that friendships among school students are strongly constrained by group-belonging and peer pressure (Bidart 2010; van Zanten 2000), transitivity being effectively a formal operationalization of this idea; and that the number of friends is an important resource in terms of peer status and prestige (Eder 1985; Merten 1997).

Figure 5-1a: Network Endogenous Effects: Mutuality, Transitivity, 3-Cycles and Mutuality*Transitivity Interaction

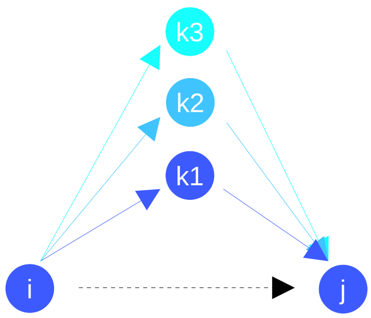
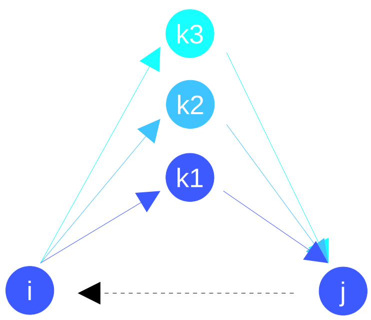
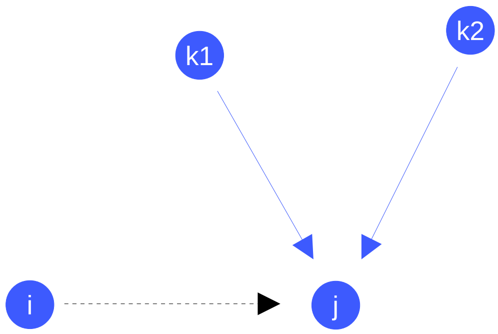
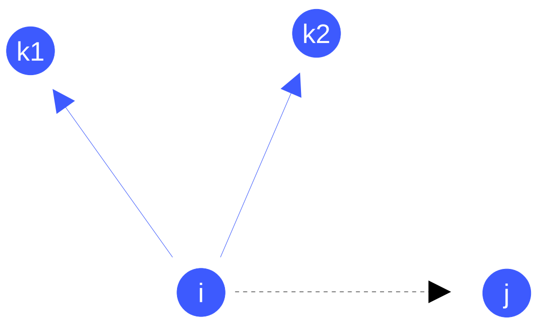
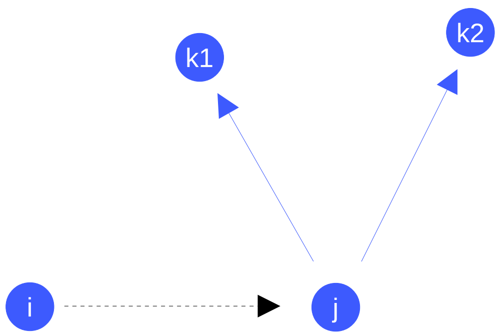
Model Term	**Tie Configuration**	**Interpretation of the Parameter**
Mutuality		Log-odd change for i -> j given j -> i
Gwesp OTP (Outgoing Transitive Partner, i.e. transitivity)		Log-odd change for i -> j given a triangle i -> k and k -> j. Decreasing weights (with decay α) applied to each additional k.
Gwesp ITP (Incoming Transitive Partner, i.e. 3-cycle)		Log-odd change for j -> i given a triangle i -> k and k -> j. Decreasing weights (with decay α) applied to each additional k.

Figure 5-1b: Network Endogenous Effects: Indegree Popularity, Outdegree Activity and Outdegree Popularity

<u>**Model Term**</u>	<u>**Tie Configuration**</u>	<u>**Interpretation of the Parameter**</u>
Indegree Popularity		Log-odd change for i→j for each additional k→j (number of ks may be squared)
Outdegree Activity		Log-odd change for i→j for each additional i→k (number of ks may be squared)
Outdegree Popularity		Log-odd change for i→j for each additional j→k (number of ks may be squared)

1.2.4. Imposed Foci

Preliminary Remark: Model Specifications for Attribute-Based Effects

While network endogenous processes are represented in the models by dyad-dependent terms, the processes pertaining to imposed foci and expressed dispositions are represented by dyad-independent terms. These terms treat tie-level observations as being independent from one another; meaning that they model the probability of a tie to be observed as a function of properties of that tie, regardless of the state of other ties in the network. The explanatory variable is therefore a dyadic covariate, i.e. a property of the (directed) dyad. In most cases, these dyadic covariates are built from individual attributes, i.e. properties of the nodes. For example, gender homophily is a property of a dyad that is defined by two nodes having the same gender.

Given an individual attribute, the three most commonly used dyadic covariates are *homophily, emission* and *reception*. Homophily is defined by the two nodes of the dyad having the same value of the attribute; emission is defined by the node sending the nomination having a certain value of the attribute; and reception is defined by the node receiving the nomination having a certain value of the attribute. Reception and emission terms therefore capture the effect of a nodes' attribute on tie formation regardless of the property of the other node of the dyad. For example, it could be that girls send more friendship nominations overall, both to boys and other girls; or that girls are more popular on average, thus receiving more nominations from both boys and other girls. These effects, if not accounted for, can cause bias in homophily estimates. If girls send and receive more ties, the probability of a girl-to-girl tie will be higher than for heterophilic ties; yet this would not truly indicate homophily, because girls are in fact nominating and being nominated by boys and girls equally. Therefore, network analysts generally include emission and reception terms as controls together with any homophily parameter.

I have talked here of categorical attributes, extending to continuous attributes is straightforward. Emission and reception terms do not consider a specific value for the attribute, but the impact of a one-point increase in its value (e.g. how much more likely one is to receive a tie for each additional point of academic performance). As for homophily terms, they look at the absolute difference between the values of the two nodes: the higher this difference, the more distant the nodes are from the point of view of the considered attribute. In SIENA, these absolute differences are re-scaled on the interval (0;1) and subtracted from 1, in order to indicate similarity rather than distance; the corresponding statistic is therefore called a *similarity statistic* (see the *Rsiena* manual).

Imposed Foci Considered in the Models

There are 5 types of imposed foci that I have information about: classrooms, special tracks, places of residence, former primary schools and, to an extent, parental networks.

Within schools, students are distributed across classrooms, whose sizes vary from about 25 to 35 individuals. These classrooms are fixed for the whole academic year, but they can be reshuffled from one year to the next (this depends a bit on the school and year but, in general, the reshuffling is only partial: there is a “core” of students for each classroom that carries over from one year to the next). Most courses take place with the entire group at once, and students’ timetables are also classroom-based (e.g. what time does school start and end each day, what time is lunch, etc.). For these reasons, the distribution across classrooms is the major formal constraint that school organization imposes on students’ sociability, and we should expect same-classroom ties to be much more likely than cross-classroom ones. Two homophily effects are added, one for the classroom at time *t* and the other for the classroom at time *t-1*. This is because SAOM models the transition between two waves, which means that students can change classroom during this transition (when the two consecutive waves occurred during different school years). In three out of the four schools, no emission nor reception effect is included: it would excessively inflate the number of model parameters, making the models harder to estimate as well as leading to potential over-fitting issues. Alternative specifications (not shown) were tested to make sure that this would not bias the value of other parameters of interest. In Paris 1, however, classroom main effects turned out to affect the value of other important parameters while not raising too many estimation issues, hence their inclusion in the final specification.

In Paris 1 and Savoie 1, a small group of students (about 15) attend a special track. In Paris 1, it is a rather “elitist” track (a French-English bilingual section), which is integrated within a regular classroom (students attend most of their courses together, except for a few special bilingual courses). In Savoie 1, it is an entirely separate classroom for students with severe learning difficulties. The track in Paris 1 mostly includes students with upper-class backgrounds, whereas the one in Savoie 1 mostly includes students with working-class backgrounds. In both cases, three dedicated effects are added to the model: two differential homophily effects (homophily of the special track, and homophily of the majority group), plus an emission effect capturing the asymmetry in nominations emitted between the special and normal tracks. Though I will use uniform homophily for most other attributes (i.e. not distinguishing between, for example, boy-to-boy and girl-to-girl homophilies), there is a clear asymmetry between students' from the special track and others, in terms of group size but also of indigenous labeling (they are seen as being the "special" or "abnormal" case). It is therefore interesting to consider two separate homophily effects. However, this means that only one out of emission and reception terms can be added together with these two effects[120].

I geocoded students' declared place of residence, then computed pairwise walking distances between their homes using Open Route Service (Oles 2021), a cartographic tool based on Open Street Map data. There are good reasons to expect students that live closer from one another to have higher chances of being friends: not only is it easier to see friends that live close-by (especially for younger students that need their parents' authorization to go out or use public transportation), the place of residence should also approximate local community networks. To an extent, it may also capture the effect of some chosen foci that are constrained by the place of residence, such as sport or art clubs, gardens, sport fields, swimming pools or libraries. One potential problem in terms of causal identification is that students living closer should be socially more similar on average, given the socioeconomic segregation of the housing market, particularly in the Parisian agglomeration (Oberti and Préteceille 2011). Nevertheless, since we are controlling for students' socioeconomic and ethnic backgrounds, this bias should be greatly reduced, such that estimates should come close to the net effect of residential propinquity. Two effects are included in the models: the first considers the linear walking distance between students' homes (in hours), and the second the same distance squared, to account for the decreasing marginal relevance of residential distance (when students live far away from one another, they need to use public transportation or car, which effectively compresses distances compared to walking).

Before coming into middle school, students attended different primary schools. Peers from the same primary are more likely to be friends, since they have known each other for a longer time[121]. Primary schools are generally smaller than middle schools and located closer

120 The special track attribute is binary: 1 for being into the track, 0 for being outside of it. This means any tie can be in either of four cases: 1-to-1, 1-to-0, 0-to-1 and 0-to-0. Since there are only four possible combinations, a maximum of three terms can be added to the model. Interpretation-wise, this implies that the emission or reception term included reflects the difference in tie odds between 1-to-0 and 0-to-1 types of ties. In this case (emission for the special track), a positive parameter would mean that 1-to-0 ties (special track to regular track) are more likely than 0-to-1 ones, net of the two groups' respective homophilies.

121 Note that since SAOM is a longitudinal model, the first wave of observation is taken for granted. This means that most, if not all friendships that were formed in the past during primary school years should already be observed in wave 1 (*former* sociability contexts should have little impact on the creation of *future* friendships). Therefore, the primary school effect can approximately be seen as an endowment effect, i.e. modeling

to students' places of residence. There is a small socioeconomic imbalance across former primary schools for Paris 1 and Savoie 1, but not Paris 2 and Savoie 2, which is probably linked to the fact that private schools are not bound by geographic districts the same way public schools are (cf. Table 7-4 in chapter 7). Since the distances between students' homes are controlled for, the parameter should mostly reflect the net effect of having been in the same school, rather than living closer to one another. However, it may also capture some threshold phenomena linked to various details of the residential configuration, that are not accounted for by continuous residential distances. A single homophily effect is added; similar to the case of classrooms discussed above, emission and reception effects would add too many parameters to the models, so they are not included.

Finally, in waves 1 and 3, students were asked to nominate the peers whose parents knew their parents. Therefore, it is possible to create a parental network, made of the ties among students' parents (as reported by the students). Ideally, we would like to test the impact of parents' relations on the odds of friendship among students. However, an issue is that parental ties are endogenous to students' friendships, as parents can meet through their children. For this reason, parental ties are not part of the main model specification, but they are included in an extended, exploratory specification. Even though we cannot be sure that parental ties are not caused by students' friendships, it is still interesting to consider how strongly they are associated to them, as well as to see whether they could have a mediating impact on socioeconomic homophily (though the model should over-estimate this mediation, due to the endogeneity issue). The parental ties from wave 1 are included as a time-constant dyadic covariate: the corresponding parameter gives the log-odd change in the probability of observing $i \rightarrow j$ when i declared that her parents knew j's parents. Note that SAOM starts to model the evolution of friendship networks after wave 1, which means that these parental ties are strictly anterior to the modeled process of friendship evolution. This reduces concern for reversed causality, though it does not entirely solve the issue (the parental ties of wave 1 could capture unobserved heterogeneity in the strength of wave 1's friendships).

Can we reasonably associate the parameters of these various effects to real-life relational processes? For classrooms, primary schools, and residential distance, I expect the parameters to capture propinquity processes, since we control for the effect of several disposition-related attributes (see below). While in theory, there could also be some selection

the greater resilience over time of friendships formed at the time (there may also be new friendships formed at a later age among students that were in primary school together while not being friends at the time, but this is marginal).

processes associated to these foci (e.g. forms of pride or "nationalistic" feelings toward one's classroom or neighborhood), I found no hint of those during the qualitative fieldwork: students did not mention those during the interviews, nor did teachers or headmasters during informal discussions. The literature suggests that one's neighborhood in particular, may be a strong element of identification in segregated working-class environments (Lepoutre 1997), but this seemed not to apply (or not much) for the socioeconomically diverse schools studied here.

The case of special tracks is more complex: in both Paris 1 and Savoie 1, there is a clear label, if not stigma, associated to these tracks, and the students that attend them are well-known among their peers. On several occasions, I witnessed some students explicitly antagonizing peers based on their belonging to a certain track; for instance, saying that internationals were condescending (Paris 1) or that SEGPAs were stupid (Savoie 1). Therefore, the homophily parameters should capture a mixture of selection and propinquity processes.

Finally, as I have already mentioned, parental networks raise important identification issues, so one should refrain from interpreting the corresponding parameters in terms of propinquity processes.

1.2.5. Expressed Selective Dispositions

Five individual attributes are considered that are not related to foci, but instead to intrinsic traits or properties of students. Therefore, the corresponding parameters are generally meant to reflect selection rather than propinquity processes – though, once again, this depends on the presence of appropriate controls within the specification. The five attributes are: socioeconomic background, gender, ethnic background, academic achievement, and cultural tastes. For each of those, the effects of homophily, reception and emission are systematically included.

The construction of the variables of socioeconomic and ethnic backgrounds has been presented in chapter 3 already (sections 1.2.3 and 1.2.5). As a reminder, socioeconomic background is operationalized through Socio-Academic Scores (SAS), a continuous measure based on the academic performance predicted by the occupation of students' parents. In other words, this is a way of ordering occupational groups based on their expected school performance, which we can see as capturing the average academic or cultural capital of occupational groups. Ethnic background is built from students' first and last names, as well as the foreign languages that they declared to speak or understand, which are used to assess migratory origins. The original variable takes 13 modalities corresponding to different world-regions, but here I use an aggregated version of the variable with 5 modalities, for parsimony (natives, Middle East and North Africa, Sub-Saharan Africa, Europe, and others). The homophily effect only considers the 3 main groups (natives, MENA and Sub-Saharan Africa), as other origins have structurally few opportunities for homophilic bonding. Robustness checks were conducted with the 13-post variables to make sure that results were not substantially affected.

There is no theoretical reason to believe that gender would introduce biases regarding socioeconomic homophily, as the gender distribution should be random with regard to social class. Still, this is one of the most important determinants of students' relations, so it is essential to include in any model.

In Paris 1 and Savoie 1, grades were obtained from the school administration. For each network observation, the trimester average was obtained at the closest time possible *before* the wave was used. In Paris 2, however, grades are only available for the cohort's first year of middle school, because the school refuses to communicate the grades for subsequent years. Therefore, only the grades of the first wave are used; they are treated as a constant covariate throughout all waves. Finally, in Savoie 2, students do not receive numerical grades, but ordinal "competences" for each school subject, such that there is not a general mean computed over all courses. The "grade" variable I use is therefore a construct of my own meant to capture students' overall academic level: I used a factor analysis on all competences, then kept the coordinates on the first axis, indicative of overall performance (cf. chapter 3, section 1.2.4).

For these four attributes, an interpretation of the parameters in terms of selection processes may be considered reasonable, given the extensive controls included in the models for structural network processes and imposed foci. In that regard, I would argue that most external constraints that shape students' sociability are at least partly accounted for – classrooms, primary schools, places of residence and, in the extended models, parental networks, thus covering a wide range of contextual factors. Of course, these controls are not perfect and, to the extent that the measures used contain some noise, part of the effect of the underlying factors may still be captured by other parameters. Still, the estimated parameters should come much closer from effective selection processes than the "naive", descriptive estimates presented in chapter 4.

One important limitation concerns the lack of information about chosen foci. Students' out-of-school activities are not considered in the models[122]; and, even if they were, it is practically impossible to track all the "mini foci" that structure individuals' social lives (going to the movies on Wednesday, to the park on Thursday, etc.) and that are endogenous to friendship networks in the first place. Nevertheless, this is not as problematic as it may seem: though conceptually distinct from selection processes, chosen foci still pertain to students' expressed dispositions (cf. figure 2-4), meaning that they are related to internal dispositions rather than external constraints (though these dispositions are always "filtered" throughout contextual moderators). Therefore, assuming that the estimated parameters capture a mixture of selection and of propinquity through chosen foci, they would still be a good measure of the strength of dispositional processes in friendship formation.

Finally, cultural tastes are measured using two different types of covariates. *Cultural agreement* is a measurement of how many opinions students share about various musicians and video-makers. C*ultural status* is a continuous score that distinguishes between lowbrow and highbrow cultural practices, based on students' declarations of their favorite music, their reading habits, and their artistic activities. The first is represented in the model by two covariates that give, for each dyad, the squared number of common opinions that students have on musicians and video-makers separately. The second is an individual attribute, hence the inclusion of a homophily, a reception and an emission term. These measures are presented in more details in chapter 3, section 1.2.6. Both are clearly imperfect: cultural agreement was defined from a restricted set of opinions, and cultural status relied on my subjective assessment of what is a highbrow or lowbrow practice. However, the use of two measures is also meant for each one to (partly) cover the defaults of the other, such that they jointly capture a good portion of students' cultural similarity. Moreover, theory-wise, similarity in cultural tastes may be an important mediator of socioeconomic homophily. Therefore, it seems important to try and consider its impact on socioeconomic homophily, even with imperfect measures.

Since there should be some peer-to-peer socialization in terms of cultural tastes, these measures are partly endogenous to students' sociability. For this reason, and similar to what was done for parental networks, cultural variables are included in an extended specification,

122 I did collect some information about sport and artistic activities outside of school, but in practice they are very hard to integrate into the models in a satisfactory way: they are strongly endogenous – students choose activities to be with their friends – and co-attendance of an activity is usually very sparse (most students have at most 1 or 2 schoolmates in their out-of-school clubs).

separate from the main one. These extended models can thus be considered as exploratory treatments.

Table 5-1 below summarizes the different models that will be estimated. “Model 1”, “Model 2” and “Model 3” are the SAOM specifications that are presented in the body of the text, which are based on data from waves 1 to 4 and whose effects have been discussed across this section. On the other hand, SAOMs for waves 2 to 6 and ERGMs for wave 1 are presented in Appendix 5D, and their specifications can slightly differ from those of the main models. I will briefly mention them in section 1.3.6 for robustness checks. Moreover, all the effects discussed in this section are summarized in Table 5-2.

Table 5-1: Summary of the Estimated Models

Presented in...	Model Name	Period	Effects
Body of the text	Model 1	waves 1 to 4	main specification
	Model 2	waves 1 to 4	main specification + cultural variables
	Model 3	waves 1 to 4	main specification + parental networks
Appendices 5D and 5E	SAOM w246	waves 2, 4 and 6	main specification
	ERGM 1	wave 1	main specification
	ERGM 2	wave 1	main specification + cultural variables
	ERGM 3	wave 1	main specification + parental networks

Note: in each case, 4 different models are estimated, one per school, with minor differences in the corresponding specification.

Table 5-2: Summary of Effects Included in the SAOM Specifications

Class of Effects	Individual Attribute	Covariate	SAOM term	Assumed Social Process
Baseline Parameters	-	Density	outdegree	-
	-	Rate Function	rate	-
Network Processes	-	Mutuality	recip	Reciprocity
	-	Transitivity	gwesp FF	Transitive closure
	-	Mutuality * Transitivity	recip * gwesp FF	Decreasing returns to relational proximity
	-	(3-cycle)	(gwesp BB)	-
	-	(Two-path)	(gwdsp FF)	-
	-	Indegree Centrality	inPop (sqrt)	Popularity
	-	Outdegree Centrality	outAct (sqrt)	-
	-	Outdegree Popularity	outPop (sqrt)	-
Imposed Foci	Classroom	Homophily (+ emission + reception)	sameX (+ egoX + altX)	Propinquity
	Track	(Differential Homophilies + Emission)	(sameX + egoX)	Propinquity + Homophilic Selection
	Place of Residence	Distance + Squared Distance	X	Propinquity
	Primary School	Homophily	sameX	Propinquity
	Parental Network	Main effect	X	-
Expressed Dispositions	Socio-Academic Scores (socioeconomic background)	Homophily + emission + reception	simX + egoX + altX	Homophilic Selection
	Ethnic Background	Homophily + emission + reception	sameX + egoX + altX	Homophilic Selection
	Gender	Homophily + emission + reception	sameX + egoX + altX	Homophilic Selection
	Academic Achievement	Homophily + emission + reception	simX + egoX + altX	Homophilic Selection
	Cultural Agreement	Main Effects: Sqrt of the Nb of Shared Opinions on Musicians + Sqrt of the Nb of Shared Opinions on Video Makers	X	-
	Socio-Cultural Scores	Homophily + emission + reception	simX + egoX + altX	-

Note: the name of terms in the column "SAOM term" is taken from the R packages *Rsiena* (for SAOM). Terms between parentheses are included in the specification for certain schools, but not all of them (see section 1.2.6 below).
The effects within shaded rows are not included in the main specification, but in a separate, extended specification.

1.2.6. Further Remarks on Model Specification

Differences across Schools

The specifications described in Table 5-2 are applied to the four schools, but with minor differences meant to account for particularities in their network structures. First, the decay parameters of the curved effects (‘gwesp’, ‘gwodegree’ etc.) differ across models and schools. Second, the centrality terms are sometimes used in their linear version, and sometimes in their square root version. Third, 3-cycle effects are used in some schools, but not all. Finally, the models of Paris 2 have both a two-path effect (‘gwdsp’) and an outdegree popularity (‘outPop’) effect, whereas the other schools only have the outdegree popularity effect. These two terms are somehow equivalent in what they mean for the network but including both greatly improved the model’s fit in Paris 2 (see Snijders et al. 2006 for a presentation of two-path effects). In general, all the differences mentioned here were based on trial-and-error to maximize the goodness-of-fit of the models.

Furthermore, special tracks are only meaningful in Paris 1 and Savoie 1. In Savoie 2, there was a special track in wave 1 (pre-vocational track) but it disappeared at the beginning of wave 2, due to a policy change from the school administration, as a result of which most students left the school. As for Paris 2, there never was any special track there. Finally, emission and reception effects for classrooms are only included in Paris 1, where their omission would bias other estimates (notably that pertaining to the international section).

The complete model specifications are available in Appendix 5B, together with goodness-of-fit tests for all schools and models.

Missing Values

SAOM can handle missing values, both for attributes and network ties. Therefore, missing value imputation was left to the internal data treatment of SIENA (the software used for the estimation of SAOM). SIENA treats missing values as occurring at random; as a rule of thumb, it is considered that the estimates remain reliable up to a maximum of 20% of missing values (Ripley and Snijders 2020, p.13), which is far more than the actual number in my sample. I use the default method implemented by the *Rsiena* package (*ibid.*, p.35). Importantly, this means that students who missed the questionnaire in one wave but answered in prior or later waves, are still included in the model: their missing outgoing ties at time *t* are inferred from their outgoing ties at time *t-1* (if available). Nevertheless, students that never answered any questionnaire are excluded from the treatments, because there is not enough information about them to meaningfully impute anything.

Note that missing values are not the same as composition change, i.e. the fact that students leave and arrive in the schools at different times across the three years. These are treated by using structural zeros at the times where students have not yet arrived in the cohort, or have already left it (Ripley and Snijders 2020, p.32).

Details about the number of missing values for the different covariates are available in Appendix 5A.

Rate Function and Time Heterogeneity

Finally, additional parameters are included in the models to account for time. First is the *rate function*, which controls the number of mini-steps simulated between each wave, effectively representing the speed at which the network evolves. It takes one parameter per period. I use a constant rate function, meaning that each node has the same probability of being selected at each mini-step to perform a change in its outgoing ties.

Second, it can be necessary to include *time dummies* in the models to account for the heterogeneity of certain effects across periods. For example, there is a decrease in network density in wave 2 for certain schools, so it is necessary to add an interaction between time and the outdegree parameter to account for this properly. Details about the dummies included for each school are in Appendix 5B. I will also come back to the interpretation of these terms in chapter 7.

1.3. Results for the SAOM Estimation

The estimated models are presented in Table 5-3. Model 1 only includes the terms from the main specification, whereas models 2 and 3 include the effects pertaining, respectively, to cultural tastes and parental ties. Only the effects of interest for the analysis are presented here; complete outputs, with all control terms, are in Appendix 5B.

A word on the interpretation of the value of parameters. These are log-odd coefficients, that give the change in the odds of a node emitting a tie as a function of the covariate value of that tie. This is similar to the ERGM coefficients discussed in chapter 4, but there are two important differences to keep in mind. First, since SAOM models network evolution, the coefficients consider the odds of a tie as being either created (switching from 0 to 1) or maintained (from 1 to 1). It would be possible to estimate separate effects for the creation and maintenance of ties, but this is not done here. Second, since SAOM simulates a series of mini-steps in-between network observations, the coefficients are applied at each simulated mini-step. Indeed, when a node is selected at a given mini-step during the simulation process, it is allowed to perform one action, either send a tie, delete one, or do nothing; and the probabilities of all possible actions are determined by the value of the parameters (using a multinomial logit model). Importantly, all the nodes will be selected several times between two network observations; this depends on the estimated parameters of the rate function, but the model will generally simulate an average of at least 20 mini-steps per node. What this means is that a given coefficient does not tell us how likely a tie is to be maintained or created from one wave of data to the other: the impact of a given coefficient will be repeated over several mini-steps, so the probability of an action at a given mini-step is not the same as the probability that this action will be performed at some point over the entire sequence of mini-steps. For example, for the gender homophily parameter, a coefficient of 0.5 would mean that, *at a given simulated mini-step*, assuming that node *i* has the choice between sending a tie to a same-gender node *j1* and a different-gender node *j2*, and assuming that *j1* and *j2* have perfectly equivalent network positions from the perspective of *i* apart from their gender, then *i* would be exp(0.5) ~= 1.65 times more likely to create/maintain a tie toward *j1* than toward *j2*. However, this does not tell us how likely a tie is to be created/maintained between *i* and *j1*/*j2* between two waves in the observed networks.

Due to this, it is not always easy to recognize whether a given coefficient should be considered as large or small, or whether its impact on the network is important. One simple solution is to compare different coefficients within the same model; e.g., by looking at how strong socioeconomic similarity is relative to gender matching. Moreover, familiarity with the literature and the effect sizes typically found in other studies can help assess the strength of the modeled effects. Therefore, in the following sections, I will try to indicate when an effect should be considered as strong, moderate or weak, on top of giving its exact coefficient value.

I first detail results from the main specification (Model 1), with network endogenous effects (1.3.1), imposed foci (1.3.2) and expressed dispositions (1.3.3) examined in that order. Then I comment on the terms specific to models 2 and 3, and how their inclusion seems to affect other coefficients (1.3.4 and 1.3.5). Finally, I briefly indicate how results differ in the supplementary models presented in the Appendices (1.3.6).

Table 5-3a: Estimated Parameters for the SAOM Models (waves 1 to 4) – Paris 1 and 2

	Paris 1			Paris 2		
	Model 1	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
mutual	2.36*** (0.19)	2.34*** (0.19)	2.32*** (0.18)	3.53*** (0.14)	3.50*** (0.12)	3.44*** (0.13)
GWESP I -> K -> J (decay 110 / 69)	1.14*** (0.06)	1.12*** (0.06)	1.12*** (0.06)	1.56*** (0.06)	1.54*** (0.05)	1.55*** (0.05)
mutual * GWESP I -> K -> J (decay 110 / 69)	-0.49*** (0.11)	-0.50*** (0.11)	-0.47*** (0.11)	-1.28*** (0.08)	-1.27*** (0.08)	-1.25*** (0.08)
indegree popularity (linear/sq)	-0.03*** (0.01)	-0.03*** (0.01)	-0.03*** (0.01)	0.21*** (0.02)	0.20*** (0.02)	0.21*** (0.02)
outdegree popularity (linear/sq)	-0.06*** (0.01)	-0.06*** (0.01)	-0.06*** (0.01)	-0.66*** (0.04)	-0.66*** (0.03)	-0.64*** (0.04)
outdegree activity (linear/linear)	0.003 (0.003)	0.003 (0.003)	0 (0.003)	0.01*** (0.0006)	0.01*** (0.001)	0.01*** (0.001)
walking distance (hours)	-0.27 (0.29)	-0.10 (0.27)	-0.16 (0.27)	-0.21*** (0.06)	-0.23*** (0.06)	-0.12** (0.06)
walking distance – squared	0.28 (0.22)	0.15 (0.21)	0.2 (0.21)	0.07*** (0.03)	0.09*** (0.03)	0.04 (0.03)
same primary school	0.03 (0.06)	0.03 (0.06)	0 (0.06)	-0.02 (0.03)	-0.03 (0.03)	-0.11*** (0.03)
same classroom (t)	0.17*** (0.07)	0.19*** (0.06)	0.18*** (0.06)	0.08*** (0.02)	0.09*** (0.02)	0.10*** (0.02)
same classroom (t-1)	0.49*** (0.07)	0.49*** (0.06)	0.48*** (0.06)	0.90*** (0.03)	0.90*** (0.03)	0.91*** (0.03)
same track – international section	0.03 (0.17)	0.00 (0.18)	-0.01 (0.16)	-	-	-
same track – regular section	0.52*** (0.14)	0.45*** (0.13)	0.5*** (0.13)	-	-	-
emission – international section	-0.12 (0.19)	-0.23 (0.19)	-0.1 (0.19)	-	-	-
same sex	0.47*** (0.05)	0.45*** (0.05)	0.46*** (0.05)	0.32*** (0.03)	0.31*** (0.02)	0.31*** (0.02)
same ethnic group	0.13** (0.06)	0.15*** (0.06)	0.11* (0.06)	0.20*** (0.03)	0.19*** (0.02)	0.18*** (0.03)
grades similarity	0.57*** (0.17)	0.42*** (0.16)	0.54*** (0.16)	0.28*** (0.07)	0.25*** (0.07)	0.23*** (0.07)
factbrev reception	0.10*** (0.04)	0.11*** (0.04)	0.09*** (0.03)	0.01 (0.01)	0.002 (0.01)	0.01 (0.01)
factbrev emission	0 (0.04)	-0.01 (0.04)	0 (0.04)	-0.04*** (0.01)	-0.06*** (0.01)	-0.03*** (0.01)
factbrev similarity	0.36*** (0.14)	0.3** (0.14)	0.33** (0.14)	0.11* (0.06)	0.09 (0.06)	0.09 (0.06)
time dummy factbrev similarity * period3	-	-	-	0.40*** (0.13)	0.38*** (0.12)	0.41*** (0.12)
cultural agreement – music	-	0.12*** (0.03)	-	-	0.08*** (0.01)	-
cultural agreement – youtube	-	0.06** (0.03)	-	-	0.03** (0.016)	-
cultural status similarity	-	0.24 (0.16)	-	-	0.13*** (0.05)	-
parental ties (wave 1)	-	-	0.47*** (0.11)	-	-	0.97*** (0.05)

*p<0.1; **p<0.05; ***p<0.01

See Appendix 5B for the full models.

Table 5-3b: Estimated Parameters for the SAOM Models (waves 1 to 4) – Savoie 1 and 2

	Savoie 1			Savoie 2		
Effect Name	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
mutual	2.39*** (0.08)	2.39*** (0.09)	2.33*** (0.08)	2.72*** (0.08)	2.72*** (0.09)	2.67*** (0.08)
GWESP I -> K -> J (decay 125 / 85)	1.02*** (0.04)	1.02*** (0.04)	1*** (0.04)	1.36*** (0.04)	1.36*** (0.04)	1.33*** (0.04)
mutual * GWESP I -> K -> J (decay 125 / 85)	-0.59*** (0.05)	-0.58*** (0.05)	-0.52*** (0.07)	-0.6*** (0.07)	-0.8*** (0.08)	-0.59*** (0.07)
indegree popularity (linear/sq)	-0.01*** (0.004)	-0.01** (0.004)	-0.01*** (0.004)	-0.006 (0.004)	-0.008* (0.004)	0 (0.004)
outdegree popularity (linear/sq)	-0.05*** (0.005)	-0.05*** (0.005)	-0.05*** (0.005)	-0.05*** (0.004)	-0.05*** (0.004)	-0.05*** (0.003)
outdegree activity (linear/linear)	0.002** (0.001)	0.002** (0.001)	0 (0.001)	0.006*** (0.001)	0.005*** (0.001)	0.01*** (0.001)
walking distance (hours)	-0.06*** (0.02)	-0.06*** (0.02)	-0.04** (0.02)	-0.05 (0.03)	-0.05 (0.03)	-0.03 (0.03)
walking distance – squared	0.01*** (0.002)	0.01*** (0.002)	0.01*** (0.002)	0.01** (0.004)	0.01** (0.004)	0 (0.005)
same primary school	0.24*** (0.04)	0.24*** (0.04)	0 (0.04)	0.17*** (0.04)	0.16*** (0.04)	0.03 (0.04)
same classroom (t)	0.18*** (0.03)	0.18*** (0.03)	0.19*** (0.03)	0.1*** (0.03)	0.1*** (0.03)	0.13*** (0.03)
same classroom (t-1)	0.72*** (0.04)	0.71*** (0.04)	0.71*** (0.04)	1.09*** (0.04)	1.09*** (0.04)	1.08*** (0.04)
same track – vocational section	-0.19 (0.17)	-0.23 (0.16)	-0.12 (0.15)	-	-	-
same track – regular section	0.41*** (0.11)	0.41*** (0.11)	0.38*** (0.12)	-	-	-
emission – vocational section	0.61*** (0.16)	0.66*** (0.16)	0.63*** (0.16)	-	-	-
same sex	0.28*** (0.03)	0.27*** (0.03)	0.25*** (0.03)	0.33*** (0.03)	0.32*** (0.03)	0.31*** (0.03)
same ethnic group	0.19*** (0.05)	0.18*** (0.05)	0.15*** (0.05)	0.19*** (0.05)	0.18*** (0.05)	0.16*** (0.05)
grades similarity	0.18* (0.11)	0.19* (0.11)	0.17 (0.11)	0.09 (0.1)	0.09 (0.09)	0.08 (0.09)
factbrev reception	-0.01 (0.02)	0.01 (0.02)	-0.01 (0.02)	0.02 (0.02)	0.02 (0.02)	0 (0.02)
factbrev emission	-0.01 (0.02)	-0.04** (0.02)	-0.01 (0.02)	-0.09*** (0.02)	-0.1*** (0.02)	-0.1*** (0.02)
factbrev similarity	0.05 (0.08)	0.04 (0.07)	0 (0.07)	-0.05 (0.07)	-0.07 (0.07)	-0.08 (0.07)
cultural agreement – music	-	0.07*** (0.02)	-	-	0.09*** (0.02)	-
cultural agreement – youtube	-	0.02 (0.02)	-	-	-0.02 (0.02)	-
cultural status similarity	-	0.07 (0.08)	-	-	0.21*** (0.07)	-
parental ties (wave 1)	-	-	0.69*** (0.05)	-	-	0.66*** (0.05)

*p<0.1; **p<0.05; ***p<0.01

See Appendix 5B for the full models.

1.3.1. Network Endogenous Effects (Model 1)

Starting with network endogenous effects in Model 1, we can see that most of these are statistically significant and with large effect sizes. This is particularly true for mutuality and transitivity, both of which are strong and positive in all four schools: students tend to reciprocate ties and to close transitive triangles (unsurprisingly). As for their interaction, it is always negative, coherent with Block's (2015) prediction that shared friends matter less for individuals who already had some prior contact.

For centrality terms, things appear slightly more complicated. In Paris 1, Savoie 1 and Savoie 2, the overall pattern can be described as consisting of a negative effect for both indegree popularity and outdegree popularity, and a positive one for outdegree activity. There are some minor differences across these three schools, but they can be considered secondary (outdegree activity is not statistically significant in Paris 1, nor indegree popularity in Savoie 2). This means that students are *more* likely to send nominations if they have already sent many in the past, but that they are *less* likely to receive nominations if they have either sent or received many in the past (net of other modeled effects, notably mutuality). As for effect sizes, they are rather small for indegree popularity and outdegree activity, but the outdegree popularity effect is much larger than the other two[123].

How can we interpret this? Regarding the positive outdegree activity coefficient and the negative outdegree popularity one, the explanation I find most likely is that these reflect the fact that different students have different conceptions of the notion of "friendship", with some being very extensive (mere comrades are considered friends) and others restrictive (only close friends are nominated). As a result, students with extensive conceptions tend to have high outdegrees compared to their indegrees, thus the negative outdegree popularity (higher outdegree means lower indegree on average[124]). To a lesser extent, the outdegree activity effect should also capture the fact that "super-nominators" send ties more easily overall.

As for the negative indegree popularity effect, this is more surprising: it contradicts the general expectation that friendship networks are marked by popularity phenomena, with high-status students attracting a lot of friends through a "the rich get richer" type of process.

123 Taking Paris 1 as an example, the coefficient for indegree popularity is -0.03. Since this is a linear effect, we can compute that, if a node i has to choose between sending a tie to a node with indegree 10 (i.e. 10 received nominations) and another with indegree 0, the log-odd difference is of -0.3, meaning that the first tie would be 1.35 *less* likely to happen than the second (1/exp(-0.3) ~= 1.35).

124 This can also be seen from descriptives about the networks: the dispersion of outdegrees is much stronger than that of indegrees (see Appendix 3B).

On the contrary, friendship networks seem to be rather "egalitarian" in Paris 1, Savoie 1 and Savoie 2, with no tendency toward the concentration of ties toward well-connected individuals. Indeed, this tendency is even negative, which suggest a decentralization of ties over time – networks become flatter still, in later years.

Moving on to Paris 2, it clearly stands apart as per its centrality pattern. Outdegree popularity and activity effects are of the same sign as in the other schools, but they are notably larger (even when accounting for the fact that some effects are based on the square root of degrees in Paris 2, which mechanically inflate the size of coefficients). Moreover, the indegree popularity effect is positive, which means that, unlike in other schools, there seems to be a process of reinforcement of high-status positions through a popularity process. Thus, the networks appear to be more hierarchical in Paris 2 than in the other schools, regarding both popularity and activity.

One piece of explanation for this may come from the very high average degree in Paris 2: students there simply nominate far more friends altogether. The dispersion of outdegrees is stronger here too (some "super-nominators" reach up to 50 or 60 friendship nominations). This may indicate a more intensive sociability in Paris 2 but is also very likely due to students having on average a more extensive conception of friendship[125]. In any case, the concentration of ties may be more pronounced because of this large number of nominations: weak ties can be more easily concentrated and accumulated than strong and intensive friendships, which generally demand significant time investment, thus imposing a practical upper boundary on the number of close friends one may have. To support this interpretation, another SAOM is estimated, this time not on the networks of very good friend nominations, but on the networks of 5-friend nominations (i.e. the five friends with whom the respondent spends the most time at school). Since each student could emit a maximum of five ties for this measure, the density of the network is mechanically much lower, and weak ties are forcefully evacuated from the answers. Results are in Appendix 5C: what interests us here is that the indegree popularity parameter is now negative, indicating a de-concentration of received ties. Therefore, popularity processes in Paris 2 seem to operate at the level of weaker ties and extended sociability circles. When looking at the most intensive friendship ties, however, the networks appear significantly less hierarchical.

125 Why this is the case remains unclear to me, and for it to be understood would demand a dedicated investigation, probably based on qualitative evidence (unfortunately, Paris 2 is the school where I could conduct the least interviews). It may be linked to a more extensive conception of friendship specific to immigrant urban suburbs (van Zanten 2000 mentions this hypothesis as well), but the evidence for this remains rather limited for now.

1.3.2. Imposed Foci (Model 1)

Four types of imposed foci are tested in Model 1: classrooms, residential distance, former primary schools and, in Paris 1 and Savoie 1, special tracks. Effects pertaining to the first three are meant to capture propinquity processes among students, whereas those pertaining to special tracks should capture a mixture of propinquity and homophilic selection.

Unsurprisingly, being in the same classroom strongly increases the odds of friendships in all schools, both for the current and previous wave's classroom. The former logically exhibits much larger effect sizes.

In Paris 1, residential distance does not seem to play a role in friendship formation, either directly or through the mediation of the former primary school. In Paris 2, there is an effect of residential distance, but not of primary schools. As expected, friendships become less likely as students live further away (the effect is negative), but with a decreasing relevance of larger distances (the squared effect is positive). The pattern is reversed in Savoie 2: no clear effect of residential distance (the parameter is not statistically significant for the linear distance), but a homophily effect for the former primary school, of a moderate magnitude (as a benchmark, it is roughly half that of gender – 0.17 vs 0.33 for coefficients, meaning approximately 1.19 vs 1.39 for odd-ratios). Finally, in Savoie 1, all three parameters are statistically significant with substantial effect sizes, suggesting that residential propinquity matters much more in this school than in the other ones. Note that the fact that the coefficient for residential distance appears smaller there than in Paris 2 (-0.06 vs -0.21) is misleading: since the coefficient gives the log-odd change for each additional hour of walking time and given that residential distances are larger overall in Savoie 1, the log-odd decrease applied to most ties is in fact much larger than in Paris 2. Moreover, the marginal decrease of the effect, represented by the coefficient for the squared distance, is larger in Paris 2.

These results are coherent with what we know of the residential structure of the four schools. Paris 1 has the smallest recruitment sector, combined with the densest surrounding network of public transportation – it is therefore not surprising that residential propinquity has no effect there. Paris 2 has a relatively good transportation network as well, but a larger recruitment sector due to it being a private school; and students come from a large number of primary schools, some of which are very close to each other (there is a high population density in the school's town). As for Savoie 2, the town has sparser transportation and, most importantly, several primary schools correspond to villages or small towns, relatively isolated from each other: thus it makes sense that residential propinquity is best captured through the discrete variable of primary schools. Nevertheless, Savoie 2 is located in a central town in which many students live, so the residential structure is not completely fragmented either (it is more of a core-periphery structure). Finally, Savoie 1 has the largest recruitment sector, composed of several small villages, and lacking almost any form of public transportation, thus the large effect sizes for both residential distance and the former primary school. It should also be noted that in all the schools where the former primary school has an impact on tie formation, this effect decreases in time, disappearing in the later periods covered by the model (see Appendix 5B with the complete model specifications for the time dummies showing this).

Finally, what of special tracks? In Paris 1 and Savoie 1, differential homophily effects are statistically significant and have large effect sizes for students in the normal track (i.e. the large majority of the population), but not for those in the special tracks. This seems to suggest that students from the normal tracks are reluctant to nominate those in the special tracks, but not the other way around. Therefore, there might be a status asymmetry, with a form of stigma being specifically associated to students from the special tracks. Interestingly, this pattern is similar for the elitist track of Paris 1 and the pre-vocational one of Savoie 1. As for the emission effect, it captures the likelihood of nominations from students in the special track toward "regular" students, relative to nominations going the other way around. The coefficient is not statistically significant in Paris 1, but is significant, positive, and large in Savoie 1, meaning that students from the special track send much more nominations toward other students than they receive – they also send more toward other students than toward their own group. This further confirms the existence of a hierarchical pattern, with this small group of students being marginalized by the majority of the student population. In Paris 1, despite the differential homophily, the hierarchy appears less clear. This may come from students in the elitist track of Paris 1 having more social resources than those of the vocational track in Savoie 1, most of whom have severe difficulties at school and are from working-class backgrounds. Alternatively, it may also reveal an internalization of school criteria by the majority of students: the vocational track is an imposed orientation, which is perceived as a failure in academic terms, so students who end up there are frequently seen as stupid or lazy by their peers (see chapter 6, as well as Palheta 2012).

1.3.3. Expressed Dispositions (Model 1)

Four individual attributes susceptible to eliciting homophilic selection are tested in Model 1: gender, ethnicity, academic achievement, and socioeconomic background. It seems reasonable to assume that the corresponding effects capture selection processes relatively well, due to the presence of many relevant controls in the models.

The coefficients for gender and ethnic homophily are positive and statistically significant in all four schools, indicating higher probabilities for ties among same-gender and same-ethnicity peers. Effect sizes are similar in Paris 2, Savoie 1 and Savoie 2: around 0.3 for gender homophily (0.47 in Paris 1) and 0.2 for ethnic homophily (0.13 in Paris 1). Thus, gender appears to have more impact than ethnic background: all other things being equal, a student is, for a given simulated mini-step, about 1.35 times more likely to nominate a same-gender peer as a very good friend, and 1.22 times more likely to nominate a same-ethnicity one (respectively 1.60 and 1.14 in Paris 1)[126]. Note that even though these differences may not seem enormous, they accumulate over several mini-steps in the simulation, so eventually the impact of gender selection on the network is clearly larger than that of ethnic selection.

Homophilic selection based on academic results also appears to be strong in Paris 1 and Paris 2. The corresponding coefficients are positive, significant, and rather large. The similarity scores used by SIENA give the log-odd change for the maximum range of the independent variable, so the best and worst student at the school are 1.77 times (Paris 1) and 1.31 (Paris 2) less likely to nominate one another than same-graded peers[127] (again, at any given mini-step). In Savoie 1, however, the effect of grade similarity is weaker (log-odd of 0.18, odd-ratio 1.20), and is only significant at the 90% confidence threshold. As for Savoie 2, there appears to be very little selection of same-graded peers, if any: the coefficient is close to 0 (0.09) and is not statistically significant. This may be due to the fact that Savoie 2 uses a non-numerical grading system, which may make academic labeling less precise and/or less pervasive within students' subcultures. Alternatively, it may simply introduce more noise in my measure of academic performance. Furthermore, I have reason to believe that Savoie 2 is the school with the least diversity in terms of academic performance, with fewer students with severe learning difficulties compared to the other schools (see chapter 7, section 2.1.1). This may also explain why grades seem to have little impact on students' friendships.

Now, what of our main variable of interest, that is, socioeconomic background? In that regard, results markedly differ across schools. In Paris 1, the homophily effect is statistically significant, and rather strong: the log-odd coefficient is of 0.36, suggesting that, at any given mini-step, the highest- and lowest-background students at the school are about 1.43 less likely

126 exp(0.3) ~= 1.35; exp(0.2) ~= 1.22; exp(0.47) ~= 1.60; exp(0.13) ~= 1.14.

127 exp(0.57) ~= 1.77 and exp(0.27) ~= 1.31.

to nominate one another than same-background students[128] – which is far more than ethnic selection, but still less than gender and academic selections. This aligns well with the descriptive result of high aggregated levels of socioeconomic homophily in Paris 1 (cf. chapter 4). In Savoie 1 and 2, however, the effect is close to 0 and not statistically significant.

As for Paris 2, the effect is positive, but small (0.11) and barely significant at the 90% threshold. Nevertheless, it is the only school where there is time heterogeneity in the socioeconomic homophily effect. To understand what this implies, remember that SIENA aims to simulate the right value of the target statistic (here socioeconomic similarity). Having time heterogeneity means that, although the model simulates the correct total value of the statistic over all waves, it distributes it unevenly across them: simulated networks exhibit too much homophily at certain waves and not enough at others. To solve this, it is necessary to add time dummies to the models, that is, to allow the value of the parameter for socioeconomic homophily to differ across waves. Using such time dummies reveals that, in the transition from wave 3 to wave 4, the homophily effect goes up, reaching a significant and large coefficient of about 0.5. In other words, in Paris 2, there seems to be very little socioeconomic selection up to wave 3, but a strong one between waves 3 and 4 exclusively. Additional tests suggest that this selection disappears again in the transition from wave 4 to 6 (Appendix 5D). This is a peculiar pattern and is difficult to explain: it may come from a sort of “catch-up” effect after wave 3, where the friendship network happens to exhibit less socioeconomic homophily than in all other waves (cf. chapter 4, Figure 4-3). In any case, this seems exceptional, and homophilic selection on socioeconomic background remains weak most of the time in Paris 2.

Altogether then, and assuming that the modeled homophily affects approximate selection processes reasonably well, these results suggest that students have, on average, a strong propensity to select same-background friends in Paris 1; an uneven but mostly weak propensity in Paris 2; and none at all in Savoie 1 and 2. Once again, we find that socioeconomic background impacts students’ sociability quite differently across schools.

Finally, a quick word about the emission and reception terms for socioeconomic background. Reception, in particular, can be seen as indicating popularity patterns in relation to social origin: a positive effect would mean that upper-background students tend to receive more friendship nominations from other students (regardless of their own background), and a negative effect would on the contrary point to a greater attractiveness of lower-background

128 exp(0.36) ~= 1.43.

students. As a matter of fact, no consistent pattern emerges across schools: the estimated coefficient is positive and significant in Paris 1, but extremely close to 0 in the other three schools. Therefore, there does not appear to be an overall clear attractiveness of any social background. Interestingly, this suggests that the greater centrality of upper-class students in the friendship networks of Savoie 2 does not pertain to a higher status or social activity directly attached to socioeconomic origin. Much like with homophily, this centrality is most likely the result of complex compositional and structural effects.

1.3.4. Cultural Variables (Model 2)

In model 2, five effects are added to account for cultural variables: two effects for the square-root of the number of shared opinions about, respectively, musicians and video makers; and three effects for the scores of cultural status, namely one homophily, one emission and one reception effects. Note that I will not comment on the latter emission and reception effects, as I have no particular prediction nor framework to interpret them – they are here as controls, and are only presented in Appendix 5B.

The only effect that is statistically significant in all four schools is that of shared opinions about musicians. It ranges from 0.07 (Savoie 1) to 0.12 (Paris 1), meaning, for example, that two students that declared 4 opinions in common (out of a list of 25) are 1.15 and 1.27 times more likely to nominate one another at any given modeled mini-step, compared to two students with no common opinion[129]. Therefore, it seems that students tend to select friends with musical tastes akin to their own, and quite markedly so: the model predicts an effect of the same magnitude as gender homophily for a number of shared opinions between 9 and 16 (out of 25), depending on the school[130].

By contrast, agreement about video makers appears less impacting: it is statistically significant in only two schools, and with effect sizes at least two times smaller than musical agreement (0.06 and 0.03 in Paris 1 and 2 respectively, vs 0.12 and 0.08 for musicians). This is likely due to the fact that video tastes are more homogeneous than musical ones: popular video makers tend to be popular among most students, whereas musical tastes are more

129 The covariate is the square root of the number of shared opinions, so in this case we have exp(sqrt(4)*0.07) ~= 1.15 and exp(sqrt(4)*0.12) ~= 1.27.

130 The square root of 9 and 16 are 3 and 4; and we can see that the coefficient of gender homophily is always 3 to 4 times larger than that of musical agreement. For e.g., in Paris 1, the coefficient of gender homophily is 0.47 and that of musical agreement 0.12, almost 4 times smaller; so 16 shared opinions are predicted to have the same importance for a student’s (simulated) friendship choice as gender matching.

distinctive, and with clearer oppositions between different styles (e.g. rap and rock music)[131]. I also have reason to believe that musical tastes are more closely associated to other forms of cultural consumption that impact students' friendship selection, notably clothing styles (see chapter 6, section 3).

The scores of cultural status are meant to distinguish between highbrow and lowbrow cultural practices, and thus serve as a proxy for students' individual levels of cultural capital. The corresponding homophily effect is positive and significant in two schools, Paris 2 and Savoie 2, with weak to moderate effect sizes (in both schools, the maximum possible amount of similarity has a predicted impact lower than that of gender matching, but close to that of ethnic matching[132]). Interestingly, Paris 1, which is doubtlessly the school where both socioeconomic homophily and selection are the highest, does not exhibit a significant effect for this form of cultural homophily. By contrast, the effect is most pronounced in Savoie 2, the school where socioeconomic homophily and selection are the lowest.

In fact, descriptively, the homophily of cultural status is much higher in Paris 1 than in Savoie 2 (cf. Table 6-3 in chapter 6 or Appendix 5F). However, in the estimated SAOM, this is accounted for by other effects, notably the parameters of socioeconomic homophily and of academic homophily. On the contrary, in Savoie 2, other model coefficients hardly generate any cultural homophily, such that the model needs to implement a direct selection effect. What this suggests is that individual-level differences in cultural capital, and their impact on friendships, are well-captured by parental occupation in Paris 1 and, to an extent, Savoie 1, but much less so in Paris 2 and Savoie 2. Indeed, considering the correlation (Pearson's R) between students' cultural statuses and their score of socioeconomic background, we find different values across schools: 0.45 (Paris 1), 0.08 (Paris 2), 0.15 (Savoie 1) and 0.13 (Savoie 2) (NB: these are descriptive correlations, without any controls nor significance tests). To sum up, then, one might say that, in Paris 2 and Savoie 2 – the two private schools –, students seem to select friends with similarly highbrow or lowbrow cultural practices, but that this makes for a poor predictor of these friends' socioeconomic background. This is an important result, which will be further explored in the next section through the use of contribution

131 On average, for each listed musician, and considering only "affirmative" opinions (i.e. "I like it" or "I do not like it", without counting "I do not know it" or "no opinion"), the modal answer is chosen by 61% of respondents. For video makers, this is 71%. Therefore, it seems that opinions are more consensual for video makers than musicians.

132 Log-odd coefficients are 0.13 (Paris 2) and 0.21 (Savoie 2), while ethnic and gender homophily have coefficients of respectively 0.31 and 0.19 (Paris 2) and 0.32 and 0.18 (Savoie 2).

scores. It may in particular hold a hint as to why socioeconomic homophily and selection so saliently differ across schools, which will be discussed in chapter 7.

Finally, one can note that, in Paris 1 and 2, the introduction of cultural covariates in Model 2 results in a reduction in the size of certain coefficients compared to Model 1, notably that of socioeconomic similarity[133]. In Paris 2 in particular, this effect was barely significant in Model 1, and no longer significant in Model 2. This seems to suggest a mediation effect of cultural practices on socioeconomic homophily – something we shall test for more thoroughly in the next section.

1.3.5. Parental Networks (Model 3)

The parental ties of wave 1 have a strong predictive power on the creation and maintenance of ties in subsequent waves. The coefficients are large and statistically significant in all schools, from 0.47 (Paris 1) to 0.97 (Paris 2). Moreover, in Paris 2, the introduction of this effect reduces the effect for socioeconomic similarity, which is no longer statistically significant in Model 3. As for other schools, things are less clear, but there also seems to be some small reduction in the effect sizes of socioeconomic similarity (Paris 1) and ethnic matching (Paris 1, Savoie 1 and Savoie 2). Therefore, it seems plausible that parental networks act as inducers of socioeconomic and ethnic homophilies – once again, this is something that will be explored more thoroughly in the next sections. Remember, though, that the covariate of parental ties may also capture unobserved heterogeneity in the strength of students' friendships in wave 1 (parents met because students were close friends), so it is unclear what causal relation really is at work here.

1.3.6. Additional Models in the Appendices

The results described so far apply to the evolution of friendship networks from waves 1 to 4. Are the relational processes of interest similar prior to wave 1, and between waves 4 and 6? The models in Appendices 5D and 5E can help to answer these questions. They should, however, be considered less reliable than the ones presented above: the cross-sectional ERGMs applied to wave 1 cannot control for temporal dependencies the way longitudinal models do, and the SAOMs applied to waves 2 to 6 are based on two time periods of one year

133 As mentioned in chapter 4, log-odd coefficients are difficult to compare across model specifications. However, as Mood (2010) points out, the introduction of additional effects to a model should generally inflate the other coefficients (because coefficients mechanically get larger as the model accounts for a larger share of the dependent variable's total variance). Therefore, a reduction in the size of coefficients can be interpreted safely – indeed, we actually obtain a conservative estimate from this, as the inflation of coefficients leads us to under-estimate this reduction.

each (w2→w4 and w4→w6), which makes them less precise than the above models, based on three periods of six months each. Pragmatically, then, we can check that the different specifications are globally coherent – that they tell the same overall story –, but we should not expect them to yield the exact same results either.

Starting with Paris 1, the cross-sectional ERGM of wave 1 also finds a net effect of socioeconomic similarity on tie odds. It is, however, less strong than what was found between waves 1 and 4. This effect is also found in the SAOM of waves 2 to 6, this time with a much larger effect size – more than twice that estimated between waves 1 and 4! At first glance, these results may suggest that socioeconomic selection increases over time (weaker effect in wave 1, larger effect in the model that goes up to wave 6). However, as I have already mentioned, the SAOM of waves 1 to 4 does not find time heterogeneity in the corresponding parameter; the model of waves 2 to 6 does not either. In other words, even though this latter model finds a very large effect of socioeconomic similarity, it also suggests that the strength of this effect is the same from waves 2 to 4 and from waves 4 to 6, which contradicts the idea of an increasing strength of homophilic selection. Therefore, and given the uncertainty associated with the different model specifications, it is also possible that this SAOM simply fails to control for network dependencies as well as the main one does, due to the larger time gap between waves, thus over-estimating the net effect of socioeconomic similarity. In any case, what is clear is that all the specifications find a net effect of socioeconomic similarity, strongly suggesting that there is indeed homophilic selection based on socioeconomic origin in Paris 1. Another interesting finding for Paris 1 is that the ERGM from wave 1 does find a significant effect of both the former primary school and the residential distance, whereas this was not the case in the SAOM presented above. Most likely, this shows that residential propinquity played a role at the beginning of middle school, but not from wave 1 onward.

Moving on to Paris 2, the ERGM of wave 1 finds a net effect of socioeconomic similarity, which is small but clear (statistical significance is very high). On the other hand, no such effect is to be found in the SAOM of waves 2 to 6. Remember that the main model found Paris 2 to exhibit a net effect of socioeconomic similarity only in the transition from waves 3 to 4, but not on the others. Altogether then, the picture remains the same: Paris 2 seems to exhibit socioeconomic homophilic selection at certain waves, but not others, suggesting that the effect does exist but remains relatively small and fluctuating.

Savoie 1 exhibits no net effect of socioeconomic similarity in the cross-sectional ERGM of wave 1. The SAOM of waves 2 to 6, however, does find such an effect, but exclusively at the transition from waves 4 to 6 – not from waves 2 to 4. The effect is large enough that it could hardly be explained away solely by the inclusion of better structural controls; thus there may have been an increase in socioeconomic homophilic selection toward the end of middle school, possibly in relation to the COVID crisis (e.g. students fell back on their out-of-school sociability networks during this period, which we know to be more homophilic than in-school ones). The finding that there is no such selection up to wave 4, however, is confirmed by these models.

Finally, Savoie 2 does not exhibit any net effect of socioeconomic similarity in the cross-sectional ERGM of wave 1. There is, however, a slight effect in the SAOM of waves 2 to 6, which is small and barely significant (90% interval only). Given the uncertainty associated with this specification, I would be inclined to believe that this is a consequence of the model not controlling perfectly for network dependencies, rather than an indication that homophilic selection appears toward the end of middle school. In any case, even if it does, the effect remains marginal, coherent with the weak socioeconomic homophily found in this school.

2. Testing for the Inducing Power of Several Processes on Socioeconomic Homophily

So far, I have tried to find a statistical model that can account for the main features of the observed evolution of friendship networks over time. This has been done by fitting SAOMs to the networks of “very good friend” nominations, mostly on the period spanning wave 1 to 4 (though extensions to wave 6 as well as to the initial first wave have been proposed in the appendices). In doing so, it has been possible to sketch out the combination of low-order processes that drive the formation, maintenance, and dissolution of ties among students. Nevertheless, this does not answer our driving research question; that is, how socioeconomic homophily emerges. Indeed, remember that SAOM aims to find a set of parameters that can jointly account for the structure of an observed network (as captured by the target statistics). Consequently, the value of these parameters can be interpreted as reflecting the strength of the relational processes that they represent in regard to the overall evolution of the networks. However, these parameters do not account for the importance of any one process *in regard to a specific structural feature*. In other words, parameters do not tell us how much transitive closure, ethnic assortative mixing, or classroom propinquity account for socioeconomic homophily specifically.

The rest of the chapter is dedicated to answering this question. This is done by means of a specific method, which I call "contribution scores". The rationale behind these scores is not new and can be found in various forms in the literature – though it remains relatively rare compared to the well-known process of specifying and estimating a SAOM. However, the exact way in which these scores are computed (embodied in equation (1) below) has not, to the best of my knowledge, been proposed before. Therefore, one might say that contribution scores are a new shell put around an old idea[134]. One consequence of this is that the method remains under-explored at this stage, and that it lacks benchmarks or recurrent results that could be identified through its use in several different studies. Consequently, the treatments that I propose below should be seen as exploratory; more research is needed to confirm the results that they offer, as well as the overall validity of the approach. In any case, beyond the particular question of the emergence of socioeconomic homophily, the following sections are also meant to explore new ways to extend the SAOM framework, and how it can be used to answer novel research questions.

2.1. Contribution Scores: A Method to Assess the Impact of Low-Order Processes on Macro-Level Outcomes in SAOM

2.1.1. Counterfactual Reasoning with Network Models

How can we assess the importance of various relational processes in explaining the emergence of a particular network feature? This question is difficult to answer because, by construction, the macro-level features of the network cannot be expressed as a sum of linear contributions from each model parameter. This is due to both the logit form of the model[135], and to its simulation-based estimation technique. Instead, all parameters *jointly* generate network features, through complex and usually intractable interactions throughout the simulation procedure (which is necessary to account for the intrinsic dependency of network observations). Therefore, the idea of a parameter's "importance" in regard to a given statistic, while intuitively seducing, is not straightforward to operationalize. This is actually a

134 Here, I must express my thanks to the research group on social networks of the University of Groningen, which gave me important feedback on an initial version of the method. This is particularly true of Tom Snijders, whose advice and insight have proven extremely valuable. Moreover, members of the DefSim project, directed by Andreas Flache, pointed out to me that the idea of deriving simulations from an estimated SAOM was similar to their own work on the empirical calibration of Agent-Based Models (ABM), which led to a very interesting discussion on this topic. See Flache et al. (2017) for a discussion of the link between SAOM and ABM. Needless to say, any mistake in the conception or implementation of the method remains entirely my own.

135 For any logit model, the absolute probability change implied by a log-odd coefficient for a given observation depends on the "baseline" probability of that observation. For example, applying a log-odd change of +0.5 will not have the same effect if the total log-odd of the observation goes up from 0 to 0.5 (corresponding probabilities are 0.5 and 0.62) or from -2 to -1.5 (probabilities of 0.12 and 0.18).

substantial question, not simply a technical one. As an analogy, one can think of multiple individuals that work together to accomplish a given task – say, building a house, or putting together a research project. It is clear that it is only their joint effort and cooperation that make it possible to complete this task, such that the end result cannot be thought of as a sum of individual efforts. Yet, we are generally willing to accept the idea that some of them are more *important* or *contribute* more to the task. The problem is the same with the importance of the parameters through the generative process: some likely impact the network, or a particular feature of the network more than others, but always through joint interactions among several parameters.

One way to circumvent this issue is to adopt a counterfactual approach. Taking the estimated model parameters for granted, it is possible to "experimentally" manipulate the value of certain parameters and to see how the predicted network structure evolves under these assumptions. This means asking questions of the form : "if the social process X changed, how much would the network feature Y change in return?" The process X is represented by one or several model parameters, and the network feature Y by one or several model statistics. Crucially, this is not a real counterfactual observation – there is no new observation beyond the original network – but rather a logical extrapolation from an estimated model. Moreover, this approach is only meaningful if the parameters are treated as capturing real-life social processes following the interpretation of SAOM discussed above.

In the literature, this general approach has been used with SAOM by Lakon et al. (2015) to assess the importance of selection and influence processes in explaining homophily of smoking behaviors among adolescents, as well as by Block (2018) to explore the implications of certain structural effects on network structure. Kruse et al. (2016) also used a similar method with ERGM, to look at the contribution of residential segregation to ethnic homophily among high-school students. Along a different, but related vein, Indlekofer & Brandes (2013) proposed a score that measures the "relative importance" of effects for SAOM. While these scores do not consider a particular network feature, but rather the network as a whole, they also rely on a counterfactual simulation, where the "importance" of an effect on the network is defined as the average probability change for tie observations when the related parameter is set to 0. Finally, Snijders and Steglich (2015) also use counter-factual simulations to study the sensitivity of network macro-features to variations in parameter values under SAOM. Interestingly, they do so not for sufficient statistics, but for other macro-features that cannot be captured by a model's statistic, being irreducible to a count of low-order configurations (e.g. the size of the largest component in the network).

2.1.2. Introducing Contribution Scores

Following the same logic as the aforementioned studies, I propose to define the *contribution* of any statistical effect to socioeconomic homophily as the predicted change in the macro-level amount of socioeconomic homophily when the corresponding parameter is set to 0 (note that this could generalize to any sufficient statistic other than socioeconomic homophily, but I will stick to this single example throughout the discussion for simplicity). To quantify the raw quantity of socioeconomic homophily in a network, I use the macro-level statistic typically used by network models, which is the sum over all ties of the absolute difference in socioeconomic scores among both nodes. In the case of SAOM, these absolute differences are further re-scaled on (0,1) and subtracted from 1 to give a score of similarity, which is the default measure used by the estimation software. By convention, I call this statistic S_{SAS} (SAS stands for Socio-Academic Scores, which is my measure of students' socioeconomic backgrounds). Therefore, given a model specification $\hat{\theta}$ with n estimated parameters $\hat{\theta}=(\theta^1,\theta^2...\theta^n)$, the idea is to set the *k*th such parameter θ^k to 0, to simulate a large number of random networks from the altered specification, and to compare the average value of S_{SAS} in these fictive networks to its observed value.

Note that S_{SAS} is highly sensitive to the number of ties in the network. Setting one parameter to 0 will likely change the predicted density of the network, thus making the values of S_{SAS} incomparable between the observed and simulated networks. To circumvent this issue, I follow Block's (2018) recommendation to adjust the density parameter in the altered specification, such that the predicted density matches that of the observed network[136].

136 For SAOM, this is done by re-estimating a model with all parameters offset to their value in the original specification, except the one whose contribution is tested, which is offset to 0, and the outdegree effect, which is the only one not offset. Therefore, the estimation will simply adjust the outdegree parameter to reach the right predicted density.
One question is whether one should re-estimate several model parameters, rather than only the outdegree one. Indeed, Snijders & Steglich (2015) suggest that, when performing simulations from a SAOM while changing the value of a parameter, one should ensure that simulated networks still have the right value of other target statistics, beyond the one(s) whose reaction to the parameter change is being investigated. Their idea is that affecting a single parameter "*may lead to quite unrealistic networks, having implausible values for the average degrees, degree variances, and so on*" (p.33). However, within the framework discussed here, this would imply somehow more complex assumptions at the micro-level: we would be assuming that a given social process "disappears" (which is the assumption made when setting a parameter to 0) *and* that other social processes adjust to compensate for the impact of this on several network features (degree distributions, gender homophily, reciprocity, etc.). Moreover, in the present study, networks simulated from altered specifications remain relatively realistic as per their overall structure (in this regard, controlling for network density seems to go a long way). Thus, I chose to let other parameters remain unchanged; however,

What we are interested in is the predicted change in the value of S_{SAS} when θ^k is set to 0. Let $\overline{S_{SAS}^{\widehat{\theta^{(-k)}}}}$ be the average value of S_{SAS} observed in simulated networks with θ^k set to 0. The quantity of interest is therefore $\overline{S_{SAS}^{\widehat{\theta^{(-k)}}}} - S_{SAS}$. However, one issue with this quantity is that it is very abstract: it is the change in the sum of similarity scores over all ties in the networks, which means that its value strongly depends on the size and density of the network as well as on the distribution of individual SAS scores. Not only is it hard to interpret, but it is also impossible to compare across different networks. Instead, we can express $\overline{S_{SAS}^{\widehat{\theta^{(-k)}}}} - S_{SAS}$ as a proportion of the total amount of socioeconomic homophily observed in the empirical network.

Here, it is important to note that S_{SAS} does not measure homophily strictly speaking: it gives the observed value of SAS similarity among ties but, in itself, does not tell whether this value is larger or smaller than what would be expected by chance. Therefore, homophily is best defined as the difference between the observed value of S_{SAS} and its expected value under a reference null distribution: we would say that the network is homophilic in regard to socioeconomic background if the observed is S_{SAS} smaller than the reference. The null I use is a simple SAOM that only considers network density, rate parameters from each wave to the next, and a reciprocity effect. It is estimated separately, then networks are simulated from the resulting specification and the average value of S_{SAS} in these simulations is taken as a reference[137]. Let $\overline{S_{SAS}^{null}}$ be the average value of S_{SAS} under the null distribution. The total socioeconomic homophily observed in the empirical network can therefore be quantified as $S_{SAS} - \overline{S_{SAS}^{null}}$.

This leads to the following formula. The *contribution score* of a parameter θ^k to socioeconomic homophily, according to the model specification $\hat{\theta}$, is defined as:

this may be either more or less appropriate depending on one's data and research questions. In general, the more parameters that are re-estimated to maintain the original value of their statistics, the lower the contribution score of the tested parameter will appear (because part of the effect on the network of setting it to 0 will be "absorbed" by the re-estimation of other parameters).

137 This is the simplest specification available from the *Rsiena* package. The fact that it accounts for reciprocity and for the amount of change in-between observations (rate parameters) means that some of the socioeconomic homophily observed in wave 1 should still carry over into later waves under this null, though it will decrease across waves very quickly. This seems like the most appropriate null, because SAOM models network evolution, not network-structure *per se*, so the null should also pertain to an evolution process from wave 1 onward. Alternatively, one could use an Erdos-Renyi model (random allocation of ties that simply maintains the density of the networks), or other null distributions such as CUG or QAP, depending on what makes sense theoretically.

$$C_{k,SAS}^{\hat{\theta}} = \frac{\bar{S}_{SAS}^{\hat{\theta}^{(-k)}} - S_{SAS}}{S_{SAS} - \bar{S}_{SAS}^{null}} \tag{1}$$

with S_{SAS} the observed value in the original network, $\bar{S}_{SAS}^{null}$ the average value under the null distribution, and $\bar{S}_{SAS}^{\hat{\theta}^{(-k)}}$ the average value observed in simulated networks with θ^k set to 0. $\bar{S}_{SAS}^{\hat{\theta}^{(-k)}} - S_{SAS}$ represents the average predicted change in the sufficient statistic S_{SAS} under the alternative model specification, and $S_{SAS} - \bar{S}_{SAS}^{null}$ is the amount of socioeconomic homophily observed empirically (i.e. the difference between the observed value of the sufficient statistic and its value under the null reference). Interpretation-wise, this can be seen as the predicted percentage change in socioeconomic homophily when the parameter of interest is set to 0. Positive contributions to homophily are therefore indicated by negative scores (homophily is reduced when the parameter is taken off).

The novelty of contribution scores compared to previous studies that also performed counter-factual simulations from network models (cf. the last paragraph of 2.1.1 above) is two-fold. First is the fact that the macro-feature whose change is investigated is one of the sufficient statistics captured by the parameters of the model, instead of being a network feature not directly related to the estimated model. This means that contribution scores can be used as a diagnosis tool for an estimated SAOM – how does the model account for specific statistics? –, not only as a fictional thought experiment. In that regard, the use of counter-factual scenarios is instrumental to the assessment of the contribution of relational processes *within the empirical networks*; the aim is to infer what has (most likely) happened in the studied situation, not to predict what would happen in alternative ones. Second, the predicted change in the statistic of interest is expressed as a proportion of its difference to a null distribution. This results in a metric that can be understood relatively intuitively, and that can be compared across different parameters within the same model, or even across different networks: for example, saying that reciprocity accounts for 20% of socioeconomic homophily in school 1, or that residential propinquity accounts for 10% of it in school 2. Note, however, that for inter-case comparison, percentage points do not refer to the same total quantity, so this should be taken into account in the interpretation of the results (10% of the total socioeconomic homophily does not mean the same depending on how homophilic a given school is).

In any case, three things are important to note when considering contribution scores. First, they are always computed based on an estimated SAOM[138]. As such, their value depends on the model specification that is used, and different specifications will yield different results. Simply put, this means the scores can only be as good as the model specification. The fit of the model should therefore be assessed beforehand, and the choice of the effects should be based on a careful mixture of theory and inductive adjustment to the data. Moreover, it is essential for the estimated model to have converged correctly in regard to the statistic of interest. In other words, the predicted value of S_{SAS} under the complete model specification should be extremely close to its value in the observed network.

Second, we should not expect the scores to reach the sum of -1 over all model parameters. The reason is, again, that micro-level processes jointly generate network features under SAOM (or ERGM). As such, when we set a parameter to 0, we effectively remove the processes by which it interacts with other modeled effects – and these interactions may themselves impact socioeconomic homophily. As a consequence, the joint contribution of two or more parameters cannot be assessed by summing their contribution scores. Instead, it is necessary to set them to 0 simultaneously and compute the corresponding score.

Finally, one should keep in mind that the pseudo-experimental manipulation of setting a parameter to 0 is not a real counterfactual, as it does not correspond to any observed data. Consequently, the scores are not meant to indicate what would happen in the real world, should certain social processes disappear. In particular, they cannot account for potential transfers among effects that may happen in empirical settings. For example, imagine that, in a school, classrooms are strongly segregated by socioeconomic background, such that classroom propinquity accounts entirely for the observed socioeconomic homophily. The corresponding contribution score would only tell us that, in the observed network, socioeconomic homophily can be attributed to the distribution of students across classrooms. However, it may be the case that students did have a preference for same-background peers, but that they did not need to express this preference, as the classroom distribution alone

138 Note that the same logic and formula could be applied to compute contribution scores from an ERGM. I favor SAOM here for the reasons expressed in section 1.1.2, that is, the fact that this class of models is more directly designed for reasoning in terms of low-order relational processes (which is essential for the counterfactual reasoning at the core of contribution scores). However, depending on the situation and on how well a given model captures the relational processes of interest, ERGM may be perfectly fitted to this approach as well.

provided them with satisfactory friendships. Should the school administration decide to redistribute students across classrooms, individuals may now perform homophilic selection to compensate for this, in which case the impact of reshuffling classrooms on socioeconomic homophily may be null. Therefore, contribution scores are not empirical predictions. They should be seen as a thought experiment informed by empirical data and meant to explore the implications of an estimated model.

To sum up, contribution scores quantify the change in socioeconomic homophily that is predicted by an estimated model when one or several relational processes are assumed to be null. This is meant to indicate the importance of the tested processes in accounting for the homophily effectively observed in the networks.

2.2. Empirical Application: Deriving Contribution Scores from the Estimated SAOMs

We now turn toward applying contribution scores to the models estimated in section 1. Out of the four schools, three clearly exhibit socioeconomic homophily in the evolution of their friendship networks, in the sense that the value of the SAS similarity statistic is higher than what would be expected by chance (remember that this statistic is the sum over all ties of their similarity score, which is equal to the absolute difference in SAS scores between the two nodes, re-scaled on (0,1) and subtracted from 1). However, in Savoie 2, this is not the case: the SAS similarity statistic observed in the networks is extremely close to the value expected under the null distribution (i.e. a simple SAOM with only the rate function and outdegree and reciprocity effects). Therefore, there is simply no socioeconomic homophily to be explained in Savoie 2, or at least not enough for a meaningful application of contribution scores[139]. For this reason, only the three schools Paris 1, Paris 2 and Savoie 1 are considered in the rest of the section.

Also, note that contribution scores are only computed for the SAOMs estimated on waves 1 to 4, not on the additional models presented in the appendices (SAOMs on waves 2, 4

139 This may seem contradictory with the results from chapter 4, where we did observe a little bit of socioeconomic homophily at certain waves in Savoie 2. However, the definition given of socioeconomic homophily is a bit different: in chapter 4, it was considered for each cross-sectional observation separately, whereas here, it is considered in the evolution of the network from wave 1 onward, with this first wave being taken for granted. In chapter 4's cross-sectional pictures, it appeared that there was socioeconomic homophily in waves 1 and 4, but not 2 and 3. Thus, starting from wave 1, there is in fact a (slight) decrease in homophily in the transition toward waves 2 and 3, though this is roughly compensated for by a (slight) increase between waves 3 and 4. Altogether then, socioeconomic homophily does not clearly appear in the overall evolution of the network, averaged over all transitions, though it may appear on cross-sectional snapshots of certain waves specifically. Additionally, this confirms that the socioeconomic homophily observed in Savoie 2 at certain waves was very weak, at least in the networks of "very good friend" nominations.

and 6). The reason is that these models are the most reliable ones at our disposal, i.e. the ones that should come closest to real-life social processes, for reasons explained earlier.

2.2.1. Choice of Effects

10 potential inducers of socioeconomic homophily are considered; that is, 10 hypothesized social processes that are represented by one or more model parameters. Strictly speaking, contribution scores assess the impact of model parameters, not of social processes *per se*. However, to the extent that these parameters can be interpreted as capturing certain social processes, we can similarly interpret contribution scores as approximating their inducing power on socioeconomic homophily.

Homophilic selection – hereafter "selection" for short – based on socioeconomic background, grades, ethnicity, and gender is represented in the SAOM by homophily effects on the corresponding attributes. The first three processes are expected to contribute to socioeconomic homophily: socioeconomic selection is of course the most direct way to induce socioeconomic homophily, whereas grades and ethnicity selection should indirectly contribute to it, since these attributes are correlated to socioeconomic background. Gender selection, on the other hand, should not contribute to socioeconomic homophily since gender and socioeconomic background are unrelated.

Propinquity (i.e. spatial proximity among individuals) is represented by the homophily effects pertaining to students' places of residence, former primary schools and classrooms. To test for the effect of residential propinquity, three parameters are jointly set to 0: the walking distance between students' homes, this distance squared, and the homophily parameter for former primary schools. Indeed, primary schools are directly dependent on students' places of residence, so they conceptually pertain to the same type of inducer, namely residential segregation. As for the effect of classroom propinquity, it is tested by setting the two classroom homophily parameters to 0 (current and previous year classrooms).

The contribution of special tracks is tested by setting the two parameters of differential homophily and the parameter of emission[140] to 0 (only in the schools Paris 1 and Savoie 1). Note that this likely captures a mixture of propinquity and assortative mixing: students in

140 Note that for other attributes, only the homophily parameter was set to 0. The reason is that the use of differential homophilies, rather than uniform homophily, changes the meaning of the emission parameter: it does not represent some inherent propensity to sociability anymore (as it would if modeled together with uniform homophily), but now pertains to hierarchical patterns among the two groups (the emission term represents the prevalence of ties going from group A to group B, relative to ties going from group B to A – and accounting for each group's propensity to self-nominate).

these tracks have more opportunities for contact among themselves (shared classes), but they are also well-identified by their peers, with an institutional label often acting as a stigma.

The contribution of transitive triadic closure (the social process by which individuals close transitive triplets) is tested by setting the transitivity (gwesp) effect to 0. Note that the 3-cycle effect is conceptually distinct from transitive closure, so it is not set to 0.

For all the parameters mentioned above, the model used to derive counterfactual scenarios is Model 1. Models 2 and 3 are only used for the parameters specific to each of them, that is, cultural tastes and parental ties respectively. Remember that these parameters raise endogeneity issues, so their interpretation in terms of selection processes is relatively risky – the related scores should be considered as indicative rather than conclusive. The contribution of cultural selection (choosing friends with similar cultural tastes) is tested by setting three parameters to 0 in Model 2: the homophily effect for cultural status (based on a continuous score that distinguishes between highbrow and lowbrow tastes) and the two homophily effects for cultural agreement about musicians and video makers (based on the square root of the pairwise number of matching opinions within closed lists of artists). Finally, the contribution of parental ties (having parents that know each other) is tested by setting the parameter for declared parental ties in wave 1 to 0 in Model 3.

For each of the 10 scenarios, I set the parameter(s) of interest to 0, then adjust the density parameter and perform 1000 simulations from the resulting specification, resulting in 1000 fictive networks for each modeled wave of the friendship networks (i.e. waves 2, 3 and 4). These are used for computing contribution scores, based on Equation 1 above.

2.2.2. Results

Table 5-4 gives the contribution scores under the 10 fictive counterfactual scenarios. These scores can be interpreted as the percentage change in socioeconomic homophily predicted by the model when the considered parameters are set to 0. Moreover, Table 5-5 shows the coefficients of the tested parameters in the estimated SAOMs – this is a reminder taken from Table 5-3 above (full model output), to make the interpretation of the contribution scores easier.

Note that some parameters are not statistically significant in the SAOM estimation, but are nevertheless tested through contribution scores, such as socioeconomic homophily in Savoie 1 (this is indicated in the corresponding cell in Table 5-4). By definition, any null effect will have a null contribution score (setting it to 0 does not change the specification whatsoever). Therefore, if we cannot be certain that a “true” effect size is different from 0 (following the usual definition of statistical significance), then we cannot be certain either that the corresponding “true” contribution score is different from 0. This means the contribution scores of non-significant parameters should also be considered statistically non-significant. These scores may still be different from 0, though, because non-significant parameters are nevertheless active in the SAOM simulation process (including the simulations that were used for the original estimation of the parameters’ value).

Furthermore, Figure 5-2 shows the distributions of the Socio-Academic Scores (SAS) similarity statistic in the 1000 networks simulated for each counterfactual scenario. The simulated distribution under the full model specification (Model 1) is also represented in rity or parental ties).

Table 5-4: Contribution Scores

Model	Scenario Name	Parameter(s) set to 0	Predicted % change in SAS homophily Paris 1	Paris 2	Savoie 1
Model 1	trans	transitivity (gwesp)	-0.35	-0.65	-0.55
	SAS	SAS similarity	-0.37	-0.28	-0.19 (n. sign.)
	ethn	same ethnicity	-0.03	-0.08	-0.18
	grades	grades similarity	-0.08	-0.03	-0.09
	sex	same sex	0.04	-0.03	0.00
	classroom	current year classroom + past year classroom	-0.09	-0.01	-0.05
	track	same * special track + same * normal track	-0.12	x	-0.03
	resid	same primary school + residential distance + residential distance squared	-0.02 (n. sign.)	-0.02	-0.14
Model 2	cult	cult_lvl similarity + music matching + youtube matching	-0.06	-0.20	-0.09
Model 3	parents	w1_parents	-0.08	-0.06	-0.26

Note: Example of interpretation: in Paris 1, when the transitivity parameter is set to 0, simulated networks exhibit on average 35% less socioeconomic homophily than in the observed network.

"n. sign." indicates that the corresponding parameter(s) in the SAOM are not significantly different from 0 at the 10% threshold (see Table 5-5). The contribution scores computed based on a model do not take into account the parameters that are only present in other models. For instance, the score of SAS similarity computed from Model 1 (-0.37 in Paris 1) comes from a model in which cultural similarity and parental ties are not considered. If we computed the same score based on Model 2 or 3, it would probably be smaller (part of the mediation effect of SAS similarity would be "explained away" by cultural similarity or parental ties).

Table 5-5: Estimated Parameters for the SAOM Models – Parameters tested through Contribution Scores (reminder from Table 5-3)

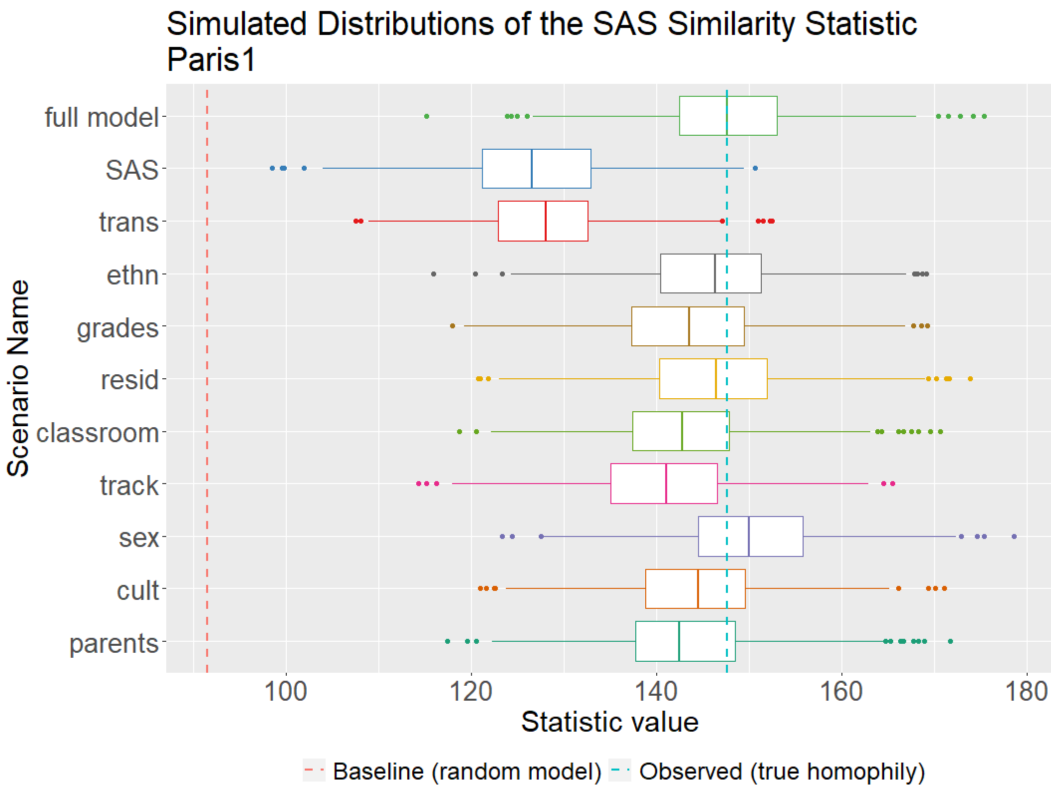
Model	Effect Name	Paris 1	Paris 2	Savoie 1
Model 1	transitivity (gwesp FF)	1.14*** (0.06)	1.56*** (0.06)	1.02*** (0.04)
	SAS similarity	0.36*** (0.14)	0.11* (0.06)	0.04 (0.07)
	SAS similarity * period 3	-	0.40***(0.13)	-
	same ethnicity	0.13** (0.06)	0.20*** (0.03)	0.18*** (0.05)
	grades similarity	0.57*** (0.17)	0.28*** (0.07)	0.19* (0.11)
	same sex	0.47*** (0.05)	0.32*** (0.03)	0.28*** (0.03)
	same classroom – current year	0.17*** (0.07)	0.08*** (0.02)	0.18*** (0.03)
	same classroom – previous year	0.49*** (0.07)	0.90*** (0.03)	0.72*** (0.04)
	special track – same * special track	0.03 (0.17)	-	-0.19 (0.17)
	special track – same * regular track	0.52*** (0.14)	-	0.41*** (0.11)
	special track – emission	-0.12 (0.19)	-	0.61*** (0.16)
	same primary school	0.03 (0.06)	-0.02 (0.03)	0.24*** (0.04)
	residential distance	-0.27 (0.29)	-0.21*** (0.06)	-0.06*** (0.02)
	residential distance – squared	0.28 (0.22)	0.07*** (0.03)	0.01*** (0.002)
Model 2	matching opinions – music	0.12*** (0.03)	0.08*** (0.01)	0.07*** (0.02)
	matching opinions – video makers	0.06** (0.03)	0.03** (0.016)	0.02 (0.02)
	cultural status similarity	0.24 (0.16)	0.13*** (0.05)	0.07 (0.08)
Model 3	parental ties (w1)	0.47*** (0.11)	0.97*** (0.05)	0.69*** (0.05)

*p<0.1; **p<0.05; ***p<0.01

Note: Standard errors of the coefficients are between parentheses. See Appendix 5B for the complete models outputs.

Figure 5-2a: Distributions of Simulated Statistics for Socioeconomic Homophily (SAS Similarity) under the Different Counterfactual Scenarios – Paris 1 and Paris 2

Note: boxes represent the values of the SAS statistic in the fictive networks. The blue vertical line corresponds to the true observed value of the statistic, and the red line to the average value observed under the null distribution. Therefore, the total amount of socioeconomic homophily observed empirically is represented by the distance between the red and blue lines, and the reduction in the SAS similarity statistic is the distance between the blue line and the center of the box. Contribution scores are defined as the reduction in the statistic divided by the total homophily, so they are represented by the proportion of the red/blue distance that stands right to the center of the box.



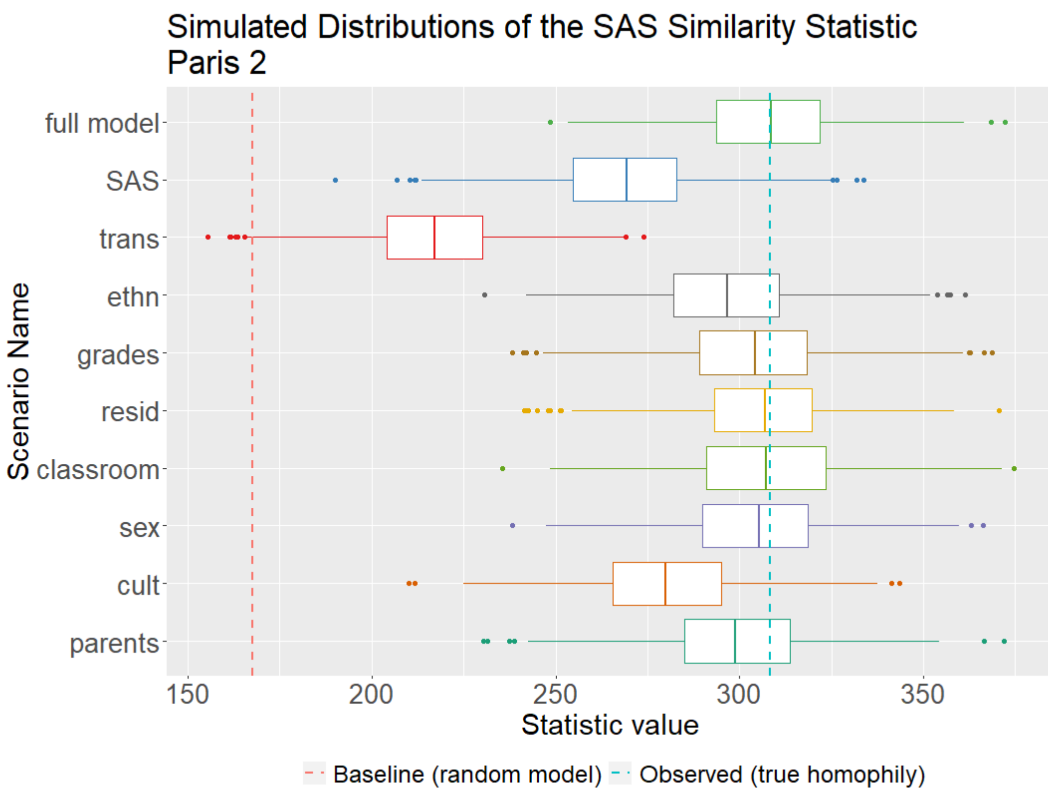
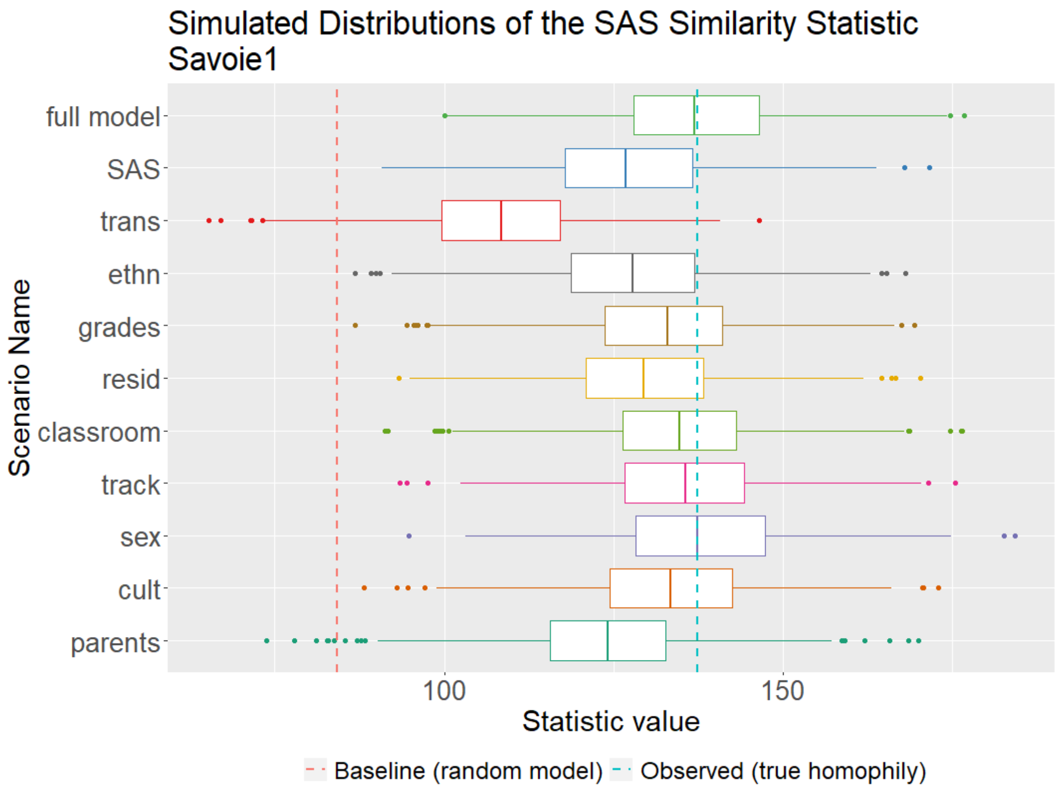


Figure 5-2b: Distributions of Simulated Statistics for Socioeconomic Homophily (SAS Similarity) under the Different Counterfactual Scenarios – Savoie 1

Note: boxes represent the values of the SAS statistic in the fictive networks. The blue vertical line corresponds to the true observed value of the statistic, and the red line to the average value observed under the null distribution. Therefore, the total amount of socioeconomic homophily observed empirically is represented by the distance between the red and blue lines, and the reduction in the SAS similarity statistic is the distance between the blue line and the center of the box. Contribution scores are defined as the reduction in the statistic divided by the total homophily, so they are represented by the proportion of the red/blue distance that stands right to the center of the box.



In Paris 1 and 2, the contribution scores of the SAS homophily parameter are -0.37 and -0.28 respectively. Therefore, assuming that students suddenly stopped selecting friends from similar socioeconomic backgrounds, the model predicts that socioeconomic homophily would diminish by about a third in these schools. By extension, this means that about two thirds of the total amount of socioeconomic homophily can be accounted for by other relational processes. Note that in Paris 2, this is largely due to the strong surge in socioeconomic selection modeled between waves 3 and 4, as the effect is rather weak and barely significant during the other transitions (cf. section 1.3.3. above). As for Savoie 1, the corresponding SAOM parameter is not statistically significant. As we cannot reject the hypothesis that there is no socioeconomic selection whatsoever, nor can we reject the hypothesis that socioeconomic selection plays no role in the apparition of socioeconomic homophily. Even assuming that it does, the contribution score is relatively low (-0.19), suggesting that less than a fifth of socioeconomic homophily is accounted for by direct homophilic selection[141].

The only parameter that exhibits consistently high contribution scores across the three schools is transitivity. When it is set to 0, the average socioeconomic homophily observed in the simulations drops by 35% (Paris 1), 65% (Paris 2) and 55% (Savoie 1). This may seem surprising, because, on its own, transitivity is not related to socioeconomic background: the process it entails (friends of friends tend to become friends) is agnostic as per the SAS of nodes. Nevertheless, transitivity can act as an aggravating factor for other social processes, exacerbating an initial uNo.alance toward homophilic ties.

Savoie 1 is also the only school where residential propinquity seems to have a strong impact on socioeconomic homophily. The joint contribution score of residential distance and of the former primary school is -0.14, meaning that homophily drops by 14% when these effects are set to 0. This probably comes from the rural location of the school, which implies that there is less public transportation and greater distances between students' homes. Supplementary analyses (not shown) indicate that this contribution is explained almost entirely by the effect of the former primary school, whose impact on socioeconomic homophily is much more decisive than residential distance itself.

Other covariates have rather low scores overall. The 'classroom' scenario exhibits a score of -0.09 in Paris and -0.05 in Savoie 1. Since the total amount of socioeconomic homophily to be explained in Paris 1 is much higher than in Savoie 1, it is fair to conclude that classroom propinquity has a much stronger inducing effect in Paris 1 than in Savoie 1. In Paris 2, there is virtually no such inducing effect (score of -0.01). Looking at descriptives, this can be easily explained: there are differences across classrooms in students' average SAS in Paris 1 and, to a lesser extent, Savoie 1, but not in Paris 2 (most likely due to language options in Paris 1 and Savoie 1, notably Latin and Italian that are favored by upper-class students).

141 This score of -0.19 remains relatively important when compared with that of other parameters which, although large and statistically significant in the SAOM, have lower contributions (e.g. residential distance in Savoie 1). The reason is that the sensitivity of the SAS similarity statistic to the SAS similarity parameter is of course very high: decreasing the parameter even by a little results straightforwardly in a decrease of the aggregated statistic. By contrast, decreasing the parameter of another effect, such as residential distance, will only indirectly affect the SAS statistic: the parameter change will decrease the number of short-distance ties in the simulated networks, and it is because these short-distance ties have *on average* a higher value of SAS similarity that the SAS similarity statistic will decrease as well. However, this decrease will be moderated by all the dyads for which the two covariates are not associated (some short-distance ties have low SAS similarity), resulting in a much lower sensitivity of the statistic to the parameter.

More so than classrooms, the special track of Paris 1 has a relatively clear inducing effect on socioeconomic homophily (score of -0.12). This may appear as a rather moderate contribution: from a qualitative standpoint, one would expect the special track to act as a powerful inducer of socioeconomic homophily, given that there are strong tensions between special track students and their peers, as well as it corresponding to a clear socioeconomic divide (this will be further discussed in chapter 7). Nevertheless, it is important to remember that there are only a few students on this track (~15 out of 80), whereas contribution scores are expressed in proportion to the homophily measured over the whole network. Therefore, this remains a notable score relative to the small number of students concerned, although it remains far from explaining the entirety of the strong socioeconomic homophily observed in Paris 1. As for Savoie 1, the contribution score of special track homophily is almost non-existent (-0.03), which is likely due to the even smaller weight of this small track relative to the entire network (~10 students out of 180).

Ethnic matching exhibits a noticeable score in Savoie 1 (-0.18) and, to a lesser extent, Paris 2 (-0.08), but virtually none in Paris 1 (-0.03). The correlation of ethnicity to socioeconomic background is relatively similar in these three schools, but Paris 1's coefficient for the ethnic homophily parameter does appear a bit smaller than in the other schools (remember that inter-model comparison of log-odd coefficients is generally biased, though). Beyond this (unclear) discrepancy in effect size, the difference in contributions may come from a difference in the structure of ethnic groups: although all three schools are ethnically diverse, the exact size and average socioeconomic backgrounds of the groups differ quite strongly across schools, such that the implications for the entire network of a given level of selection may differ (for instance, natives are in a majority in Savoie 1 but not in Paris 1 and 2). Finally, remember that contribution scores refer to an aggregate quantity (socioeconomic homophily) that is itself different across schools. Therefore, a similar "raw" amount of homophily that is explained away in a scenario will appear as a smaller score in Paris 1 than in the other schools, since there is more homophily altogether in this school.

Academic selection appears as a weak inducer: setting the corresponding parameter to 0 reduces socioeconomic homophily by 8% (Paris 1), 3% (Paris 2) and 9% (Savoie 1). As mentioned in chapter 3 (section 2.3.2), there is almost no correlation between grades and socioeconomic background in Paris 2, which explains this very low score. Even in Paris 1 and Savoie 1, though, the contribution appears to be small, as it accounts for less than a tenth of socioeconomic homophily.

The gender homophily parameter exhibits very low scores, hovering between 0.04 (Paris 1) and -0.03 (Paris 2). Since we know that there should be no contribution of gender selection to socioeconomic homophily apart from random variations in the average socioeconomic background of boys and girls in each school (which will always be slightly different, even if just by chance), these small scores can be seen as an indication of the typical error margin that one may expect from contribution scores. As a rule of thumb, it seems reasonable to consider scores inferior to 0.05 (in absolute value) as virtually null.

Homophilic selection based on cultural tastes – either cultural status or cultural agreement – seems to markedly contribute to socioeconomic homophily in Paris 2, with a predicted drop of 20% when the three corresponding parameters are set to 0. Since the estimated SAOM shows but a small effect of matching on opinions regarding video makers, this is mostly explained by the joint effect of matching on opinions about musicians and on having similar scores of cultural status. Supplementary analyses (not shown) suggest that matching on musicians is the most important inducer here, as cultural status is only weakly correlated to socioeconomic background in this school. In Paris 1 and Savoie 1, the contribution appears smaller, with scores of -0.06 and -0.09 respectively. Looking at Table 5-5, one can see that only the parameter for matching on musical opinions is statistically significant in these schools, not that of cultural status similarity. This indicates weaker cultural selection overall compared to Paris 2, which logically results in lower contribution scores.

Finally, parental ties account for a small part of socioeconomic homophily in Paris 1 and 2 (-0.08 and -0.06) but a large part of it in Savoie 1 (-0.26). How can we explain this? From Table 5-5, we can see that the SAOM coefficient for parental ties in Paris 2 (0.97) is much larger than in Savoie 1 (0.69). Moreover, the statistical association between parental ties and the socioeconomic similarity of students' pairs is stronger in Paris 1 than in both Paris 2 and Savoie 1: on average, students' pairs whose parents know each other have 0.42 less points of SAS difference in Paris 1, against respectively 0.19 and 0.20 in Paris 2 and Savoie 1[142]. In other words, parental ties do not seem to impact children's friendships more in Savoie 1 than in other schools (SAOM effect size), nor are parental networks more segregated socioeconomically there (pairwise correlation with SAS differences). Why, then, do we find a contribution so saliently higher in Savoie 1? This comes from the fact that parental networks

142 Dyads whose parents know each other have an average SAS absolute difference of 0.83 (Paris 1), 0.89 (Paris 2) and 0.92 (Savoie 1), whereas those whose parents do not know each other have an average SAS absolute difference of 1.25 (Paris 1), 1.08 (Paris 2) and 1.12 (Savoie 1). Thus, a difference between both types of ties of 0.42, 0.19 and 0.20 (all three are statistically significant to a t-test).

are denser in Savoie 1, both in absolute terms and relative to the density of students' friendship networks. Indeed, students' networks are about 3 times denser than parental networks in Paris 1, and 6 times denser in Paris 2, but only 1.6 times denser in Savoie 1. As a result, the increase in the odds of a friendship among students associated to the existence of a parental tie concerns more student pairs in Savoie 1 than in the other schools, hence its bigger impact on socioeconomic homophily.

2.2.3. Discussion: Low-Order Processes and the Emergence of Socioeconomic Homophily

How important is socioeconomic homophilic selection – the social process by which individuals select friends similar to them in socioeconomic terms – in explaining the observed homophily? In one out of the three schools (Savoie 1), the estimated Stochastic Actor-Oriented Model showed little sign of students selecting their friends based on socioeconomic background, once other relevant relational processes were accounted for (structural effects, as well as the effects of gender, ethnicity, grades, classroom, tracks, former primary school, and place of residence). This suggests that socioeconomic homophily can emerge even in the absence of a deliberate selection of same-background friends. As for the other two schools (Paris 1 and 2), counterfactual simulations derived from the estimated SAOM suggest that socioeconomic selection accounts for about one third of the total homophily – or, to be more precise, that homophily is divided by three under the assumption that there is no such selection. Therefore, homophilic selection appears as an important process in explaining homophilic tie formation but is far from being the only relevant explanation.

The other major social process contributing to the emergence of socioeconomic homophily appears to be transitive triadic closure. In all three schools, counterfactual scenarios with the transitivity parameter set to 0 predict much lower amounts of socioeconomic homophily, from 35% to 65% less. This is coherent with previous studies that found triadic closure to reinforce initial levels of homophilic selection, both in empirical contexts (Goodreau et al. 2009) and through stylized simulations (Asikainen et al. 2020; Foster et al. 2011). Crucially, triadic closure cannot, on its own, induce homophily: it simply reinforces an initial imbalance in tie probabilities. In that regard, the socioeconomic homophily predicted by counterfactual scenarios with no socioeconomic assortative mixing cannot be explained by triadic closure alone: other social processes must be at work as well. What is particularly interesting is that most other SAOM parameters, considered in isolation, exhibit relatively low contribution scores. Classrooms in Paris 1, ethnicity in Paris 2 or former primary schools in Savoie 1 do not account for more than 15% of the total socioeconomic homophily in each, and some other effects even less so (not even 10% for grades). In Paris 2 and Savoie 1 in particular, the total amount of socioeconomic homophily to be explained is not even that large in the first place, so these are small contributions indeed. Yet these relational factors seem to be sufficient to "tilt" the network structure toward socioeconomic homophily, such that transitive processes then aggravate this initial imbalance. Therefore, it appears that, even in the absence of direct assortative mixing, homophily can emerge as the joint product of several minor inducing processes, that get combined and reinforced by transitive triadic closure.

This strongly echoes with previous findings that strong segregation can emerge from weak homophilic inclinations on the part of individuals. This is notably the case in Schelling's famous model of spatial segregation (Schelling 1969), and the same idea has been extended to homophily in social network contexts (Stadtfeld 2018). What we see here is that it is not even necessary for this weak assortative mixing to directly pertain to the attribute of interest: assortative mixing on correlated attributes, or propinquity processes, can be enough for noticeable homophily to emerge.

Conclusion

The aim of this chapter was to identify the relational processes that entail socioeconomic homophily among students. These processes cannot be observed directly, but they can be inferred from the observed networks: by making certain formal assumptions (e.g. using the logit function), it is possible to predict the network structure that *should* be observed under a certain set of processes, and, from there, to derive the combination of hypothetical processes that matches the observed network structure. This is the rationale behind the Stochastic Actor-Oriented Model (SAOM), which was used to identify the set of low-order processes most likely to explain the observed evolution of friendship networks in the four schools. Moreover, I proposed to push the logic of these models one step further, by manipulating the parameters of an estimated model to explore counter-factual scenarios. The idea behind the resulting contribution scores is that the importance of a given parameter in explaining a certain network feature can be approximated by the predicted change in this structure that follows from neutralizing the effect of the parameter. Crucially, this is not meant to serve as an empirical prediction, but rather to assess how much the estimated model relies on the different parameters in order to account correctly for the structure of interest – in this case, socioeconomic homophily.

The first important result that came out of these treatments is that socioeconomic homophily does not primarily result from a propensity of students to select same-background friends. In all the schools but Paris 1, the estimated models found the effect of socioeconomic similarity on friendship formation to be moderate at best once other relevant factors were controlled for. The effect appeared at certain time periods only (between waves 3 and 4 for Paris 2, and 4 and 6 for Savoie 1), or not at all in the case of Savoie 2. By contrast, the parameters for gender matching, ethnic matching or academic similarity appeared larger and more consistent across waves, suggesting that these attributes are more important in orienting students' friendship choices (directly, at least). Moreover, even in Paris 1, which is the one school where socioeconomic selection appears to be clear and strong over the entire period, the use of contribution scores revealed that it accounts for less than half of the total socioeconomic homophily.

This result has two immediate implications, the first theoretical and the second methodological. Theory-wise, it suggests that discrimination and explicit preferences are not the main drivers of socioeconomic homophily – at least not in the case of young adolescents. Upon observing segregated friendship networks, one should therefore refrain from immediately assuming that this reflects the state of individuals' preferences, or even of some unconscious psychological inclinations: there is no necessary correspondence between the friends a person has, and the friends they look for. The methodological consequence of this is that descriptive estimates about the aggregate level of homophily on the one hand, and inferential estimates obtained from multivariate models on the other, are complementary measures that cannot substitute for one another. In that regard, much of the Social Network Analysis literature on adolescent friendship tends to favor complex models, sometimes at the expense of simpler descriptives; this may be one of the reasons why socioeconomic homophily has largely flown under the radar so far.

If purposive selection does not explain socioeconomic similarity, then what does? From the theoretical framework of chapter 2, three classes of explanations can be put forward. First, same-background students attend the same foci, most notably imposed ones (remember that, in the models of this section, the effect of chosen foci is essentially subsumed within the estimates of socioeconomic selection – thus these are not the main explanation either). Second, disposition-driven selection can operate on other attributes that are, for various reasons, correlated to socioeconomic background, thus entailing socioeconomic homophily through compositional effects. Third, network processes entail complex dependencies between other relational factors, impacting the network structure in sometimes counter-intuitive ways.

Overall, the role of imposed foci does not appear to be decisive. In the most homophilic out of the four schools, Paris 1, residential propinquity has virtually no impact on friendship formation after wave 1. Students, it seems, are old enough to cross spatial boundaries relatively easily when they live in urban zones with good public transportation. In the Savoyard schools, spatial distance matters more, but the socioeconomic homophily that this entails remains small, because residential segregation itself is relatively weak[143]. In a sense, this relatively weak role of imposed foci can be seen as good news: it shows that the spatial divides within the recruitment sector of a school are not insurmountable, as many students have seemingly no issue reaching out to peers that do not live close to them. This supports the idea that mixing socially different neighborhoods within the recruitment sector of a given school can effectively result in contact opportunities among students from different backgrounds – though whether these opportunities are converted into effective friendships depends on several other factors. As for the distribution across classrooms, it strongly impacts friendship pairings, but there seems to be enough socioeconomic diversity within the classrooms of the studied schools for this to be a minor inducer as well. Of course, this implies that strongly segregated classrooms would likely generate high levels of socioeconomic homophily; this will largely depend on the policy of a given school administration.

Compositional effects resulting from homophilic selection on attributes that are correlated to socioeconomic origin probably play a more important role overall. In each school, at least one of ethnicity, academic results and cultural tastes elicit strong homophilic selection, resulting in a noticeable inducing impact on socioeconomic homophily. However, taken in isolation, each of these remain a secondary inducer, not much stronger than imposed foci are. In fact, these compositional effects remain less important overall than direct socioeconomic selection, at least in Paris 1 and 2. Therefore, though direct selection cannot account for most of socioeconomic homophily, no other process really can either; there

143 In Savoie 1, the absolute difference of students' SA scores accounts for less than one percent of the walking distance between students' homes (R2 of a linear regression as the level of the dyad). On average, same-background students live at 1.36 hours from one another by foot, whereas this is 1.49 hours for a working-class and an upper-class student (SAS difference of 2.44).

simply does not seem to be any factor that would single-handedly entail high levels of socioeconomic segregation in friendship networks.

This is where the role of network processes appears decisive. By aggregating, combining and reinforcing several minor inducers, they markedly aggravate the networks' initial imbalance toward socioeconomic homophily. This is why transitivity is the one parameter that exhibits the highest contribution scores across the different schools, even though it is not directly related to socioeconomic origin. In substantial terms, this suggests that enlarged sociability circles, friends of friends, or peer pressure, can push many students toward a personal network that is more homogeneous than what they would have actively wished for (which is not to say that they are unhappy with these homogeneous networks – I do not have evidence for or against this idea). In that sense, socioeconomic homophily can be said to be a social and relational phenomenon, more so than a psychological one.

This raises one final question: is all this specific to socioeconomic homophily? Is it the case that socioeconomic origin specifically is attached to so many derivative properties that such indirect inducement can occur? Or, on the contrary, are most homophilic patterns in friendship networks the product of a similar combination of minor inducers – perhaps due to the way adolescent sociability works in general? One simple way to answer this is to apply the same method of contribution scores to other types of homophily, such as gender and ethnic homophily. This is done in Appendix 5G, where I present the results in detail. The conclusion is straightforward: both gender and ethnic homophilies are primarily driven by direct homophilic selection. Similar to what it does to socioeconomic homophily, transitive closure acts as an aggravating factor for these as well; but what it aggravates is mostly the effect of direct homophilic selection, as imposed foci and compositional effects induce very little homophily. In fact, in the case of ethnic homophily particularly and for certain schools, the total impact of other relational processes beyond direct selection is to induce ethnic *heterophily*: when the effect of ethnic selection is set to zero, ethnic homophily is predicted to decrease over time even faster than it does under the reference model. In other words, it seems as if raw ethnic selection forcefully maintains a level of ethnic homophily well above what other relational processes lean toward.

The distinction between socioeconomic origin on the one hand, and attributes such as gender or ethnic origin on the other, will be further discussed in the next chapter, when considering the mental categories that students use to organize their relations. What we can already say is that homophilic selection appears much more pregnant for these attributes than for socioeconomic background (which can also be seen from their overall larger effect sizes in the SAOMs of Table 5-3 above). This is even more noteworthy given that ethnic and socioeconomic homophily tend to be of similar magnitudes in most schools, as we saw in chapter 4. Yet the mechanisms through which these homophilies emerge are different: ethnic background elicits more direct selection than socioeconomic background does – in a sense, more discrimination –, but it also correlates less to other structuring factors of students' social lives, such as academic results, places of residence and parental networks. Therefore, this pattern of several indirect inducers combining to produce homophily despite modest levels of direct selection may well be a distinctive trait of socioeconomic homophily specifically.

Chapter 6: How do Students See One Another? Indigenous Perception Categories and Socioeconomic Origin

In the previous chapter, we found that imposed foci and propinquity could not account for most of the socioeconomic homophily observed among students. This suggests that there are selection processes at work: students actively pair with others that are similar to them on important characteristics. This is partly the case of socioeconomic origin –depending on the school and wave – and, more clearly so, of ethnicity, academic results and cultural tastes. However, selection estimates in multivariate models are always tricky to interpret. Technically speaking, they do not provide an estimation of selection *per se*, but rather of the homophily that could not be accounted for by the other parameters of the model. This means it is always unclear whether they truly reflect selection, or simply some omitted foci. In particular, *chosen foci* are a tricky category: they are dependent on students' internalized dispositions but, at the same time, do not entail an inclination toward a certain type of friend – indeed, the mechanism of friendship formation at work is propinquity, not selection.

Among these chosen foci, many are what can be called "micro foci"; that is, small scale and time limited activities that are virtually impossible to trace, but nevertheless impact the opportunities for friendship formation. For instance, at school, some students like to play football, and others do not; some can eat outside with friends during lunch time, or leave the school earlier and hang out when a teacher is absent, whereas others are forced to stay at school or to go straight home; some participate in school trips and others do not. All of these micro foci are, from a model perspective, contained within selection estimates. This is not necessarily a bad thing, because such daily micro-choices are part of how dispositions are expressed, and, in a way, could be said to be constitutive of what we call friendship preferences in the first place – I want to have a friend that plays football with me, I want to have a friend that goes to eat lunch outside with me, etc. However, they are still different from discriminatory beliefs, labels, social stigma or cultural norms and taboos, all of which could also be at work in producing homophilic selection among students from different groups.

Moreover, quantitative models do not tell us exactly how selection operates, if it does. The relationship between meeting opportunities and tie formation seems straightforward, that individuals would bond with the persons they see more often is almost self-explanatory. By contrast, it is less clear what lies behind homophilic selection. Implicit to the concept of "selection" is the idea that students are active in sorting out their peers. Whether consciously or unconsciously, they choose who to be with, or, at the very least, behave differently with different peers, such that they end up being friends with some of them, and not with others. If students behave differently depending on their peers' characteristics, then it must be the case that they *perceive* these characteristics. How they do so, however, is something that cannot be known through statistical associations alone.

Indeed, perceptions are not spontaneous, but are mediated by psychological categories, most, if not all of which have been socially constructed. Consequently, the ability one has to select certain friends rather than others depends on the mental categories available to them, and through which they can differentiate among various alters. Crucially, we should not expect these categories to be a faithful representation of reality, but instead we should see them as perception filters, which stress some traits and mask others, granting them different levels of subjective significance. The statistical estimation of the effects capturing homophilic selection suggest that students perceive *something* that is related to the attribute of interest, but this does not mean that they directly perceive it, nor that their perception is accurate. How does one "know" – or feel – which peer is similar to them? Is it through conspicuous consumption? Musical tastes? Language habits? Or through trial and error, by progressively sorting the friends with whom one gets along – and if so, what are the behaviors that make two students get along?

This chapter aims to find the perceptive categories that mediate selective behavior among students. Incidentally, this is also a way to support, or reject, the idea that quantitative estimates do indeed reflect selection, rather than unobserved micro foci. To do so, I analyze the 38 semi-directive interviews that were conducted with students in three schools in wave 1 (Paris 1, Savoie 1 and Savoie 2). In essence, the point is to find traces of students' selective dispositions in their discourses, which should be seen as complementary from the reconstruction of these dispositions' effects through quantitative models. As I will explain, students' indigenous categories for discriminating among peers are especially apparent when they talk of their animosities, much more so than when they talk of their friendships. Therefore, the whole chapter is geared toward the analysis of conflict and tensions within peer groups. Nevertheless, the point remains to get a sense of the "tools of thought" (Lignier and Pagis 2014) through which students make sense of their social environment.

Section 1 details the methodological choices made in the conduct and analysis of the interviews. Then, sections 2 to 5 each cover a particular perceptive framework that is related to socioeconomic origin.

1. Method

1.1. Discursive Categories: Connection to Mental frameworks and Behaviors

It is impossible to directly observe the perception categories that students have in mind. Instead, the interviews allow for an examination of the *discursive* frameworks that are proposed by students. The connection between these discourses and the underlying perceptive categories is not straightforward: one may say things that they had never thought about before (in which case the question creates its answer), or they may not be able to put words on things that they did think about (this is particularly problematic given that there may be a socioeconomic asymmetry in students' ability to verbalize their thoughts in front of an adult). Nevertheless, there are good reasons to treat discursive categories as a relevant indicator for the examination of perceptive categories.

First, there are relations of mutual influence between discourses, practices and tools of thought. In particular, it seems quite unlikely that certain categories would be extremely salient in students' discourse, yet entirely inactive in shaping their perceptions of their environment. While we should not conclude that a category does not exist just because it has not come up in the interviews, it seems much more reasonable, on the contrary, to consider that a perceptive category exists if students talk about it (given it has not been suggested to them by the interviewer, another issue that I discuss below).

Second, my objective here is not to uncover the mental categories of individual students, but rather to examine the categories that may circulate in the social space of the school. The categories used by students to talk about one another are not indicative of what happens in their heads, as much as they are indicative of what happens in their peer groups. Indeed, students' sociability is intrinsically discursive, in the sense that the sorting, differentiating and ranking of peers are not performed in isolation by each individual, but instead in the course of social interactions, of which talking is an important part (typically, the case of gossips). The assumption I am making is that the sorting and judgment of others that underlies selective attitudes occurs primarily through collective interactions: students identify their friends (or enemies) by exchanging information about them with others, not just by directly observing and judging them through their own mental categories. Consequently, discursive categories have a key role in making perceptive categories active, operational; and probably in shaping them as a whole.

The connection between perceptive and/or discursive categories on the one hand, and selective behaviors on the other, is more problematic. The causal relationship could go both ways: students might select certain friends because of their perceptive categories, but they could also rationalize and justify existing relations by applying the categories that are available to them. This means that selection might operate on a given, unobserved characteristic, then be re-interpreted by students through another, correlated characteristic, which is for some reason easier or more legitimate to articulate discursively. Unfortunately, there is no easy way around this problem. Still, I would argue that there is a similar asymmetry to the one I mentioned above about the discourse/perception link: while the absence of a discourse does not mean that the corresponding behavior does not exist, it seems highly likely that the presence of a discourse will have an effect of some kind on behaviors. For example, students might strongly select their peers based on ethnicity, while never presenting those as relevant categories in their discourse. However, if it is the case that students mobilize ethnic discursive frameworks to differentiate between their peers, then we would expect these frameworks to impact their selective behaviors, at least in the absence of contradictory evidence.

1.2. Conduct and Analysis of the Interviews

For the analyses of this chapter, I use the 38 semi-directive interviews conducted between waves 2 and 3 (second year of middle school, age 12-13). These include 13 interviews in Paris 1, 9 in Savoie 1 and 16 in Savoie 2. There are no interviews from Paris 2, as the school administration did not give me the opportunity to conduct them. Moreover, the 14 follow-up interviews that took place one year later (after wave 4) are not considered here: there are too few of them to investigate the evolution of students' answers over time in a robust way, and, since they were mostly conducted with students that already answered in phase 1, they do not increase the sample size either. We shall come back to some of these in chapter 7, though. Appendix 6A gives sociodemographic information about the interviewed students.

All interviews were fully transcribed. Then, specific extracts were isolated in order to find the discursive categories used by students that may serve as proxies for the perception of socioeconomic similarity or dissimilarity. To that end, I selected the extracts where students expressed *a feeling of distance or disliking toward a peer framed through collective categories*. The rationale behind this coding rule is explained in the next three subsections.

1.2.1. Schemes of Perception used to Perform Collective Categorizations

In the interviews, two types of questions were asked that were meant to investigate social similarity among students. First, open questions asked respondents who they were friends with; whether there were students who they disliked; and whether they found some peers strange or hard to understand, regardless of whether they liked them or not[144]. For each of those, I would then ask for more details so as to lead the student to elaborate on the considered relationship. The aim was to inductively determine which types of reasons students put forward to explain their friendships, dislikes, or feelings of distance. Second, later on in the interviews, questions were asked about specific topics which I had reasons to believe could mediate socioeconomic perceptions. They pertained to academic results, bullying, popularity, musical tastes, clothing trends and ethnicity – but this time were not framed in terms of liking or disliking other students (e.g. "*what music do you listen to?*" or "*do you follow clothing trends?*", systematically followed by "i*s this something you talk about with other students?*"). Note that in some cases, not all of these could be asked during an interview, when some topics took too long to cover. Moreover, some of these subjects could also come up in the course of the interview without me specifically asking the dedicated question.

Then, I selected the extracts where students used collective categorizations to sort out their peers. By "collective categorization", I mean any instance where the respondent suggested some type of group-level differentiation among schoolmates (even implicitly). More precisely, this means that they (a) relate the behavior of certain peers to the fact that they belong to a broader group or category and (b) differentiate among different positions or attitudes regarding this group or category. For example, "*I dislike X because she is weird, she doesn't dress like a girl should*" is treated as a manifestation of generic gender norms, which are not consensual within peer groups (the respondent seems not to agree with X about how girls should dress). By contrast, something of the form "*I dislike X because she betrayed me,*

144 "*We have talked about your friends, as well as the people you dislike. Now, this is a difficult question, so it's okay if you don't know how to answer. Are there some students – which you may or may not like – that you find, I don't know, a bit strange, or different, or that you have trouble understanding? For example, that you feel are not on the same planet as you* [in French, "être dans le même délire", an expression often used by teenagers], *or that sometimes do things you don't understand?*"

we were friends and she bad-mouthed me" is excluded. While it does pertain to a collective norm – one should not bad-mouth a friend – there is no hint that the respondent perceives this conflict in terms of more generic positions (it could still be the case that the conflict is causally driven by macro-social differences, but inasmuch as this effect is not mediated by discursive categories, it falls outside the range of the present investigation).

Extracts were included even when they did not directly relate to socioeconomic positions: these include differentiation expressed in terms of gender, popularity at school, ethnicity, age, clothing, musical tastes and academic attitudes. Then, all the extracts were analyzed in order to identify the perceptive schemes used most frequently. Whether these could act as proxies for socioeconomic positions was only considered *ex post* (see below).

1.2.2. Distance and Disliking

One important result that came out of this process is that categories expressing inter-personal distance were much more frequent in students' answers than categories expressing proximity. Respondents generally had a lot of trouble offering reasons for their friendships; questions such as "*why do you think you two get along?*" would often puzzle them. By contrast, they could always offer some reason for disliking someone – and, in many cases, did so spontaneously when naming the people they disliked, without the need for me to specifically ask why. This is the reason as to why this chapter focuses on distance and animosity, rather than likability[145].

With that said, *distance* is not the same as *disliking*. Not all disliking nominations are indicative of distance. In fact, many animosities stem on the contrary from proximity: former friends with whom one got into a fight, or friends of friends that are seen as rivals for the time and attention of one's best friend, make up for a significant portion of the animosities expressed during the interviews. Conversely, distance does not necessarily imply animosity. In the interviews, respondents could mention someone as "weird" or hard to understand, while

145 Why this is the case is an interesting question in its own regard. Clearly, negative feelings are depreciated in most contexts (in particular for school children, as there are explicit incentives from teachers to "be nice to others"), making the need for justification more pressing. It might also be the case that the usual social representations of friendship sanctify it as a "pure" and "elective" relationship (similar to love), thus discouraging analytical justifications for its formation (note, however, that Bidart 1991 did manage to have adults justify their friendships, meaning that the inability to offer explanations of friendship formation might be specific to teenagers rather than intrinsic to the generic cultural conception of friendship). Moreover, beyond the sole case of friendship and disliking, it is often the case in sociological analyses that social norms are revealed through the analysis of deviance (Becker [1963] 2008). Indeed, the normal functioning of a consensual norm generally implies the forgetting of this norm by individuals, as they take it for granted. Likewise, similarity in friendship formation should induce a feeling of smoothness, ease or be self-evident for students, thus dissimulating the reasons for the normative convergence among them. By contrast, normative dissensions make it clear that there is something about which individuals disagree.

explicitly saying that they had no negative feeling toward them. At times, they could also mention a feeling of strangeness toward some of their friends' attitude. By definition, these non-antagonistic feelings of distance should not be captured by quantitative analyses of disliking networks. However, they are likely shaping friendship networks, as they may prevent students from selecting certain friends.

Therefore, in the coding process, I isolated extracts that were indicative of distance (i.e. without animosity) and of disliking, and that were framed through collective categories. Extracts pertaining to friendship or likability were not considered, neither were those indicative of "proximity dislikes" not framed through collective categories, and therefore not indicative of social dissimilarity (e.g. betrayal from former friends).

1.2.3. Subjective Significance of Expressed Opinions

Finally, I tried to objectify the subjective significance that each student seemed to attach to the judgments they formulated. For example, consider a student that expresses their disapproval and hostility toward a peer that chats during class. It is clear that "*I don't really like X because she talks in class, it can be annoying sometimes*" does not have the same weight as "*I hate X because she always talks in class, it drives me insane*". Surely, both are worth taking into account, especially given that one's virulence in expressing one's animosities also depends on one's personality and relationship to the interviewer, as students that were shy or uncomfortable during the interview tended to under-state most of their answers[146]. However, it is also important to differentiate among these differences, as they imply different levels of distance or conflict among students.

Therefore, for each extract, I coded whether it seemed to have a high or low subjective significance for the respondent. I relied on verbal markers of importance (e.g. "*really*", "*very much*", "*a lot*"), or on the contrary, of moderation ("*kind of*", "*somehow*", "*sometimes*", "*a little*"). I also considered the personal involvement of the student in their answer: whether they fully took on the considered classification framework ("*I think that rap music sucks*"), or showed their ability to perceive it, but mostly linked it to other students' behavior ("*many people think that rap music sucks*").

146 While I do not have the statistical power to back it up, I also suspect that there is a social class bias in the expression of animosities. Upper-class students seemed to express their animosities in attenuated ways, compared to working- or middle-class students – perhaps responding to the expectation of teachers (and probably parents) of peaceful relationships among students (which they might have internalized, or that they might simply display in front of me) (after all, and despite my very best efforts, I am an adult).

Finally, note that this question of subjective significance is distinct from the opposition of disliking and distance: one may express a dislike with low intensity ("*sometimes I find X a bit annoying*"), as much as a non-antagonistic distance with high intensity ("*I really can't understand X!*").

1.3. Categories of Social Dissimilarity

Eventually, seven discursive frameworks emerged from the analysis of the interviews, through which students could express their feelings or animosity. Note that these might not be the only frameworks ever used by students: it is clear that the "appearance" of a discursive category during an interview depends on how it has been conducted. As some likely inducers of socioeconomic distance were already part of my theoretical framework upon starting the fieldwork, I specifically designed some questions to capture these (e.g. cultural tastes); and having an awareness of these factors also meant that I would more easily notice when a respondent spontaneously used them, thus giving me the opportunity to encourage them to provide more details. Nevertheless, I did try to not force a category or answer upon respondents (e.g. asking "*how do you feel in class?*" rather than "*are you bothered when other students talk in class?*"), as well as the open questions on disliking and distance which were meant to allow for unexpected categories to emerge.

The seven discursive frameworks were the following:

1 *Academic framework*. Expressions of dissimilarity that pertain to school attitudes, performance, and formal distinctions (one's grades, track options, behavior in class or attitude toward homework).

2 *Cultural framework*. Expressions of dissimilarity that pertain to cultural practices and more broadly life-styles, including clothing, music, ways of speaking, tastes and leisure interests.

3 *Ethnic framework*. Expressions of dissimilarity that pertain to ethnicity, including migratory background, religion, language or race.

4 *Money/Consumption framework*. Expressions of dissimilarity that pertain to the economic resources of students and/or of their families.

5 *Age framework*. Expressions of dissimilarity that pertain to age-specific norms of behavior (e.g. being "mature" or "childish").

6 *Gender framework*. Expressions of dissimilarity that pertained to gender norms (e.g. being “girlish” or a “tom-boy”).

7 *Popularity framework*. Expressions of dissimilarity that pertained to popularity within peer groups (“popular” and “unpopular” students).

Among those, there was no clear evidence that the last three (age, gender, and popularity) mediated the perception of socioeconomic positions. While I did formulate certain hypotheses about how they may have done so, empirical analyses neither corroborated nor disproved them, mostly due to sample limitations (there were too few extracts that mentioned these frameworks)[147]. Consequently, they are not considered in the following treatments. This does not mean that they have no role in mediating socioeconomic perceptions, but simply that the current study is inconclusive in that regard (most likely due to limitations in the empirical design: I suspect these frameworks were not covered as well as the others by my interview grid).

This leaves us with four frameworks, namely the academic, cultural, ethnic, and economic ones. For each of those, I systematically counted how many interviews they appeared in, differentiating between disliking and distance as well as high low subjective significance. This is summarized in Table 6-1 below. For each framework (columns), the table indicates which interviews it came up in (rows). The code “dl” (dislike) indicates the expression of animosity, while “dist” (distance) indicates a non-antagonistic feeling of distance (i.e. strangeness or inability to understand). Both codes can be in the same cell, when this happens it means that the framework appeared at least twice in the interview. Moreover, each code is followed by a “(+)” or “(-)” sign to indicate, respectively, a strong or weak subjective significance.

Note that this coding does not differentiate between the different positions or opinions held by students regarding each framework. For example, the item “school attitudes” comprises cases of students complaining about classroom disruption, as well as students

147 In short, I suspect that:
(a) several elements of the transition from childhood to adolescence occur earlier for working-class students (the sexualization of relationships in particular), making them look more “mature” on average while upper-class students appear “childish”.
(b) gender norms widely differ across social backgrounds, with the distinction between boys and girls being more pronounced among working-class students (working-class boys in particular seem to have more of a virility ethos than their upper-class peers).
(c) preferred sociability contexts are different across social backgrounds, with working-class youths favoring large groups and collective activities, and upper-class ones favoring dyadic relationships and activities.
However, the empirical material did not cover these topics well enough, so I am not able to reject or confirm these intuitions.

mocking "serious" peers. Indeed, two students that completely disagree as per legitimate school attitudes still agree on the fact that it is a relevant category of relational orientation and moral judgment. How students positioned themselves regarding each framework will be discussed in the body of the text below.

Having laid out the generic rationale behind the analysis of the interviews, I will now proceed to discuss in turn the four considered discursive frameworks. At each step, the reader may refer back to Table 6-1, which is meant to act as a general guide throughout the discussion. However, in each section, I will also discuss specific interview extracts or present simple quantitative treatments.

Table 6-1: Discursive Schemes Related to Socioeconomic Distance in the Interviews

Student	School	School atttitudes	Culture	Ethnicity	Money
Ahmed	Paris 1	dl(+)	-	-	-
Marion	Paris 1	dist (-)	dist(-)	-	dist (-)
Assa	Paris 1	-	dist(-)	-	-
Iris + Kéziah	Paris 1	dl(+)	dist(+)	-	-
Kevin	Paris 1	dl(+)	dist(+)	-	-
Inès	Paris 1	dist(+)	-	dl(+)	-
Laurence	Paris 1	dl(+)	-	-	-
Paul	Paris 1	dl(-)	dist(-)	dist(-)	-
Lou + Jasmine	Paris 1	dist(-)	dist(-)	-	-
Aly	Paris 1	-	-	dist(-)	-
Anna	Paris 1	-	dist(+)	dist(-)	dist(+)
Barnabé	Paris 1	dl(+)	-	-	-
Manon	Paris 1	dl(+)	-	-	-
Théophile	Savoie 1	dist(+)	-	-	-
Laura	Savoie 1	-	-	-	-
Charlotte	Savoie 1	dist(+)	dist(-)	-	-
Quentin	Savoie 1	dist(-)	-	-	-
Brieuc	Savoie 1	dist(-)	dl(+)	-	-
Haruhi	Savoie 1	dist(+)	dist(-)	dl(-)	-
Yann	Savoie 1	-	dist(-)	-	-
Jean	Savoie 1	-	-	-	-
Céline	Savoie 1	-	dist(-)	-	-
Jade	Savoie 2	dist(-)	-	-	-
Maëlle	Savoie 2	-	-	-	-
Rayene	Savoie 2	-	-	-	-
Noémie	Savoie 2	dist(+)	-	-	-
Ninon	Savoie 2	-	-	-	-
Aurianne	Savoie 2	dist(-)	-	-	-
Asan	Savoie 2	dist(-)	dist(-) / dl(-)	-	-
Ali	Savoie 2	-	-	-	-
Dahli	Savoie 2	-	-	dl(+)	-
Mathilde	Savoie 2	dist(-)	-	-	-
Mailys	Savoie 2	dl(-)	-	dist(-)	-
Rémi	Savoie 2	dist(-)	dist(-)	-	-
Julie	Savoie 2	-	-	dl(+)	-
Ilan	Savoie 2	dl(-) / dist(-)	-	-	-
Sylvain	Savoie 2	dl(+)	dl(+)	-	-
Hugo	Savoie 2	dl(-)	-	-	-
Total		25	15	8	2
(+) / (-)		12 / 13	5 / 10	3 / 5	1 / 1
dislike / distance		11 / 15	3 / 13	4 / 4	0 / 2

"dl": expression of social distance with an element of animosity or disliking
"dist": expression of social distance without an element of animosity or disliking
"+": strong subjective significance
"-": weak subjective significance

2. Academic Framework

The most salient perception categories that emerged out of the interviews were, by far, academic ones. As can be seen in Table 6-1, they came up in 24 out of the 38 interviews – being the only framework that was observed in the majority of the interviews. It is also present in relatively even proportions in all three schools, and across all coding categories (i.e. dislike and distance as well as high and low significance). Without doubt, students' result and attitudes in class are one of the main orientation frameworks used to differentiate among peers – as far as formal categories known to all students go, I can only think of gender as being even more important.

2.1. Distance and Conflicts in Relation to School Attitudes

Upon closer inspection, there are three main and related frameworks through which academic distinctions are expressed: students perceive and judge their peers' results, disciplinary attitudes, and work ethos. Unsurprisingly, disagreements about legitimate school attitudes can lead to tensions or conflict among students, as well as a feeling of distance or lack of understanding. This can go both ways: "good" or "quiet" students judging "bad" or "disobedient" ones, or vice versa.

First, high grades are recognized by most students as something desirable. Even though grades are, in the formal discourse of teachers, a private matter, the grading system is public, and students' results are known to their peers, which effectively creates a form of competition among classmates. In that regard, conflicts mostly emerge when one is seen to be bragging or acting pretentiously.

> *"Q: Are there students that you dislike? Or that bother you, or annoy you, something like that?*
>
> *A: (disgusted tone) Aly! [...] I don't like him, I don't know, ever since elementary school he's been pissing me off! He has issues with me [...]. Outside he is not like... when he is outside, he is, well, it's okay. But at school he is annoying.*
>
> *Q: Yeah? Why? What does he do differently at school, or what is different at school?*
>
> *A: (silence) Pff. For example, when I'm at home, I play with him online [on a playstation], and all. At school, I don't know, he's too serious, I don't know, he has issues. [..] He's too much, "yeeeeah, my grades, bla bla bla".*
>
> *Q: He speaks a lot about grades?*
>
> *A: Yes."*
>
> **Kevin, Paris 1**

> *"A: Manel, I don't like how she's bragging.*

Q: She brags when she has good grades?

A: Yeah, a little. She does, hmm, [does a hypocritical, high-pitched voice] "oh, I am so sorry for you" [breathes out strongly with an upset tone]. And me, well, I am quite competitive, so I say "okay, fine, you're gonna see, you're sorry for me, well I am going to be sorry for you".

Marion, Paris 1

Second, even more so than grades, how one behaves during lessons and toward teachers is a public fact in the classroom. Not only is it visible to everyone, but it can also, unlike grades, have direct repercussions for all other students. Indeed, students do not interact with teachers individually, but as a group: the context of the classroom structurally induces forms of cooperation and/or competition among them as per the generic classroom order and relation to the teacher. For example, if many students heckle during the class, or simply talk with their neighbors, this creates a certain collective protection, limiting the ability of the teacher to isolate and sanction each individual talker (by contrast, if a single student is making noise, they will most certainly be spotted and sanctioned); but it may also lead to collective sanctions, which therefore also fall on well-behaved students.

As a result, animosity may arise among students that have different conceptions of the ideal classroom order. Even though this can go both ways, it is more often the case that students complain about peers who misbehave. Complaints about peers being too quiet or too dedicated to the classroom rules do happen, but they are less frequent (which might be due to the fact that students answered to me, an adult, thus potentially under-stating these attitudes).

"A: Sometimes, Jeanne [...], as soon as someone in the classroom laughs a little, she says "no stop, we need to get back to work". [...] Also Chahima. You can't laugh 5 seconds with her. She is always serious, serious, serious."

Kéziah & Iris, Paris 1

"A: I'd like the teachers to have a good opinion of me. For example, you see, when some people are fighting in the hallway, when I see that and the teacher is coming, I say "but stop, we're going to make a bad impression in front of the teacher!". Especially the new ones or those I like.

Q: The new teachers?

A: Those that we got this year, hmm, and which I find nice. I don't want them to think "well, this class is not as nice as I thought."

Manon, Paris 1

"Q: So, we talked about the people you like. Now on the contrary, are there some students that you don't really like, or that bother you?

A: Hmm... well, in the class, it's not that I hate them, it's just that, well, they bother me a little sometimes. For example Hugo, because, well, he really makes a mess in class. Sometimes I'd like to work, that's why we're here. And the teachers, they focus on him too much, to try to make him stop and all, instead of focusing on us and leaving him do his things... well, even though he shouldn't be doing these things. But I mean, it's not that I hate him. He's nice, sometimes it's okay."

Mailys, Savoie 2

Finally, students also clash with regard to their work ethos. Unlike conflicts about classroom order, this does not pertain to students' interests, but rather to moral judgments. A common indigenous term used to refer to good students is that they are "serious". While "seriousness" primarily refers to attitudes rather than results – that is, listening to the teachers, doing one's homework, caring for the lessons etc. – it is striking to see how easily both are associated through that concept. For most students, it seems natural that good results would be the necessary consequence of hard work (which incidentally shows that academic discourse and values are largely internalized, even by "bad" or deviant students). Therefore, excellent students are frequently assumed to be boring, because their high results can only come from tedious work, with the exclusion of leisure.

"A: Last year, sometimes, everyone said he [Colyn] was super strong, he's really lucky, because he has 19 on average. That's why. [...] Sometimes they said "ooh, he's so lucky Colyn. He works a lot that one, he must work so much to get grades like that."

Aly, Paris 1

Conversely, bad students are judged because of their lack of work and effort. In that regard, low-performing but hard-working students are not always perceived as "bad" students *per se*.

"Q: So, in your class, would you know who is a good student and who isn't, for example?

A: Yes. Well, often, the teachers say the grades out loud. Also, it shows, who's a good student and who isn't.

Q: How does it show?

A: Well [waits a moment], a good student, he's not gonna talk in class. Well, he's not gonna mess around. And... he's not necessarily super clever, but... he is... [a moment] I don't really know. In any case, me, I don't know how I did but last year, for example, I saw who was good company and who wasn't. Because when there is... because I saw that some students never raised their hand, or... or they made annoying noises and all. Well, I thought theseare not the students I should follow. [...]

Q: And so on the contrary, good students, you can recognize them, because, you told me...

A: Well, I don't know how to explain. Well, me at least, I can tell. And even, when there is, like... when the teachers ask, when they ask a question, often it's... there are those who manage to...

Q: To answer.

A: ***Because there are some people who have difficulties, but they don't do it on purpose. Others do it on purpose, and they don't learn their lessons*** [I underline]."

Théophile, Savoie 1

2.2. Interaction of School Attitudes and Socioeconomic Background

That students disagree and fight over school attitudes is hardly a surprise and has been documented in several studies (Cousin and Felouzis 2002; Gayet et al. 2003; van Zanten 2012). Nevertheless, as we saw in chapter 4, academic homophily, though it exists, far from implies a complete separation between "good" and "bad" students; which in turn implies that the inducing effect it has on socioeconomic homophily remains limited (cf. chapter 5). If grades and school attitudes really are one of the main perceptive frames through which students sort one another out, why do we not observe levels of segregation on a par with that of gender in friendship networks? A first piece of the explanation might be the difference between attitudes and results: the models of chapters 4 and 5 considered students' grades but, as we just saw, homophilic selection likely pertains to attitudes more than results strictly speaking, and the two are only partly correlated. A second plausible explanation, which I will now further explore, is that there is a negative interaction effect between social and academic distances. In other words: conflicts pertaining to school attitudes may be more severe among same-background students than among different-background ones.

2.2.1. The Impact of Social Distance on Conflicts Related to Grades, Attitudes and Ethos

Looking through the interview extracts where students complain about disturbing peers, it appears that many salient expressions of animosity toward bad students came from working- or middle-class students, some of whom have relatively low results themselves.

"Q: Are there students that you dislike? Or, that bother you a little, that you don't feel right with?

A: (wait a little) Farouk, in our classroom. He makes... a mess (laugh). He bothers people. He says words that are not really... well, not very nice to the teachers. And he heckles the class."

Charlotte, Savoie 1

(father administrative clerk, mother secretary)

Considering the aforementioned three frames through which academic conflict occurs (results, classroom order and school ethos), there are indeed reasons to expect academic distances to be more salient among same-background peers. First, regarding academic results, we saw that conflicts primarily emerged when good students were perceived as bragging. This can lead some of them to adopt a strategy of discretion: by not overly mentioning their results, they can mitigate part of the animosity felt toward them by some of their comrades.

"Q: So this year, they don't say that [brainy] anymore?

A: No.

Q: Because you changed?

A: No, it's... I think the teachers changed a bit as well, but me, now, I am still as clever, but I show it less. I still have good grades, but I show it less.

Q: This means you're gonna participate less in class? Or you're gonna show your grades less to other people?

A: Yeah, that's it, I show my grades less to others, and I say less, like, "oh, I got an 18 [out of 20]!", all that.

Q: You did that last year?

A: Yeah.

Q: And not anymore?

A: Yes. So now I feel much better.

Q: You feel better because the others stopped calling you brainy?

A: Yes."

Haruhi, Savoie 1
(father landscaper, mother legal professional)

Although this could concern upper- and lower-background students equally, the first are more likely to be consistently high-performing, and thus to develop this type of mitigation strategy. Students that perform well occasionally might not need these discretion tactics; and the working-class students that get good grades are more often "occasional" than "consistent" high-performers (simply due to the fact that the very best students tend to be from upper-class backgrounds).

Second, regarding the classroom order, high-performing upper-class students may have less of an interest in regulating the behavior of their peers. Some of them clearly indicate that they enjoy occasional disturbances in class.

"Q: So, what do you think of that? The ones who do it [heckling the class] *all the time, do they bother you, or not? Do you like them, or?*

A2: Well... If it weren't for them, there would be no fun in class.

A1: Yeah, it would be like... just silence."

Jasmine & Lou, Paris 1

(fathers paramedical worker & university teacher, mothers high school teacher & pharmacist)

"Q: There are other students who are often in conflict with the teachers, like him [another student we talked about before]*?*

A: Well, all of his group. Ayoub, Kevin...

Q: So what do they do, in class? Like, they talk, or...

A: They talk, mostly. A lot. And sometimes, they throw paper balls at each other.

Q: And you, does it bother you when they do this?

A: No. I don't like to work when it's completely quiet."

Paul, Paris 1

(father legal professional, mother trip organizer)

These three students have excellent grades and, most importantly, do not struggle to get them: the courses feel relatively easy to them. By contrast, some working-class students express harsher judgments toward the peers that heckle the class or behave badly toward teachers.

"Q: So you still see your friends from primary school.

A: Hmm. But there are some people, they have changed. Because when I was with them, they were nice and all [No.: the French word is "gentil", which can be translated either as "nice" or "kind"]. *For example Wanis, he was really nice, he worked well and all. Now I think he has changed.*

Q: He changed, what do you mean? He's more...?

A: He says more swearwords, he works less. He hangs out with people that are not good. [...] I know that, them, they will, if I am friend with them, they will make my head, they will... you understand?

Q: Hmm, I think so, but I'm not sure... Go on?

A: They're gonna change it.

Q: Change your head?

A: Yes."

"Q: In what sense?

A: Bad sense. [...] I'm not gonna work, I'm gonna do crap.

Q: Okay. Because that has already happened to you, to, not completely but to start having that type of change, a little?

A: Hmm, in the beginning of last year. [...] And then, I begun to understand. If I stay with them, it's not gonna go well."

Ahmed, Paris 1

(father construction worker, mother secretary)

In the second extract, Ahmed proposes an explanation as to why peers' attitudes might matter more to working-class students: he needs to consciously stay away from bad-behaving peers because he is scared of being pulled toward anti-school behaviors himself.

Several studies have pointed to the fact that classroom heckling was subject to strong imitation effects, as many students get drawn into the general atmosphere of the group at a particular moment, despite their individual desires to work. These studies rely on ethnographic observations in segregated working-class schools, where students are culturally very distant from academic demands, thus making schoolwork a counter-intuitive struggle. This tension between "pro-school" rational aspirations and "anti-school" internalized inclinations manifests itself during collective heckling (Testanière 1967; van Zanten 2012). In that regard, it is likely the case that upper-class students, whose family socialization makes academic demands somehow self-evident, are less at risk of getting dragged into these groups. More generally, their school performances are more robust, and less subject to brutally failing because of a lack of work than those of working-class youths. Thus, they can more easily afford occasional drifting from the ideal working conditions.

Finally, moral judgment of the work ethos of peers may also be mitigated by socioeconomic distance. Students tend to associate high grades with tedious work, justifying moral judgment on the basis of grades (good students are boring, bad ones are lazy). However, this meritocratic view on grades seems more frequent among working- or middle-class students. Upper-class students, on the other hand, having strong cultural affinities with the demands of school, often do not need to work hard to get results. The disconnection of work and grades is therefore apparent to them, grounding discourses about "abilities" or "talent" (Bourdieu and Passeron [1970] 1990).

> *"Q: We talked about grades already, but how do you feel yourself about the lessons? Do you like it? Are you interested? Does it sometimes bore you?*
>
> *"A: Well, maths, it's my favorite subject, but the thing is... the thing is, actually, maths... I mean, it's not that I want to brag, but it's kinda simple actually. Because we almost never learn new things. All we do, it's logic, actually. [...] Most of the time, in maths, I'm like this and I look through the window, because I am bored."*

Haruhi, Savoie 1

(father landscaper, mother legal professional)

> *"Q: Now I'd like us to talk a little bit about the courses. You already told me you had a good mention, that it went fine?*

A: Well, actually... I have very good grades, because I've had abilities since I was little [nb: litterally "I have facilities" or "I have ease" ("avoir des facilités" in French)]. [...] There are some teachers who don't really like it, because since I have abilities, I have very good grades even though I don't work much."

Rémi, Savoie 2

(father company owner, mother school teacher)

Given that the association of grades and work appeared as an important element of negative moral judgment among peers that have different school results, it might therefore be the case that high-performing, upper-class students tend to be less judgmental about their low-performing peers, as they interpret low grades in terms of a natural lack of ability, rather than a moral failure to work hard enough (note, however, that this is a guess I am making through extrapolation; there is no interview extract in my corpus where an upper-class youth would explicitly hold something like a patronizing discourse toward low-performing peers).

By contrast, working-class students that are strongly dedicated to schoolwork do condemn the lack of work of their peers (including close friends).

"Q: So your pals, they have good grades like you?

A: No, it depends which friend. Ayoub, him, he's not someone that tries to work, he is mostly into video games. Especially a game that is called Fortnite, he spends his time playing that game. As a result, he is really strong at the game, but he has very bad grades."

Aly, Paris 1

(father job unknown, mother cleaning worker)

"A: Sami, he's nice. The thing is, though, he needs to get to work a bit more."

Ahmed, Paris 1

(father construction worker, mother secretary)

2.2.2. Two Salient Cases

These two last extracts come from Ahmed and Aly, two working-class students from Paris 1 who both actively try to have good grades at school. Their case is interesting to consider because it illustrates the argument developed in this section. Both tend to have conflicting relationships with other peers from their neighborhood and tried to reach out to higher-background peers for support. Importantly, this does not apply only to classroom discipline, but also to the choice of friends in general.

In his interview, Ahmed strongly insists on the fact that Nathan, his best friend (who he met in middle school), is more reliable and trustworthy than his former primary school friends, as Nathan helps him with schoolwork. Moreover, one notable trait about Ahmed is that his mother strongly dedicates herself to his school success. In a second interview one year later (during phase 2), he explained to me that she had enrolled him in an after-school homework association and started monitoring his outings when his results dropped. Thus, parental pressure (and/or support) may explain why Ahmed is so keen on having better results, and in turn why he seems to be intentionally selecting school-oriented friends.

The case of Aly is even more telling. From a working-class or lower-middle-class background[148], he lives close to Younès and Kevin, both of whom are in frequent conflict with the teachers and have low results. In his interview, Aly clearly said that he disliked both of them. I also interviewed Kevin, who named Aly as the person he disliked most in school. Furthermore, Aly's parents strongly control his outings (he is generally not allowed to go outside alone or to meet his friends outside of school) and are highly concerned about his school success. He has two brothers that are both in higher education, and of whom he talks with pride. He also mentioned his own grades a lot during the interview; and another student said to me that Aly's mother would punish him when he got poor results.

The next extract comes from Anna, an upper-class student from the international section who was friend with Aly during the first year of middle school. She mentions the fact that Aly cares a great deal about his grades, and that Younès and Kevin bully him for it (which was also apparent in both Aly and Kevin's interviews). She also explains that Aly became friends with students from the international section, and that they sometimes helped him with school work (most students from the international section have very high results, and they are all from an upper-class background).

> *"Q: And are there some students that are obsessed over their grades?*
>
> *A: Yes.*
>
> *Q: A lot?*
>
> *A: Hmm, actually, me too, I am... I mean, I want to have good grades but Aly, last year, he was really being harassed by Younès and Ayoub. Now I think it got better. But ~~since~~ they are much stronger than him, because him he is very weak, very fragile and sensitive. Therefore, last year, they bullied him a lot but now, I think it's okay. And I know that at home, his parents are really... really strict. So, he was really into his grades and he wanted to do everything perfectly in his life, because he had a lot of problems. So he spoke a lot about his grades.*
>
> *Q: A lot of problems? What do you mean?*

148 Her mother is a cleaning worker, but the declared job of his father is not clear, so there is a bit of uncertainty as per his exact background.

A: Yes, well, for example he had people bullying him at school. And also, he was friends with the internationals. And then, some boys told him he was friends with us just to get good grades and I don't know what, whereas I know it wasn't true, because he was a really good friend of ours. [...]

Q: So you think, when Younès and Ayoub bullied him, it was because he was focused on his grades?

A: Well, yes, because... [...] one day he sent me a message, because I used to help him a little, because he, he was afraid to go see the teachers, because he was afraid that Younès would come for him, since they live in the same building."

Anna, Paris 1

In his own interview, Aly himself strongly insisted on his friendship with students from the international section, which seemed to matter a lot to him.

As for Kevin, the boy that antagonizes Aly, his hostility toward the international section clearly came up during his own interviews (he mentioned them to answer the question about "weird" people, then strongly stated that he did not want to be friends with them). This is also something other students confirmed in their own interviews. However, when asking him about who he disliked, his first, spontaneous answer was for Aly. Kevin's resentment seems stronger toward a peer from the same working-class background that hangs out with upper-class students, than toward these students themselves.

Therefore, Ahmed and Aly appear to be torn between their desire to succeed at school, fueled by their parents, and the influence of the peers from their neighborhood which distract them from schoolwork (Ahmed) or even antagonize them for trying to "break apart", so to speak (Aly). Consequently, both tried to reach out to more school-oriented peers they had met in middle-school (Nathan for Ahmed, who is from the middle-class, and international students for Aly).

Nevertheless, this does not mean that Ahmed and Aly did break out from their former friends entirely, nor that they are unambiguously dedicated to schoolwork. Ahmed does heckle the class himself sometimes, despite his voluntarism in trying to refrain from such attitudes. As for Aly, though his grades remained relatively good, they did drop between the 6^{th} and 3^{rd} grades. Perhaps more importantly, by the end of middle school, he spent almost no time with students from the international section, but had made up with Kevin, who stopped harassing him (we shall come back to this case in chapter 7). Therefore, rather than constant and open conflict between low- and high-performing working-class students, it seems more appropriate to speak of recurring tensions, that sometimes burst out into clear conflict, but can also be mitigated by belonging to close social circles.

2.2.3. Centrality in Friendship Networks: The Interaction of Grades and Socioeconomic Background

In any case, from these examples, we can understand why peers' attitudes toward school may be more decisive for working-class students than for middle- or upper-class ones, whose friends and social environment at large are on average much more favorable to school performance, thus making them less sensitive to the behavior of other students. Can we find quantitative evidence to support this interpretation? As the reader may remember, in chapter 4, I mentioned the fact that higher grades were associated with less received friendship nominations in two schools, Paris 1 and 2, according to dyad independent ERGMs. By interacting this popularity effect of grades with the socioeconomic background of the student, we can see that it is indeed the case that the interaction is negative: the indegree penalty of high-achieving students gets weaker as their socioeconomic background increases. This is shown in Table 6-2 below.

Table 6-2: Interaction of Grades and Socioeconomic Background Reception Effects (ERGM) – Networks of Friends and Very Good Friends, Wave 1

	Paris 1		Paris 2		Savoie 1		Savoie 2	
	Friends	Very Good Friends	Friends	Very Good Friends	Friends	Very Good Friends	Friends	Very Good Friends
reception SAS * reception grades	**0.132*** (0.036)**	**0.098* (0.055)**	**0.052*** (0.012)**	**0.068*** (0.017)**	**0.015 (0.022)**	**-0.056 (0.039)**	**-0.037** (0.017)**	**-0.021 (0.029)**
absdiff SAS	-0.212*** (0.038)	-0.330*** (0.062)	-0.145*** (0.014)	-0.203*** (0.021)	-0.066*** (0.024)	-0.148*** (0.042)	0.028 (0.020)	-0.112*** (0.034)
emission SAS	-0.037 (0.035)	0.174*** (0.056)	0.014 (0.012)	-0.014 (0.017)	-0.044** (0.022)	-0.068* (0.038)	0.166*** (0.017)	0.245*** (0.030)
reception SAS	0.049 (0.037)	0.081 (0.058)	0.025** (0.012)	-0.009 (0.018)	-0.050** (0.022)	-0.055 (0.038)	0.135*** (0.017)	0.147*** (0.030)
absdiff grades	-0.332*** (0.042)	-0.417*** (0.066)	-0.034** (0.014)	-0.043** (0.019)	-0.170*** (0.024)	-0.230*** (0.042)	-0.149*** (0.020)	-0.177*** (0.035)
emission grades	-0.342*** (0.039)	-0.562*** (0.060)	-0.094*** (0.012)	-0.164*** (0.017)	-0.174*** (0.022)	-0.168*** (0.038)	0.006 (0.017)	0.007 (0.030)
reception grades	-0.312*** (0.039)	-0.383*** (0.059)	-0.037*** (0.012)	-0.048*** (0.017)	-0.057*** (0.022)	0.032 (0.038)	0.031* (0.018)	0.055* (0.032)

***: p<0.01; **: p<0.05; *: p<0.1

Note: results are similar overall after wave 1. See Appendix 4F.

Parameters not shown here density, gender homophily, gender emission and gender reception (gender and grades are related so it is important to control for the former).

Once again, these are simple models with a descriptive purpose, so they do not necessarily indicate a causal relationship. Moreover, it holds in two schools only – Paris 1 being the one from which the two cases discussed above come from. In Savoie 1, there is no clear association of grades and indegree centrality[149], and, in Savoie 2, high-achieving students are actually more central on average (more on this below). In both cases, there is no significant interaction with socioeconomic background. Still, this means that, in schools where high grades are associated with low centrality, the relationship only concerns working- and middle-class students. Upper-class ones are equally central no matter their results (on average, of course). This does support the idea of a greater relational cost of performing well at school for working-class students.

149 In the network of friends in wave 1, the association is negative (log-odd of -0.057), but not in the network of very good friends; and this disappears after wave 1 anyway.

Similar insights are to be found in the literature on schooling inequalities among ethnic groups in the US context. Researchers have theorized an “acting white” phenomenon; that is, a statutory penalty of school achievement within Black students’ peer groups, which is seen as a betrayal of the group (Fordham and Ogbu 1986). However, other studies have questioned the relevance of this notion (Cook and Ludwig 1997; Harpalani 2002), so the case for this is not clear cut yet. Regarding socioeconomic background in the French case, some studies did report forms of bullying of high-achieving students in segregated working-class contexts (Cousin and Felouzis 2002; van Zanten 2012), which, to the best of my knowledge, was not observed (or not to the same extent) in dominantly upper-class schools. However, this may be a school-level effect, rather than one that depends on the socioeconomic background of individual students (indeed, my own hypothesis rests on the idea that, *within the same school*, the relational cost of academic achievement is higher for lower-background students). Finally, Passerieux (2010) makes an argument much similar to my own, but with relatively limited empirical support. Altogether then, this idea of differing statutory penalties associated with school achievement depending on students’ background appears promising, but it would demand more robust empirical confirmation.

2.3. A Regulation of Normative Divergence within Peer Groups?

Remember that we are trying to understand why academic tensions among students do not result in high levels of academic segregation within peer groups. As we just saw, this may come from the fact that academic tensions are most pronounced among working-class students. Since academically dissimilar dyads are, more often than not, socioeconomically dissimilar as well, this would tend to reduce the impact of academic distances on average.

A second plausible hypothesis is that academic distances are moderated not just among different-background students, but within peer groups in general. Since there is no normative consensus among students regarding school attitudes, neither pro- nor anti-school norms being hegemonic within peer subcultures (even among same-background students), it may be necessary to compromise in order to maintain some level of cohesion within friendship groups.

2.3.1. The Statutory Ambiguity of High and Low Results

A first element supporting this interpretation is the ambiguity of status norms regarding grades. In the students’ discourse, the relationship between school performance and popularity appears unclear. On the one hand, some respondents clearly stated that students that have low results and clash with teachers are seen as popular. Conversely, “brainies” appear to be mocked or marginalized.

> *“Q: What does ‘brainy’ mean?*
>
> *A: Well, someone that has good grades, who has glasses, and that, hmm, that cares about the lessons?*
>
> *Q: And that’s not good?*
>
> *A: Well, for them [the popular girls of her classroom] no, it’s not, but for me it’s good!*
>
> *Q: For them it’s not good? Why?*
>
> *A: Because for them, you gotta be popular, you should not care about the lessons, or skip classes.”*
>
> **Inès, Paris 1**

> *“A: Some people say she [a girl in her class] is popular because she is very... very silly in fact. Because she does everything to make herself look silly, even though she’s not silly at all.*
>
> *Q: Meaning what, what does she do to look silly?*
>
> *A: Meaning she falls, she falls on purpose. She throws herself against the walls, on purpose [No.: during class]. She does a lot of things that... that tend to piss the teachers off. In class, she takes her phone out. [inaudible] She wants to get expelled from the school to go to another one.*
>
> *Q: But so, when she does it, it makes the others laugh [she approves].”*
>
> **Jade, Savoie 2**

Nevertheless, most students do agree that good grades are something desirable and, more importantly, a sign of either merit (work) or intelligence. In French, the word for “brainy” is *“intello”*, which is built from *“intelligent”* (in English as well, “brainy” comes from “brain”). In Jade’s extract above, the term “silly” (*bête*) is an antonym of intelligent and is used to refer to a student that messes around during class. Furthermore, even though no student explicitly mocked a peer during the interview because of their low results, other studies suggest this can be the case, particularly in primary school (Gayet et al. 2003; Lignier and Pagis 2014). Thus, the association of grades and intelligence seems latent within the perceptive framework that students have.

This can also be seen from the way “brainies” are described by respondents. The term has a very clear negative connotation. However, the interviewed students, while clearly aware of this pejorative sense, struggled to explain why exactly being called “intelligent” was not positive.

"A: There were some people last year who called others "brainy", and they didn't like it. For example, when the teacher had forgotten about the homework, some students reminded her, and the others, they didn't like that.

Q: What does "brainy" mean?

A: It's someone that doesn't speak [in class], who has good grades, for example if the teacher wants to ask something they will always raise their hand.

Q: Is it good to be a brainy, is it positive?

A: Yes, it's positive.

Q: But you said they didn't like being called like this? So it's a bit negative? Is it nice to say this to someone?

A: [hesitates a moment] It's annoying to be called like that every day, "brainy, brainy, brainy" [seems unsure]."

Assa, Paris 1

"Q: [talking about how other students perceive grades] What do they think? Is it better to have good grades or bad grades?

A: Well, it's good to have bad grades, for... because those that have good grades, they call them brainy.

Q: What does it mean brainy?

A: Well, that you are super clever, and all. Well, that's about what they say.

Q: And that is not good?

A: No, it's not good.

Q: It's negative to be a brainy?

A: Yes.

Q: So who is called brainy in your class?

A: Well, me, last year. This year also a little bit, I think. Apart from this, I don't know, I don't have an example.

Q: So when they called you that, you didn't like it, last year?

A: No, I liked it actually.

Q: But you just told me it's negative.

A: Well, they were mocking us, but for me it's positive."

Quentin, Savoie 1

One of the interviewed students, Haruhi, even suggested that some peers meant the word "brainy" as a compliment. Nevertheless, he still disliked being called this, as it made him feel excluded from the rest of the students ("*I don't like being called a model*", which in this case also pertains to the fact that teachers' interventions are typically seen as illegitimate among teenager groups). Note that we already met Haruhi above: he was the student that developed discretion strategies to avoid being called a brainy.

> *"Q: If you had to compare your classroom this year and last year, which one do you prefer?*
>
> *A: Hmm... well, it depends. Last year it was more... a bit more about playing around, and me, I was kind of the "brainy", so to say. This year, I feel more at ease. I guess I am older now. And that's it. So I prefer this year. […] Now I am more in line with the class, so I feel better.*
>
> *Q: And so, brainy, what is it, compared to others?*
>
> *A: Clever, a little. Basically... in fact, it's good to have good grades and all, but then it is a bit annoying when everyone tells you you're a brainy. So now I feel better. In fact, I am still clever, but I am less doing the... doing the brainy... well...*
>
> *Q: What is it, doing the brainy?*
>
> *A: Hmm... How can I say? In fact... [pauses a moment] well, for example, I... [pauses again] well for example, I participate in class, and all, and in fact, for example now, the reference... because last year, all the teachers said "go see Haruhi, he's a model". And I don't really like being called a model actually. [...] Well, they didn't really harass me. It was just a little, I don't know, I guess to joke around, but they said for example "you are the brainy one in the class. We can count on you, you're the model". And I don't like being called a model. Actually, for them, I don't think it's an insult. They say that I'm clever, but actually, me, I don't like being called a model.*
>
> *Q: But you don't think they meant any harm when they said it?*
>
> *A: I really don't think so."*
>
> **Haruhi, Savoie 1**

2.3.2. Grades: An Illegitimate Means of Differentiation?

More than stigmatizing either high or low results, many students express a lack of interest for their classmates' results. While acknowledging that they often ask for their friends' grades, they also insist on the fact that they do not care much about them, nor spend a lot of time discussing it.

> *"Q: So, with your friends, do you sometimes talk about the courses? Do you work together, do the homework, sometimes?*
>
> *A: Hmm, I sometimes do my homework with Clara, but since we're not in the same class it's difficult. Otherwise no, we don't necessarily talk about courses. [...]*
>
> *Q: But like, do you tell each other what competences [equivalent of grades in Savoie 2] you have, or not?*
>
> *A: We do.*
>
> *Q: You speak about it still. A little.*
>
> *A: But it's very rare.*
>
> *Q: And do you know how things are going for your friends?*
>
> *A: Weeeeell... not really. Grades, they really don't matter."*
>
> **Mathilde, Savoie 2**

> *"Q: So sometimes, you ask about each other's grades [with his classmates]. Is it something that people get mocked for sometimes?*
>
> *A: No.*
>
> *Q: Or that they argue about?*
>
> *A: Sometimes, we say to ourselves that we have shitty grades, to laugh."*
>
> **Brieuc, Savoie 1**

> *"Q: With your friends, grades, homework, it's something that you talk about? [...]*
>
> *A: Well, for example, when there has been a test, everyone talks about it, "ah, how much did you get? How much did you get?". Every time there is a test, everyone talks about it. [...]*
>
> *Q: So after a test, everyone says their grade. So you, you have good grades most of the time, I guess. How do others react to it?*
>
> *A: To me having good grades?*
>
> *Q: Yeah, do they congratulate you, or are they jealous?*
>
> *A: They don't care, I think they don't care.*
>
> *Q: Yeah? But then why is everyone asking, if they don't care?*
>
> *A: Well, no, they just want to know to compare their grades, that's it. [...]*
>
> *Q: But it doesn't really call for comments, either positive or negative?*
>
> *A: No."*
>
> **Paul, Paris 1**

There is an apparent contradiction between these discourses that present grades as something private, and the fact that academic hierarchies are extremely apparent in the classroom and are a clear part of students' perceptive frameworks for differentiating among peers. I believe that the above extracts should not be interpreted as descriptive so much as normative: what the respondents express here is essentially that grades *should not* be a legitimate distinction among peers.

The attempts to keep grades a private matter, and the relative moderation with which disapproval toward other people's results – either high or low – is expressed, might therefore be a form of adaptation to the normative heterogeneity among students. Additionally, this could also be due to the fact that students in middle school develop a "critical distance" (Cousin and Felouzis 2002) toward teachers and school, thus becoming more reluctant to integrate academic norms within their frames of judgment –even though they still do, at least to an extent.

2.3.3. Co-Existence of Conflicting Norms within Peer Groups

Finally, a last important point is that the heterogeneity of normative attitudes toward school is not just the result of different groups of students having different attitudes. This heterogeneity exists within the same friendship groups, and sometimes even within an individual student.

Indeed, most respondents expressed mixed feelings as per the most desirable classroom order. In section 2.2., we already saw that some high-performing students (Jasmine, Lou and Paul) said they enjoyed some measure of "fun" in class. In addition, some working-class students dedicated to schoolwork (Aly and Ahmed) were torn between "anti' and "pro" school aspirations. Similarly, in the following extract, Yann is a student with medium grades but who is dedicated to schoolwork, and

confronts peers who are openly in conflict with teachers. Nevertheless, he mentions that he himself can be talkative in class.

> *"Q: How is the mood in your class, actually? Do you find it nice? Is it calm or rather agitated?*
>
> *A: We are noisy. Very noisy, yeah. Very, very noisy, even.*
>
> *Q: That's what the teachers tell you, or it's what you think?*
>
> *A: That's what the teachers tell us, and I also think that a little. It's true we're often talking."*
>
> *[No.: he clearly includes himself in that "we"]*
>
> **Yann, Savoie 1**

In other words, the majority of students agree on (a) giving some legitimacy to school rankings and academic goals and (b) finding classes boring at times, which is why many students wish for occasional disturbance. In fact, hostility or disapproval occurs when a student perceives another as going too far in either of those directions: either being too dedicated to schoolwork or breaking the classroom order too often. This also contributes to explaining the ambiguity of status norms regarding school attitudes.

In that regard, a plausible hypothesis is that the socioeconomically and academically diverse nature of the studied schools specifically contributes to the described process. This diversity might encourage students to moderate the expression of their own school attitudes and preferences, either because of positive contact with dissimilar peers, or simply because the open expression of certain negative judgments would create heightened tensions and retaliations – so to speak a “balance of power” between students with different levels of performance. Consequently, it might be the case that in academically homogeneous schools, statutory norms regarding school attitudes are clearer.

In fact, this is quite possibly the case in Savoie 2, to an extent at least. Remember that it is the only school where high grades (and high socioeconomic backgrounds) are associated with higher indegrees and outdegrees in the friendship networks. Moreover, in the interviews, Savoie 2’s respondents tend to be less ambiguous about the desirable school attitudes than those of Paris 1 and Savoie 1: adhesion to school norms is generally clearer. For instance, in the following extract, Sylvain describes Hugo as a “troublemaker” in class. Of particular interest is the fact that Hugo is marginalized by his classmates for this reason, standing apart from the large group of friends to which Sylvain belongs and which is the central friend group in the classroom.

> *“Q: Ok, well, let’s get started! Who are you friends with in your classroom?*
>
> *A: Well, half of them! No, even more! Almost everyone. Except perhaps a few that are a little… troublemaking. Just a few. Hugo, him, he really is a troublemaker when he is put next to Adrien, Leo or Remi. But… those three, if they are not next to Hugo, they are super nice.*
>
> *Q: Troublemaking, you mean… in class, or with the other students?*
>
> *A: In class! In class he is the worst! Like, in History, the girls were doing a presentation, and Hugo kept on shouting. He was saying “yeah, I don’t care!!”. He just didn’t stop.*
>
> *Q: So that was annoying to you?*
>
> *A: Yeah, really annoying. Especially because it was interesting, the presentation. And he just didn’t stop talking. And in a loud voice,on top of that.”*
>
> **Sylvain, Savoie 2**

Other students shared similar hostility toward Hugo in their interviews, and confirmed his relative isolation. This contrasts with similar stories in Paris 1 and Savoie 1, where students that disrupt the class often act as a group, and do not particularly suffer from a lack of friends. Also notable is the fact that Sylvain spontaneously mentioned Hugo as soon as we started talking about his friends, without me specifically asking him about who he dislikes, about marginalized students or about the mood of the classroom.

Beyond this specific case, it is also clear that classrooms are much quieter and more obedient toward teachers in Savoie 2 compared to the other schools[150]. Furthermore, as was

150 Including when they are asked to fill sociological questionnaires!

mentioned in chapter 3 (section 2.3.4.), the school administration tends to operate an academic selection of students (more or less officially), as low-achieving or disruptive students are at risk of being refused acceptance into the school from one year to the next (which is what eventually happened to Hugo). Altogether then, academic attitudes are more homogeneous in this school than in the others, with a clearer adhesion to school norms from the majority of the student body. This likely explains the positive association of grades and network centrality that is specific to Savoie 2.

2.4. Institutionalized Academic Differences

To conclude this discussion of academic frameworks of perception, it is important to mention the case of institutionalized academic differences. More specifically, I refer to “special” tracks, with options or curricula that are different to those offered to the majority of the student population. While grades are already an institutional distinction that schools operate among students, special tracks go one step further, for two reasons. First, they strongly constrain the structure of meeting opportunities among students, as individuals that are in the same classroom spend most of their time together. Second, unlike grades, which are essentially a continuum of positions, these classrooms operate a clear symbolic distinction between two discreet and unequal groups: the majority of the student body, and the “special” students.

Paris 1 and Savoie 1 offer two interesting cases to compare in that regard. In Paris 1, the atypical option is the international section, a French-English bilingual program that prepares students for the international baccalaureate. Almost all students in this program come from the upper-class and have high results. Depending on the year, there are around 15 students in the program, which makes up for about half a classroom; the rest of the classroom is made of “regular” students, but a section of the course is attended by the international program students separately (international students have a few extra-hours of language classes). On the contrary, in Savoie 1, the atypical classroom is a SEGPA[151]. This is a program that was historically designed for students with mild intellectual disabilities, before being extended to students with strong academic difficulties; it now effectively acts as an anticipated vocational track (Palheta 2015). SEGPA students have a smaller classroom of their own (~10 students, depending on the year), and attend most of their classes in a separate building.

151 Adapted Section for General and Professional Teaching *(Section d’Enseignement Général et Professionnel Adapté*).

Therefore, in both cases, there is a strong institutional label as per the academic capacities of the students of the atypical track: the international section is clearly presented as an elite track, whereas the SEGPA is explicitly designed for students that struggle at school. Unsurprisingly, these labels are then imported into students' peer cultures, creating forms of stigma or distinction among them. In Paris 1, "the internationals" are clearly identified as a distinct group from the other students. They are commonly seen as being privileged by the school and teachers and are accused of staying with one another and not opening up to others.

> *"Q: Because it's, how can I say, sort of a tight group, the internationals?*
>
> *A (together, immediately): Yeah!*
>
> *A1: Yeah, the internationals, they stay together a lot.*
>
> *A2: I mean, I feel that the internationals, they are afraid to talk to others. For example, in our classroom, when there are internationals and non-internationals, in that case the internationals manage to talk to non-internationals, because they are in the same classroom. But when it's people from other classrooms, they don't really open up, they're more together. [...]*
>
> *A1: Yeah, I feel the same. For example, well it's just an example, but Ayoub in physics, he said that he would never date an international.*
>
> *Q: Ayoub said that?*
>
> *A1: Yes! Like, he said, "with an international", you see! Like it's a group, or a political party. I don't know. It's kinda strange.*
>
> *A2: But then, like, often, you guys from the international section, you don't know that the non-internationals, they think that it's a bit unfair, everything that you have, like the school trips and all. [...]*
>
> *A1: The internationals, we are kinda privileged. We have more trips and more activities than the non-internationals. I don't understand why."*
>
> **Jasmine (non-international) & Lou (international), Paris 1**

In other interviews, some non-international students mentioned that they disliked the internationals, accusing them for example of intentionally speaking English between themselves so that others would not understand. This is a well-identified problem that the school administration is aware of, and they have attempted different strategies to try and pacify relations among the groups[152].

152 In particular, a tutorship system has been put in place, where international students could be paired with non-international ones in order to help them with their homework (this was done on a voluntary basis). I heard some positive feedback from a few students and teachers following the installation of that system, but I did not have the opportunity to systematically investigate it. Interestingly though, this tutorship system, which was meant to improve inter-group relations by turning international students' school performances into a resource for non-international ones, nevertheless reaffirmed for internationals their status as an academic elite (since they were the tutors, and non-internationals the ones receiving help).

I will not spend much time on this issue, as it is specific to Paris 1 – we shall come back to it in more detail in chapter 7. In any case, it is clear that the awareness that students have of the distinction between internationals and non-internationals, and their willingness to mobilize it to identify or judge their peers, goes far beyond what was discussed before about other academic frameworks (grades, attitudes and work ethos). This is also reflected in friendship networks, where international students form a dense, clearly identifiable cluster.

As for the SEGPA classroom in Savoie 1, it has not been mentioned in the interviews I conducted. Paris 1 being a small school (three classrooms per grade level), the international section represents a fair portion of the student body; by contrast, Savoie 1 is large (seven classrooms per grade level), meaning that the significance of the SEGPA section relative to the student population is much smaller. Moreover, the SEGPA students attend most of their classes in a different building, meaning that they have relatively little contact with non-SEGPA peers. As a consequence, they form a dense cluster that stands almost completely apart from the rest of the school in friendship networks.

The impact of special tracks on friendship formation was confirmed in the SAOM models of chapter 5 (cf. Table 5-3). Moreover, similar patterns are reported in the literature. Palheta (2012) interviewed former SEGPA students that declared having been frequently mocked by other students, while van Zanten (2015) mention that headmasters in schools with CHAM programs (music classes with special timetables, which is an elitist track) regret a lack of mixing among CHAM and non-CHAM students. Altogether then, it seems clear that atypical programs, formalized through official labels and embodied into separate tracks and classrooms, have a strong impact on students' relationships and perceptive frameworks. While this is a relatively expected result, it does point to the strong power of the school institution in shaping youths' sociability through organizational and symbolic distinctions.

3. Cultural Framework

After academic attitudes, the most frequent perceptive categories that were mobilized by students to express socioeconomic distances were cultural ones. Here I use "culture" in the restrictive sense of the term: it pertains to students' artistic, clothing and leisure practices. We can see in Table 6-1 (I show it again in the next page for ease of reading) that these types of categories came up in 15 out of the 38 interviews. However, occurrences are not evenly split across schools: they are rare in Savoie 2 (only three, two of which have low subjective significance) but frequent in Paris 1 (more than half of the interviews, with about half of the mentions having a high subjective significance). Savoie 1 seems to stand in between these extremes.

3.1 Working- and Upper-Class Styles

Prominent trends in the sociology of culture regard cultural tastes and practices as an expression of class positioning, and a means to distinguish oneself from other social groups (Bourdieu [1979] 2016; Coulangeon 2010). Coherent with this view, students in my sample can have different cultural "styles", which can be related to certain socioeconomic positions.

The easiest opposition to identify is that of rap and rock music, which largely overlaps with a French/American distinction. Working-class students overwhelmingly listen to French rap, whereas American rock, but also pop-rock, electro or "black" music (i.e. soul, blues, jazz and the likes) are quite distinctive of upper-class tastes. This opposition is not only about music, as it defines overall "styles" that are also apparent in clothing. Rap-oriented tastes generally go along with a so-called "street" clothing style: tracksuits, sneakers, and sport brands in general[153]. On the contrary, students that listen to rock music have a preference for jeans (for girls in particular, high-rise jeans that leave the ankles showing) and city shoes.

Importantly, these styles are not defined by explicit and systematic rules (e.g. all jeans are not indicative of a preference for rock). Rather, there are a number of small hints that, taken together, define overall styles, but which can nevertheless be combined in different ways and with a relative fluidity. For example, a student could wear shoes that are usually indicative of a "street" style together with a jacket that has a more "urban" or "rock" feel. Girls in particular seem to have a broader diversity of clothing styles, and some of them alternate between or mix different styles.

153 This "rap" style is also linked, for boys at least, to football (both of which are core elements of a broader "suburb" style, very prevalent among working-class youths with an immigrant background).

Table 6-1: Discursive Schemes Related to Socioeconomic Distance in the Interviews

Student	School	School atttitudes	Culture	Ethnicity	Money
Ahmed	Paris 1	dl(+)	-	-	-
Marion	Paris 1	dist (-)	dist(-)	-	dist (-)
Assa	Paris 1	-	dist(-)	-	-
Iris + Kéziah	Paris 1	dl(+)	dist(+)	-	-
Kevin	Paris 1	dl(+)	dist(+)	-	-
Inès	Paris 1	dist(+)	-	dl(+)	-
Laurence	Paris 1	dl(+)	-	-	-
Paul	Paris 1	dl(-)	dist(-)	dist(-)	-
Lou + Jasmine	Paris 1	dist(-)	dist(-)	-	-
Aly	Paris 1	-	-	dist(-)	-
Anna	Paris 1	-	dist(+)	dist(-)	dist(+)
Barnabé	Paris 1	dl(+)	-	-	-
Manon	Paris 1	dl(+)	-	-	-
Théophile	Savoie 1	dist(+)	-	-	-
Laura	Savoie 1	-	-	-	-
Charlotte	Savoie 1	dist(+)	dist(-)	-	-
Quentin	Savoie 1	dist(-)	-	-	-
Brieuc	Savoie 1	dist(-)	dl(+)	-	-
Haruhi	Savoie 1	dist(+)	dist(-)	dl(-)	-
Yann	Savoie 1	-	dist(-)	-	-
Jean	Savoie 1	-	-	-	-
Céline	Savoie 1	-	dist(-)	-	-
Jade	Savoie 2	dist(-)	-	-	-
Maëlle	Savoie 2	-	-	-	-
Rayene	Savoie 2	-	-	-	-
Noémie	Savoie 2	dist(+)	-	-	-
Ninon	Savoie 2	-	-	-	-
Aurianne	Savoie 2	dist(-)	-	-	-
Asan	Savoie 2	dist(-)	dist(-) / dl(-)	-	-
Ali	Savoie 2	-	-	-	-
Dahli	Savoie 2	-	-	dl(+)	-
Mathilde	Savoie 2	dist(-)	-	-	-
Mailys	Savoie 2	dl(-)	-	dist(-)	-
Rémi	Savoie 2	dist(-)	dist(-)	-	-
Julie	Savoie 2	-	-	dl(+)	-
Ilan	Savoie 2	dl(-) / dist(-)	-	-	-
Sylvain	Savoie 2	dl(+)	dl(+)	-	-
Hugo	Savoie 2	dl(-)	-	-	-
Total		25	15	8	2
(+) / (-)		12 / 13	5 / 10	3 / 5	1 / 1
dislike / distance		11 / 15	3 / 13	4 / 4	0 / 2

"dl": expression of social distance with an element of animosity or disliking
"dist": expression of social distance without an element of animosity or disliking
"+": strong subjective significance
"-": weak subjective significance

These styles are perceived by students. This is particularly the case in Paris 1, where the French/American distinction in musical tastes is made even more salient by the fact that most upper-class students are part of the international section, therefore fluent in English.

> *"Q: So, are their certain students, that you may like or may not, but that feel... not sure how to say, that feel a bit different. Like they'renot really on the same planet.*
>
> *A1: But they're not our friends?*
>
> *Q: Friends or not friends, both. But just people that make you think "him, I don't understand how he works".*
>
> *A2: Milana. She has changed. Last year she wasn't like this.*
>
> *A1: She changed.*
>
> *A2: She became a bit... weird. Like Jasmine. [...]*
>
> *Q: So how did they change, Milana and Jasmine?*
>
> *A2: No, Jasmine she was always... weird let's say.*
>
> *Q: How so? [...]*
>
> *A2: I don't know, the way she is, how she thinks, how she dresses. She's really into, like...*
>
> *A1: The American style, all that.*
>
> *A2: Yeah, she's more into American things […] She comes with the sweaters, you know, that stop before the belly button. And large pants, ridden up up 'till there* [note: high-rise jeans are an ongoing fashion trend among French youths, in particular middle- or upper-class ones]. *[...] And big shoes.*
>
> *Q: Whereas you, you wear...*
>
> *A2: No, we are normal. I mean, jeans, sneakers, a sweater, a t-shirt and a down jacket* [note: "doudoune" in French, generally evokes a rap/street clothing style]. *[...]*
>
> *Q: So, Jasmine was always like this? [...]*
>
> *A2: Yeah. And then Milana, at the beginning of the year, she was... well, jeans, sneakers.*
>
> *A1: Because Jasmine was her best friend, she contaminated her. [...]*
>
> *Q: And so apart from clothes, is it also different behavior?*
>
> *A2: Yeah. Actually, we don't like the same things. For example [...], we are going to talk about a French rapper, and she is going to talk about an American rapper, that we won't know, and she won't know."*
>
> **Keziah & Iris, Paris 1**

Beyond music and clothing, haircuts are also a way of manifesting one's style to others, as shown in Yann's extract below.

> *"Q: Do you follow fashion trends, for clothing?*
>
> *A: No.*
>
> *Q: And in the school, do you think it is something important?*
>
> *A: Yeah, I think so. Me, I can't stand strands... I can't stand them.*
>
> *Q: Strands? Of hair, you mean?*

A: Yes. I don’t like that.

Q: Because, what is it? The fashionable hair style is to have a strand?

A: Yeah. Well, almost... a lot of people have it. Me, hair gel already, I can’t stand it [...]

Q: Where do you think it comes from, that trend?

A: Pff. Football, probably.”

[note: the haircut Yann is referring to is one that is popular among the boys of his school, where one has shaved hair on the side of the head, and a longer strand of hair on top of the head that sometimes hang down on the side, or that can be held in place with gel. This type of haircut is widespread among football players, but also certain rap artists.]

Yann, Savoie 1

The rock/rap distinction in musical tastes is also apparent from the questionnaires. In wave 1, 3 and 6, students were asked to name three artists that they liked. Answers were coded in about 30 different categories of musical styles (see Appendix 3F). On average, the students that declared one or more artists belonging to the category “French gangsta rap” (n=429) had a Socio-Academic Score (SAS) of 0.11. This was 0.16 for the category “French musical rap” (n=274)[154]. By contrast, students that named an artist from the “rock and electro-pop” category (n = 113) have an average SAS of 0.78, and 0.84 for the category “classical rock” (n=55)[155]. The category with the highest average SAS is classical music, at 1.40, but it concerns a very small number of students (n=10).

As a side note, it is interesting to see that the upper-class connotation of rock music seems more pronounced than the working-class connotation of rap music. By construction, SA scores are a standardized measure within a national French sample (not within my own sample). This means that the students that listen to rock music are, on average, close to one

154 “Gangsta” rap refers to a rap style where the singer presents himself as being a gangster or outlaw, with a virility ethos of power and success, themes revolving around money, crime, weapons, drugs or women, and lyrics that are generally characterized by short punchlines with plays on words, without necessarily a narrative continuity (i.e. there is not a particular “story” that is told throughout the song). Examples in French are Kaaris, Ninho or Koba laD.
What I call “musical rap” is a style that is not necessarily completely separated from the previous one (some artists can alternate between both styles), but that generally tackles lighter themes (e.g. love or friendship), has more narrative continuity (one song is a given story) and puts more emphasis on the music, with dancing or cheerful tunes (it is also more often a style in which the artist sings, rather than only rapping). Examples are Black M or Aya Nakamura.
Both are generally working-class styles (with a strong ethnic dimension as well), but they are popular enough among youths to also be listened to by middle- or even upper-class students. “Gangsta” rap is more prevalent among boys and “musical rap” among girls.

155 “Rock and electro-pop” comprises different styles that were grouped together because each one has a low representation in the sample, but which share the fact that they come from recent, English-speaking artists, with an element of pop music in their style. They can then be oriented more toward rock (e.g. Imagine Dragons) or electro (e.g. Daft Punk, which despite being French typically use English for their lyrics).
“Classical rock” is a category for famous rock groups from the 80s and earlier (e.g. Queen, the Rolling Stones, the Beatles, AC/DC etc.).

standard deviation above the national average (0.78 and 0.84, as I just mentioned). Students that listen to rap music, on the contrary, are close to this average (0.12 and 0.16)[156]. We can also see that the total number of students that listen to the style is simply higher for rap (429 and 274, vs 113 and 55 for the categories indicative of rock and electro-pop). Therefore, upper-class youths seem to have more uncommon and distinctive tastes, whereas working-class tastes are more frequently widespread.

In any case, these differences in musical and clothing styles can elicit forms of moral judgment among youths. In general, those are not antagonistic strictly speaking: as can be seen in Table 6-1, in the large majority of cases, cultural frameworks were mobilized during the interviews to express distance, not disliking (3 instances of disliking, 13 instances of distance). Still, there is a clear element of disapproval, if not disgust, in some of these expressed distances.

The judgment frameworks used to denigrate upper-class music or styles are not entirely clear from my interviews alone, as there were few mentions of them altogether (also, remember that there is a socioeconomic and academic bias in the willingness that students have to open up to the interviewer, and in their ability to verbalize their feelings or opinions). Nevertheless, the following extracts suggest two possible lines of criticism: upper-class music might be deemed boring and depressing; and upper-class clothes uncomfortable, uptight and too formal.

> *"Q: Do you have friends that do not have the same musical tastes as you?*
>
> *A: Yeah. Milana last year, she listened to the same music as me, but there were also some things that were... beeerk.*
>
> *Q: Beerk?*
>
> *A: I don't know, it was sad. In their voice, it's like they are depressed.*
>
> *Q: It was in French or in English?*
>
> *A: English.*
>
> *Q: And you, you don't like it when it's sad like this? Like, you prefer it more...*
>
> *A: Yeah, me I like it when it's fact-paced, when it's fun."*
>
> **Assa, Paris 1**

> *Q: How do you dress? [...]*
>
> *A: Well, casual. A tracksuit... [...]*

156 In my sample, considering all schools at once, the mean of SA scores is 0.25. In that regard, the scores of 0.11 and 0.16 for students that listen to rap music are inferior to the sample mean (the difference in mean is statistically significant at the 99% threshold under a student's t-test). Still, this includes students with scores clearly above 0, thus from a middle-class rather than working-class background.

Q: And them [students from the international section], how are they dressed?

A: With jeans. [...]

Q: And you, you wear jeans sometimes?

A: I don't like it! I don't like. [makes a hand movement to express a feeling of tightness and discomfort]

Q: Why, because it sticks to you?

A: Yeah, it sticks, I don't know, it's weird. Anyway, I don't like it. But I wear jeans, for example, if I go to a marriage, stuff like that. There I wear jeans. But otherwise, no.

Kevin, Paris 1

The criticisms addressed to rap music – "gangsta" rap in particular – are much clearer and have two main components. First, it is seen as "vulgar", a word that is often used by students. Following this line of thought, rap artists are believed to promote violence, sexism or disrespect toward others. Second, it is seen as simplistic, if not stupid. Both the music and lyrics are presented as something easy to perform, that does not hold much artistic deepness.

"Q: So, what music do you listen to?

A: I listen to a bit of everything. New things, mostly. I like Kenji Girac or Marwa Loud. Well, what I hate more than everything, is JUL. It's really something I can't listen to, not even a single song.

Q: Yeah, why? What do you dislike so much about it?

A: Well, we can't even hear his voice, it's just auto-tune [note: a sound effect to warp one's voice, that some rappers use a lot]. I can do that as well! I do whatever, then I put 300 kilos of auto-tune, and that's it. I've done some JUL. And his lyrics, I mean... "yo yo yo yo yo", all the song. Great song!"

[note: she doesn't literally say "*yo*" in French, but "*wesh alors*", which is a way to say "*hi*" in slang, and has a strong rap / street culture connotation]

Marion, Paris 1

"Q: What do you like [to listen to]?

A: I really like rock music. Electro a little as well, but really, I am more into rock. And then, it hasn't been for so long, my brothers started listening to a little bit of rap, and me too. I kinda like it, actually. I really, really like the music. And rap, honestly, I like... I don't like things like Damso [a French rapper], *it's a rap that is a little... I know that deep down, there is a message, but I find it a bit violent, the way he expresses that message. And it's not what I like at all.*

Q: What kind of rap do you listen to, then?

A: I like American rap. Old things like Tupac or Fifty Cents, stuff like that. I also like French rap. So, I like BigFlo & Oli, I like Nekfeu. I like Alpha one. But I like small, like, rappers like that, who... it is sometimes violent, but well, it's already... It's not just about what a lot of rappers talk about, like, what's his name? Koba la D, who talks about drugs all the time. And honestly, at some point, pff, okay, we get it. [...] Nekfeu, he talks about a *lot of things. For example, his new album, I really liked it because he talks about himself a lot. Even if we don't always understand, he plays with words a lot. And that's why you need to listen to the same song multiple times, because you don't always understand everything from the start. Whereas someone that screams insults, insults about women, stuff like that, well we understand the message right away. And... well honestly, I don't find that interesting."*

Rémi, Savoie 2

The case of Rémi is quite interesting, in that he listens to rap music as well, even though he has an upper-class background. However, he likes specific artists: both BigFlo & Oli and Nekfeu have a style that is different from that of Koba laD, with an overall more "intellectual" touch[157]. As for Tupac or Fifty Cents, these are American rap artists that are relatively old and therefore considered as classics. Moreover, these are not exclusive tastes, as Rémi primarily listens to rock music (this very much corresponds to the model of the upper-class having "omnivorous" tastes; Coulangeon 2010). Even though rap is, by far, the most popular style among all students, social distinction thus remains present in the details of the artist choice and in the modes of listening[158].

3.2 Cultural Tastes: Conformism or Distinction?

Nevertheless, these cultural distinctions are not systematic, nor are they omnipresent in students' daily lives. In fact, compared to academic frameworks, the prevalence of cultural frameworks of judgment appears much more limited. I already mentioned the fact that, in the interviews, most expressions of distance using cultural frameworks were not antagonistic and had a low subjective significance (cf. Table 6-1: 3 "dl" vs 13 "dist" codes, and 5 "+" vs 10 "-"). Moreover, the distinction between cultural styles is only meaningful among students that have an apparent style. Indeed, many students simply do not care much about clothing nor music.

157 Neither of them really goes along with the "gangsta" style, and they both reached out, at some point in their careers, to other, more legitimate musical styles (e.g. Nekfeu has been rapping along with a jazz artist, while BigFlo & Oli have an instrumental formation, sometimes playing the piano or trumpet). Their flow is also clearer, more articulated, and with a natural voice (which is an important opposition with many rap artists that use auto-tune, including Koba laD).

158 This is something that many social science studies fail to account for, using a generic "rap" category when investigating musical tastes. This has led some of them to conclude that there was not much social distinction in musical tastes anymore, as some musical styles – rap in particular – where now listened to by the vast majority of youths (Glevarec and Pinet 2009). I would argue that this is a misconception born from poorly delimited categories: socioeconomic distinction does operate among the (many) different types of rap music. In my own questionnaires, students that named one artist belonging to the "gangsta" style had a SA score of 0.11 on average (n=429); those that named an American rap artist had an average score of 0.27 (n=108); the category that I called "anti-gangsta" artists (BigFlo & Oli, Orelsan, Soprano), a score of 0.50 (n=131); and rap artists that are closer to a pop style (in their instrumental tracks and samples) a score of 0.51 (n=54).

> *"Q: So, there are many people who listen to French rap?* [the respondents just explained that they do not like French rap, because it is vulgar]
>
> *A1: Yes. But then, people don't really judge what you listen to.*
>
> *Q: Is it something you talk about with one another?*
>
> *A2: Not so much.*
>
> *A1: Well yes, a little bit.*
>
> *Q: Just a bit?*
>
> *A2: Sometimes, well, we sing songs.*
>
> *A1: Yeah.*
>
> *A2: But apart from that, it's not, hmm... [makes a fake voice] "Yeah, I'll put that in my playlist tonight!". It's not like that.*
>
> *A1: Yeah."*
>
> **Jasmine & Lou, Paris 1**

> *"Q:* [after talking about how she dresses] *And do you see some students that have a certain style? Or that dress in a certain way, that you notice?*
>
> *A: No.*
>
> *Q: Like in an American TV show, with the gothics, the rockers...*
>
> *A: (both laugh) No!"*
>
> **Céline, Savoie 1**

Music tastes are discussed among friends–and songs are sung together–but not across the broader peer group. Furthermore, the discovery of new music seems to happen as much through family members or generic media (in particular, algorithmic suggestions from internet platforms; TV and radio to a lesser extent) than by horizontal socialization among peers[159].

As for clothing, only a minority of students have a clear, distinctive "style"; most students dress in a rather discreet and conformist way. A quick look at the playground or in the classes does not reveal strongly delimited groups. On the contrary, ordinary jeans, sweatshirts and sneakers seem to represent "normal" clothing, common to many students[160]. While upon a closer look social differences do appear through dress codes (e.g. the brand of sneakers a student might wear), they are not immediately apparent, and do not concern all students. Consequently, whereas some of them are able to "decipher" the dress codes of their

159 Out of the 38 interviews, 36 students said that they listened to music, and I asked to 25 of them how they discovered the artists they had named. 15 mentioned a sibling or parent, 10 an internet platform, 4 the TV or radio and 12 a friend or classmate.

160 There is probably more diversity in girls' than in boys' clothes, but this is something I did not investigate thoroughly.

classmates (as we saw in the above extracts), others seem not to rely on this means of orientation to differentiate among peers.

> *"Q: At school, is it something important, who wears what?*
>
> *A: A little bit, I mean, not always. Some care about it, some don't. [...] You have those that are always on-trend. You have those that don't really care that wear... like, pants, they just take them... they take it, they don't care what they wear; I mean, they choose at random. And you have those that are a bit in between.*
>
> *Q: You mean that pay attention, but sometimes?*
>
> *A: Who pay attention sometimes, and sometimes they don't care at all."*
>
> **Jean, Savoie 1**

> *"Q: Do you care about the current trends? For clothes, I mean.*
>
> *A: No.*
>
> *Q: And it's something that students talk about with one another, or...*
>
> *A: No.*
>
> *Q: You don't have any discussions about brands? Like "I want Nike shoes", or something?*
>
> *A: Oh yes, sometimes, they do speak of... between Nike and Adidas, which one is the best, things like that, I don't know. Some say it's Nike, others Adidas.*
>
> *Q: And you? You're neutral?*
>
> *A: Me, I don't mind the brand. What matters is to be dressed (laugh)."*
>
> **Théophile, Paris 1**

3.3 An Increased Relevance of Cultural Categories over Time?

Regarding cultural distinction, a student's age probably matters a lot. The literature suggests that children's tastes are mostly conformist, with fashions and trends being common to an age group, whereas teenagers, and in particular high school students, resort to clothing or music as a means of distinction and to assert an original identity (Dubet and Martuccelli 1996; Galland 2011; Pasquier 2005). At the time of the interviews, students were 12 or 13 (second year of middle school), thus probably somewhere in the course of this transition. The chances are that cultural tastes will become more and more important over the years – during middle school perhaps, and with the shift to high school most certainly. One would therefore expect cultural tastes to grow stronger as an inducer of socioeconomic homophily over time. This is what Kéziah & Iris's extract suggested, when they accused Milana of "having changed" compared to the first year of middle school (cf. section 3.1. above).

Nevertheless, there might be a contradictory trend as well. As I mentioned before, the interviewed students have tastes that are still largely inherited from their family (music discovered through siblings or parents; moreover, several students told me that they still bought clothes together with their parents). However, another transition that takes place during adolescence is that youths' normative reference shifts being their parents to their same-age peers, for cultural tastes and leisure in particular (Octobre et al. 2010). Therefore, the relative importance of tastes imported from the family sphere might decrease over time; and students from all social backgrounds might get socialized into common tastes through their interactions at school (of course, this can only happen in diverse schools, and if socioeconomic homophily in friendship networks is not too high).

Ideally, this is something that should be investigated through longitudinal interviews; those, however, have not been collected, for reasons already exposed (cf. chapter 3, section 1.1.2.). As for quantitative measures, they offer a rather mixed picture. Descriptively, the homophily of cultural status (i.e. highbrow vs lowbrow practices) does seem to increase across waves, except in Savoie 1. This is shown in Table 6-3, which presents dyad-independent ERGMs that only have parameters associated to the scores of cultural status (homophily, emission and reception, plus the density parameter; only the coefficients for homophily are shown here).

Table 6-3: Log-odd Coefficients for Cultural Status Homophily (ERGM) – 'Very Good Friend' Nominations

	w1	w2	w3	w4	w6
Paris 1	-0.674*** (0.069)	-0.724*** (0.063)	-0.803*** (0.061)	-0.874*** (0.064)	-0.840*** (0.074)
Paris 2	-0.265** (0.021)	-0.221*** (0.022)	-0.265*** (0.021)	-0.339*** (0.023)	-0.329*** (0.028)
Savoie 1	-0.274*** (0.039)	-0.341*** (0.036)	-0.189*** (0.033)	-0.227*** (0.033)	-0.243*** (0.038)
Savoie 2	-0.140*** (0.034)	-0.065* (0.035)	-0.120*** (0.036)	-0.173*** (0.038)	-0.243*** (0.046)

***: p<0.01; **: p<0.05; *: p<0.1

Note: the coefficients correspond to the ERGM parameter of the absolute difference between the nodes' scores of cultural status. Other effects included in the model: density, emission and reception. See Appendix 5F for the complete models. Scores of cultural status were only computed based on students' answers to the questionnaires of waves 1 and 3. For the other waves, the score obtained at the closest previous wave is used (i.e. wave 1's scores for wave 2's networks, and wave 3's scores for wave 4 and 6's networks). This should tend to slightly reduce the size of estimates in the later waves, particularly wave 6, as students may have changed some of their cultural practices in the meantime (i.e. the measure becomes more noisy).

However, multivariate SAOMs similar to those of chapter 5 do not indicate any clear time trend regarding selection processes based on cultural factors, once other relevant relational processes are controlled for[161]. Therefore, it is not clear whether students are indeed more sensitive to the cultural styles of their peers over middle school years. In that regard, the shift between middle school and high school may be more decisive than a simple linear effect of time.

4. Ethnic Framework

There is a pronounced association between students' ethnic and social backgrounds. On average, students with French or European origins have a higher socioeconomic background than those with non-European origins (Table 6-4). However, it is clear that identification (from oneself) and labeling (from peers) are much more frequent for ethnic than for social background. In fact, I never heard a student, neither in interviews nor during classes or in the playground, explicitly qualify someone in socioeconomic terms. Slang terms meant

161 I added time dummies for the parameters pertaining to cultural homophily, i.e. the parameters specific to Model 2 in the SAOMs from Chapter 5 (cf. Table 5-3). Most of these time dummies were not statistically significant (models not shown).

to refer to social classes simply seem absent from students' vocabulary[162]. By contrast, several slang and non-slang terms are used to refer to someone's ethnic background. These can be framed in racial as well as migratory terms: "*Renoi*" (slang for Black), "*rebeu*" (Arabic), "*babtou*" (White), but also national origins ("Algerian", "Senegalese", "French") are commonly used by students (see the methodological note below).

Table 6-4: Linear Regression of SA scores on Migratory Background (all waves)

	Paris 1	**Paris 2**	**Savoie 1**	**Savoie 2**
Intercept (= mean of natives)	0.947*** (0.162)	0.514*** (0.137)	0.322*** (0.083)	0.392*** (0.072)
non-French European	0.577* (0.314)	-0.098 (0.267)	-0.329* (0.175)	-0.393** (0.163)
Middle-East and North Africa	-0.951*** (0.231)	-0.297* (0.156)	-0.684*** (0.209)	-0.663*** (0.219)
Subsaharian Africa	-1.293*** (0.273)	-0.335* (0.198)	-0.889* (0.477)	-0.463 (0.468)
Others	-0.829* (0.493)	-0.813*** (0.237)	0.793* (0.477)	0.059 (0.384)
Nb of Observations	99	293	196	237
Adjusted R2	0.303	0.029	0.071	0.040

***: p<0.01; **: p<0.05; *: p<0.1

Note: "others" include all the groups whose sample size was too small to have a separate category. These include students with an East-Asian, Indian, South-American and Turkish Background.

Methodological Note: Ethnic, Racial and Migratory Classifications Among Students

The delimitation of the ethnic categories used by students is not straightforward at all. In fact, they are constantly alternating between different levels of ethnic grouping. They can resort to what are effectively racial classifications: "White" (*blanc* or *babtou*), "Black" (*noir* or *renoi*), "Arabic" (*arabe* or *rebeu*), "Asian" (*asiatique, chinois* or *jaune*), etc. They can also aggregate those into higher order groups, especially by opposing native students to all others – which can, again, be framed in different terms: "Whites" vs "non-Whites", as well as "French ones" vs "immigrants" (meaning that depending on the classification, students with a European immigrant background – e.g. Roma or Italian – could fall in either category).

162 In French, one could for example use "*bourge*" (bourgeois), "*fils à* papa" (daddy's kid), "*clochard*" (homeless) or "*prolo*" (prole). None of these were mentioned in the interviews or during my fieldwork in general (granted, these are terms that students would probably avoid in front of adults in general, but the same could be said of ethnic categories).

There are also particular affinities between certain groups: typically, “Blacks” and “Arabs” are generally considered closer to one another than they are to “Asians”. On top of all this, youths can also refer to one’s religion, in particular for Muslims (most of which are considered “Black” or “Arabic”, but some of them could also be, for example, from Pakistan, India or Indonesia, in which case it is not clear at all what racial group they fall into). Finally, they can go down to a smaller level, generally that of the country: one can be referred to as an Algerian, Moroccan, Senegalese, etc. This is particularly relevant when it comes to national sport teams (cf. Marion’s extract later in the section, where a student mentions the football World Cup), but also to specific rivalries among neighboring countries (e.g. Morocco and Algeria). There are further specific cases, with some categories reaching sub-country levels (notably Algerian Kabyles).

This makes it difficult to properly name the relevant groups, because their delimitation varies depending on the classification framework that students are willing to mobilize in a given situation. For example, talking about students “with a non-French background” may be problematic: students with, say, a British background, will be in most cases identified as “White”, therefore belonging to the same group as native French students. Moreover, students from the French overseas territory are frequently non-White, yet they have no non-French background strictly speaking. Yet talking of students “with a non-European background” would not be ideal either: students with an immigrant background from Southern or Eastern Europe (e.g. Spain, Portugal, Romania), although formally White, are susceptible to claiming (or having imposed upon them) an ethnic/immigrant label, which would distinguish them from native French students. For that same reason, the racial criteria – “non-White students” – is not satisfactory either.

There is no easy way around this issue, which is intrinsic to the way ethnic differentiation operates among students (i.e. with relatively flexible labels that can vary depending on the situation). In the rest of the section, I will alternate between different expressions meant to refer to ethnicity, depending on the context (“native”, “non-European”, “non-French”). In general, though, I will favor a migratory framework (rather than racial or religious) as it comes closer to my quantitative measures. I urge the reader to keep in mind that these are imperfect terms, and that the actual classification frameworks used by students are more diverse, fuzzy, and changing than what these expressions alone suggest.

The relatively small number of interviews with an expression of social distance framed in ethnic terms (cf. Table 6-1) is somehow misleading. Students do refer to ethnicity quite a lot; but, in the majority of cases, they do so as a joke, and mostly between students from the same ethnic background. Therefore, it is not social distance, but on the contrary a proximity is implied by these discourses. This use of ethnic categories marks the belonging to a common ethnic group, of course, but also to a group of friends that can joke about what would otherwise be considered a sensitive topic, therefore demonstrating mutual trust.

> *"A1: Racism, I think I never saw it in the school. I mean, me, like, for example I am Algerian, but it's not racism to me, it's just for fun. When, like, I'm going to Kenza for example, she is, like, Tunisian as well, like, I say, like, "you filthy Arab", but since I am an Arab as well, you see, it's... It's just for fun.*
>
> *Q: Alright, so you make jokes on people's origins, but it's not aggressive.*
>
> *A1: Whereas like, if I were to go to like, Aly for Example, and say to him "you filthy Black", something like that, well, even if I say it for fun, I think he would take it the wrong way, because, since I am not Black myself, it's a bit strange. [...] For example me, everyone says "Yeah, the Arabs are thieves", so I make fun of it. Like I take someone's pencil case, and I say "oh well, I'm an Arab anyway" [she laughs]. But, you know, for fun. But otherwise no, racism, I never saw it* [at school]*."*
>
> **Jasmine & Lou, Paris 1**

> *"Q: Has it already happened to you at school, either from adults or other students, to suffer from or witness forms of discrimination? Or racism, perhaps?*
>
> *A: Hmm, no, never. It never happened to me. Then sometimes, with my friends we joke around and all that, but it's for fun. But otherwise no, never.*
>
> *Q: That is, you joke around? About people's origins?*
>
> *A: Yes. Well, sometimes Maissam she calls me Mamadou, I call her Mohamed, and all that. But it's like a private joke. There's nothing mean."*
>
> **Mailys, Savoie 2**

Importantly, this "light" usage of ethnic categories seems to occur mostly among students with an immigrant background: the "Whites" (*babtous*) or "French ones" (*français*) do not seem to positively invest these labels themselves (no mention in the interviews; and I never heard them use ethnic categories when interacting with one another, unlike non-natives[163]). Rather – and contradicting what Jasmine just said – they are called as such by their non-native peers. Marion's extract below suggests that this can result in a feeling of

163 Given that I did not do a thorough ethnographic fieldwork, but simply over-heard a few conversations here and there (mostly in the corridors), this does not necessarily mean that native students *never* use these categories. Still, it is noteworthy that, even with a relatively shallow fieldwork, I witnessed several conversations where students with an immigrant background would use ethnic terms, or even explicitly discuss their relative backgrounds, but never for native students.

discomfort, even though she immediately moderates it by saying that these are only jokes. Moreover, she is the only respondent that spontaneously complained about this.

> *"Q: Are there some students that are a bit different from you, that you feel you don't understand, like they're not on the same planet?*
>
> *A: Yes. […] The boys, sometimes I don't understand why they argue for years about their country, like, "Algeria!!", hmm... Even in class this morning, there was a battle "Algeria sucks!", "no, Tunisia sucks!". Seriously, I mean...* [breathes heavily] *And me, I don't care, I was telling them "well yes, I'm French, I'm White* ["babtou" in French]*, so what?". And that's it.*
>
> *Q: Alright. So it's something you talk about a lot between students, Whites* [babtou], *hmm?*
>
> *A: Yeah! Yes, it's like that all the time. Actually, we're all a bit classified. There is the side of Algerians, of Tunisians... And then we all play together of course, but still, every time we talk about our religions, or our…* [looks for her words]
>
> *Q: Your origins?*
>
> *A: Our origins, it always goes south, kinda. Because there is always someone who says* [she takes a deep, boy voice] *"yeah, anyway the Algerians they all suck" and then another who says "what, what did you say about Algerians?" And then I say "but calm down, anyway you are always saying that the French suck and all that, it's not right, even when I say 'go Blues'* [the French football team] *you say 'oh but you're a White'* [babtou], *well guess what, we won!* [the 2018 worldcup]*" [...]. We are always arguing, sort of.*
>
> *Q: But these arguments, how can I say… Sometimes it creates actual conflicts between students, or it just comes and goes?*
>
> *A: Noooo. It just comes and goes! It lasts 10 minutes, and then…*
>
> *Q: Alright. And so is it something that make people split up? Hang around in groups? Is it a little, like, the group of Tunisians, the group of Whites* [babtous], *the group of, I don't know.*
>
> *A: Nooooo, no. For example Ylarie* [her best friend] *she doesn't have the same origin as me. I am White, of origins of origins of origins* [No.: this is a way to say that all her family is White, up to previous generations], *and she is Black of origins of origins of origins. But it doesn't block anything."*
>
> **Marion, Paris 1**

One of the questions I systematically asked in the interviews were whether students had witnessed or experienced a form of discrimination at school (either from adults or peers; note that I did not initially say what kind of discrimination but would then ask a second question on "origins" specifically if the respondent did not mention those in their answer). Out of the 38 interviews, there are three cases where a student reported a case of open discrimination among peers (plus Marion in the previous extract, although it was less clear whether she considered it a joke or an actual discrimination). The majority of respondents explicitly declared that there was no issue of racism or discrimination among students (even though they could accuse certain teachers of being racists).

What about the three students that reported discriminatory speech, then? Interestingly, in all cases, the concerned student belongs to an ethnic group that is in a clear minority in the school: Haruhi and Julie have an Asian background, which is the case of very few students in Savoie 1 and 2, and Dahli has African origins, which again is the same only for a small number of students in Savoie 2. On the contrary, no such case was reported in Paris 1, where students with non-European backgrounds are a majority.

> *"Q: Has it already happened to you, either from teachers or students, to suffer from or witness forms of discrimination?*
>
> *A: [...] Well since I am Japanese, often, well not often, but, a little, there are people that say... especially Damian. Now I don't really like him anymore, because he often says that I am yellow, whereas I'm not yellow. He says I am Asian, he says that. For example there are some people, to mimic me, they do this* [he squints].*"*
>
> **Haruhi, Savoie 1**

> *"A: At the beginning of the year, some boys were always bothering me. Well, last year as well, but it's about my family name* [her family name sounds Asian]. *[...] They always bothered me, and all that."*
>
> **Julie, Savoie 2**

> *"Q: Has it already happened to you, either from teachers or students, to suffer from or witness forms of discrimination?*
>
> *A: [...]* [No.: her first answer is about someone being called fat. She only moves on to racism after I specifically ask for discrimination about origins] *Toward me, yes, but I don't care. And other people, no, I never saw it.*
>
> *Q: And toward you, what have you been told?*
>
> *A: "Filthy Black", things like that. But I don't care.*
>
> *Q: Who said that to you?*
>
> *A: People that left. And then there was Bëtul, but actually, it's because... Actually she didn't mean it, but it just came out on its own. [...]*
>
> *Q: You were having an argument when she said that?*
>
> *A: Yes. But she didn't mean to say it."*
>
> **Dahli, Savoie 2**

Dahli's story is particularly interesting, in that the person that insulted her, Bëtul, is a good friend of hers. Therefore, this instance of racism is not necessarily indicative of same-ethnic preferences in the choice of friends. In fact, Bëtul and Dahli were originally in conflict for another, unrelated reason (of course, it is difficult to know whether the ethnic difference somehow aggravated this conflict – in any case, it was very similar to the kind of quarrel that can arise among many other students). Ethnic categories were imported into that conflict by Bëtul in an effort to insult Dahli; she "ethnicized" a conflict that originally stemmed from a different source.

The case of Inès is a bit different but can be read in the same light. She got into a dispute with a group of former friends and classmates; then, at a given point in that dispute, a seemingly incidental remark gave an ethnic turn to their argument. Note that Inès has Algerian origins herself.

> *"A: Marianne and I, we used to be friends. It started in biology class, she said "eeeeh, you're in love with Nathan, you're in love with Nathan". So at first I told her to stop, that it wasn't funny. Then she started making songs about it, and she kept going. [...] Then it was, I'm in love with Noam, then someone else. And then I told here, "I'm going to go to my mother, you won't be laughing after that". Then she did, she did lots of things, stirring trouble, and then she started to cause trouble on snapchat. [...] She sent a lot of messages [...], saying I was a whore, a pussy because I told my parents, that I... well. [...]*
>
> *Q: But how did Kéziah, Chahima and all the others get into it? Marianne asked them to send you messages as well?*
>
> *A: So, actually, I sent a picture to Iris, on snapchat. And then, Marianne was with Iris at that moment, but I didn't know* [she laughs nervously]. *So Marianne sent me a vocal message, she said, like, "oh, you're a shame", nanana, I don't know what. And then Kéziah, Kéziah too, she was with them* [laughs], *and so Kéziah started sending me messages as well. [...] So I told my mother. My mother sent a vocal message. And then she said, hmm, my mother she said "we have to tell you in Chinese or what? Just stop!"* [nb: French expression, can translate as "How do I need to say it for you to understand?"] *And then, Marianne sent a vocal message, she said "no, me I don't speak French, I speak Kabylian". So me I told her [...] "but then why are you in a French school?". And then she told me "oh, but you're racist, you don't want me to be in my country", like, "we're going to see that with a judge", nanana. She said "I am going to report to the police with my parents", blablabla."*
>
> **Inès, Paris 1**

As for cases of "blunt" racism (i.e. seemingly without any particular reason), this was reported to me only once, again by Dahli. Students she did not know insulted her, apparently without reason, which clearly does not fit the interpretation about conflicts being "ethnicized" *a posteriori*. Given what other students reported, this seems relatively rare, however.

> *"A: There are also some people, but not in the school, people I don't know. Like, they come and insult like that, but...*
>
> *Q: Who are they?*

A: I don't know them. They're often 4th graders, they come and insult you.

Q: They come, they say "filthy Black" and then they leave?

A: Yes. But they're strange. [...]

Q: And it was always 4th graders?

A: Yes, most of... actually, it was always the same people.

Q: And you don't know who they are?

A: No. I know some 4th graders, but not them. They're losers.

Q: How do you know they are losers if you don't know them?

A: Well, because no one likes them. Like, when I go to the 4th graders and I ask, 'you know the ones that insulted me?', because I want to know. They tell me 'him, he's a loser. We never talk to him'."

Dahli, Savoie 2

Moreover, on the quantitative side, multivariate models on disliking networks show no effect of ethnic dissimilarity on the probability of a nomination (see the models in Appendix 4C).

All in all, students' use of ethnic perceptive categories appears somehow ambiguous. In most cases, these categories are used to joke around among friends. On the one hand, this may actually contribute to reducing the social distance among students. Given that ethnic frameworks exist outside of the school already, and can be antagonistic in some instances, turning those into a joke might be a way to defuse possible tensions or, for students that belong to a stigmatized group, to turn the stigma around. Moreover, since the use of ethnic categories is generally discouraged by teachers, a plausible hypothesis is that these jokes contribute to a shared youth culture which, by opposition to the adults, creates a feeling of sameness among students, including those that have different ethnic backgrounds.

On the other hand, these ethnic categories, because they are part of students' perceptive framework, can be mobilized negatively in case of conflicts. I would therefore expect potential tensions born from socioeconomic and/or academic differences to be sometimes interpreted in ethnic terms by students, as these are the available "tool of thoughts" at their disposal for expressing a feeling of injustice (the strong correlation between academic results, ethnic background and socioeconomic background making either of those a satisfactory grid of interpretation for many situations).

This pattern of findings seems coherent with the interpretation proposed by Bonnéry (2006). Considering primary and secondary school students in working-class and immigrant suburbs, he found that ethnic categories circulate within students' peer groups in ways similar to what I described above, i.e. often as a joke and among non-natives (similar patterns are found by Lepoutre, 1997 or van Zanten, 2000). Furthermore, some primary school teachers try to attach value to the migratory origins of their students, as a strategy to make them feel positively attached to the school and classroom (e.g. organizing cultural events about different countries in class). However, this circulation of ethnic categories means that they are available for interpreting antagonistic or painful experiences as well: as they grow up and increasingly struggle at school, many students start to "ethnicize" their difficulties, i.e. to interpret them through ethnic lenses (calling teachers racists when they get punished, for example). This is further supported by the fact that, outside of the school, several instances of racism or discrimination contribute to making this grid of interpretation relevant to students.

Although the situation described by Bonnéry appears more extreme than what I found in my sample (which is probably due to the fact that these processes are more pronounced in segregated than in diverse schools), this idea of ethnic categories being used *a posteriori* to re-interpret conflicts stemming from various sources fits well with some of the cases described above.

5. Money and Consumption

The final framework we will explore is that of economic capital; that is, money and conspicuous consumption. As can be seen in Table 6-1, it is very rare for respondents to express social distance by referring to a peer's income or goods. Moreover, when they do so, it is generally misleading as per one's social background. Indeed, conspicuous consumption, in particular through clothes or smartphones, is actually more typical of working- or middle-class students than of upper-class ones. In the questionnaires, students were asked whether they had a personal smartphone. The probability of having a smartphone, rather than not, actually *decreases* as social background increases. In a logit model that takes as its outcome ownership of a smartphone, and controlling for gender as well as school, a working-class student with an SA score of -0.96 is predicted to be more than two times more likely to have a smartphone, rather than not, than a student with a score of 1.44 (see Appendix 6B).

As for clothing, one's money is primarily expressed through wearing clothes "with a brand" (i.e. a famous brand). Yet in Paris 1 and Savoie 1, the brands most often named by students in the interviews are sport brands typical of working-class tastes.

"Q: What do you mean "having brands"? What kind of brands?

A: I don't know, like, nike, adidas. [...] Not just because it's brands, but because people are going to think: it's expensive, so it means they can afford it, they're not poor."

Laurence, Paris 1

The situation is a bit different in Savoie 2, where other brands are more frequently cited, particularly by girls (sport brands remain popular among boys). It is less clear to me whether conspicuous consumption is also typical of students from a lower social background in this school. What is interesting is that having brands specifically seems less important here, especially compared to Paris 1; while students can also get mocked for how they dress, this seems to pertain to the actual look of their clothes, not the fact that it comes from a famous brand or not.

"Q: You have brands that you like?

A: Hmm, Fila, Nike, Adidas, Calvin Klein... well, all the brands.

Q: And is it something you talk a lot about with other students, or with your friends?

A: Hmm, we don't talk about the brands, but we talk about the clothes. For example, when there are sales, we talk about what we're going to buy, but not the brands."

Dahli, Savoie 2

Not only are sport brands less popular, but the overall dressing style also favored by most students in Savoie 2 is further away from working-class standards than in Paris 1 or Savoie 1. Simply looking at the playground, one can see that most students wear jeans, much more so than in other schools. A few interviews even suggest that tracksuits, for girls at least, are stigmatized by other students.

"Q: Are there students that are more often, like, a little rejected? Or even that get bothered or victimized?

A: Yes. Coline and Mathilde, in my class. They always dress with sport clothes."

Jade, Savoie 2

Therefore, while there is social distinction through clothing in Savoie 2, it does not pass by conspicuous consumption strictly speaking. The fact that brands are least important in the school where middle- or upper-class tastes are prevalent (Savoie 2) but are most important in the school where working-class tastes are popular among the majority of students (Paris 1), thus supports the idea that displaying one's wealth through branded clothes is more typical of working- or lower-middle class students than of upper-class ones.

There could be various reasons behind this seemingly reversed pattern of conspicuous consumption. The rejection of an excessive display of wealth, grounded in an ethos of subtlety in tastes that tells one apart from *parvenus* or *nouveaux riches*, is a distinctive trait of the French upper-class, particularly those with high levels of cultural capital (Bourdieu [1979] 2016). The same principle might be at work here. There is also probably a selection bias in the upper-class families that attend socially mixed and, for Paris 1 and Savoie 1, public schools; in particular, families with left-wing political values may well be over-represented. Perhaps more simply, upper-class students might also be adapting to the context of social mixing by not displaying their wealth in front of lower-background peers. It also makes sense that the need for distinction would be more salient among students that belong to close socioeconomic groups[164].

In any case, it is clear that the display of economic capital through conspicuous consumption is not the mechanism through which students' socioeconomic backgrounds are signaled to peers.

Conclusion

Out of the perceptive categories that we have explored, academic ones were clearly the most salient in students' discourses. A long-standing debate in the sociology of education pertains to the specific role of the school in the formation of socially homogeneous peer groups among students, as opposed to family socialization or communal ties (see van Zanten 2000 for a review). The results of this chapter strongly point toward academic labels and rankings being imported into students' subcultures, such that they are used as a perception tool for differentiating among peers.

The statistical models of chapter 5 clearly supported this idea for Paris 1 and 2, where the parameters meant to capture homophilic selection based on academic results were large and statistically significant. However, the evidence was less clear in Savoie 1 (estimated significant at the 90% threshold only) and was simply lacking in Savoie 2 (non-significant estimate; cf. Table 5-3). Moreover, even in Paris 1 and 2, contribution scores suggested a rather modest inducing power of academic selection on socioeconomic homophily (cf. Table 5-4). Therefore, the quantitative and qualitative approaches seem to disagree over how

164 This should result in a curvilinear pattern: conspicuous consumption would be typical of working- or lower-middle classes, opposing them to both the upper-classes (because they have other means of distinction) and the poorest fraction of the working-class (because they simply do not have the means to purchase branded clothing). Nevertheless, this is something I am unable to test for with my current data (not enough interview extracts dedicated to these topics).

important academic factors are in student sociability, and, consequently, in mediating socioeconomic homophily (though they both agree that it has some importance at least). Nevertheless, this might not be as contradictory as one may first think.

Indeed, what the interviews suggest is that homophilic selection may not so much operate on academic results as on academic attitudes; and while both tend to be correlated, they remain distinct in many cases. In that regard, the models of chapter 5 were based on results exclusively, not attitudes. Moreover, despite the clear impact of academic categories on students' sociability in Paris 1 and Savoie 1, there are also hints of individual strategies and collective norms that mitigate the effects of academic heterogeneity. Students, it seems, are sometimes trying to reject academic categories as an illegitimate form of differentiation among peers, though they also endorse them on many occasions. Finally, it should be noted that, just like with socioeconomic background, homophily and selection are distinct concepts: academic homophily likely has other sources than academic selection. In that regard, the prevalence of academic labels in the respondents' discourse may be a consequence of academic homophily, not just a cause of academic selection. The fact that their friendship networks are homophilic in terms of academic attitudes may push students toward academic frames of perception, as these simply fit the state of their networks as they observe it.

Finally, there are reasons to believe that disagreements about academic attitudes can be more antagonistic among working- or lower-middle-class students, than between working- and upper-class students. This may imply that academic distinctions do not induce as much socioeconomic homophily as the correlation between socioeconomic background and academic performance would predict. The evidence available, however, particularly points toward implications in terms of peer status, as having better grades may come at greater relational costs for lower-background students specifically, under certain circumstances at least (Parisian schools in this case).

The second main perception framework that was observed in the interviews is made of cultural categories – clothing and music in particular. Unsurprisingly, these can act as markers of socioeconomic origin. This seems rather coherent with the estimates from chapter 5: cultural tastes appeared as a notable factor of selection in all schools, and they markedly contributed to socioeconomic homophily in Paris 2 – to a lesser extent in Paris 1 and Savoie 1. Nevertheless, tastes and styles remain a secondary concern for many students, and they usually do not crystallize into perfectly coherent styles nor into well-bounded groups of students. While this is something that I lack the space to develop here, it is also noteworthy that students, when asked to define friendship ("*what is a friend, for you?*"), rarely mention a similarity in tastes or personality (7 respondents out of 38). In that sense, musical and clothing styles are not (yet?) a key marker of personal and group identity, and their impact on friendship choices likely remains implicit and partly unconscious. Based on the literature, one may expect these categories to gain importance as students move on to high school.

Ethnic categories are clearly present in students' perceptive frameworks. They came up relatively rarely in the interviews, and few students reported instances of blunt racism at school; but they seem to be commonly used as material for jokes within peer groups, particularly for students with a non-European background. As a result, ethnic labels circulate within peer groups, and are available for interpreting antagonistic situations when they arise. This is coherent with the models of chapter 5, which suggested that ethnic homophilic selection exists in all four schools and that it is the primary determinant of ethnic homophily – meaning that the (moderate) ethnic segregation of friendship networks is due to students' propensity to pick friends with similar ethnic backgrounds, not to the indirect impact of correlated factors as is the case of socioeconomic homophily (cf. the conclusion of chapter 5 as well as Appendix 5G). In short then, there are likely ethnic preferences and/or discrimination at work in the choice of friends, even though, for the most part, this does not seem to result in ethnic tensions in the schools. In any case, ethnic labels may sometimes act as proxies for other types of homophilies or feelings of distance, in cases where they are correlated to academic results or to socioeconomic background.

Finally, evidence suggests that conspicuous consumption does not play much of a role in mediating socioeconomic hierarchies. In that regard, cultural capital, rather than economic capital, should be the main dimension of socioeconomic background that shapes students' sociability.

Altogether, what do these different elements say about students' perceptions of their friends' socioeconomic origins? For the most part, this perception seems to be indirect, implicit, and probably relatively fuzzy. Compared with gender, ethnicity or academic categories, social class is not a clear frame of distinction known to all students and does not elicit conscious homophilic selection. In that regard, the effects of socioeconomic selection that were estimated – for certain schools and waves – in chapter 5 probably pertain to "micro foci" for a good part, i.e. daily choices in the activity or spaces that students attend, or the discussion topics that they bond around[165]. Similarity in tastes and personality among same-background peers may play a role, but most likely in a diffuse way, for instance by impacting the resilience of friendships in the long run.

165 To qualify a discussion topic as a "foci" is probably to stretch the concept a little too much, as I have mostly used it in a spatial sense so far, but it still makes sense in opposition to psychological dispositions such as preferences, discrimination, disgust of certain others, moral judgments, etc.

Chapter 7 – Differences between Schools

So far, the analysis has mostly consisted of considering all schools together, in order to identify shared patterns among them. This relates to how the research design was thought of in the first place: the chosen schools share a common property of being socioeconomically diverse, but they differ on many other aspects. As a result, common results among these schools have a good chance of being true for most diverse schools. However, in the course of the previous chapters, we also encountered striking differences between the four cases. This begs the question of where these differences come from, and of which properties of the schools they relate to. In particular, one has to wonder why there is such strong socioeconomic homophily in one case (Paris 1) and almost none in another (Savoie 2).

From a methodological point of view, this is a rather difficult question, because the size of the sample makes it impossible to consider school-level associations between variables in a statistical fashion. Instead, one can resort to the logic of case studies: by obtaining qualitative details about each case, it is possible to assess the plausibility of different theoretical accounts of the differences among them. The idea is that the alleged mechanisms that may generate different outcomes across schools – that is, different levels of homophily – ought to have other impacts as well, leaving various “traces” beyond socioeconomic homophily itself. Whether these expected impacts are observed in the data can therefore lend support, or on the contrary reject, the hypothesized explanation. For example, if one were to assume that there is more socioeconomic homophily in Paris 1 than in Savoie 2 because of higher levels of residential segregation, then it should be the case that (a) levels of segregation differ across schools and (b) residential distance accounts for a large part of socioeconomic homophily (which, based on chapter 5’s estimates, is not the case: there is no evidence for residential propinquity in Paris 1 once other factors are controlled for).

Because there are many different factors that may explain between-school differences, and because the possibility exists that the observed networks have a strong element of randomness that makes them unique even within a given school context (i.e. other cohorts in the same school may exhibit different features), this approach is far from perfect; but, at the very least, it may help to narrow down the range of plausible explanations, by ruling out those that are inconsistent with the observed data. Altogether then, the analyses of this chapter are mostly suggestive: the aim is to formulate plausible hypotheses, that future studies may then test on larger data sets.

Among the four studied schools, Paris 1 in particular stands apart from the others: not only does it exhibit the highest levels of socioeconomic homophily, it is also the only school where this homophily is not stable over time, but strongly increases between waves 1 and 6. Section 1 is therefore entirely dedicated to understanding this peculiar pattern of increasing homophily. Section 2 then considers the association between socioeconomic homophily and various school-level features, with a particular focus on the two extreme cases of Savoie 2 (very low homophily) and Paris 1 (very high homophily).

1. Socioeconomic Homophily in Paris 1: from Moderate to High Levels of Segregation

This first section focuses on the case of Paris 1 and, more precisely, on the pattern of increasing socioeconomic homophily observed in that school. Indeed, even though Paris 1 already seemed somehow more homophilic than the other schools in wave 1, it is this drastic increase that results in it having markedly segregated friendship networks by wave 6. The aim of this first section is therefore to understand where this increase came from, and through which mechanisms it unveiled in the networks. Section 1.1. is mostly descriptive, offering a visual inspection of the friendship networks to try and understand how their structure changed between waves 1 and 6. Section 1.2. then dives into more specific hypotheses about the relational processes underlying the observed evolution, first considering the role of tie formation and dissolution, and then linking this to the generative models presented in chapter 5.

1.1. A Closer Look at the School's Relational Structure

Before trying to figure out the how and the why of the homophily pattern observed in Paris 1, it is first necessary to understand what this pattern consists of exactly. Indeed, saying that socioeconomic homophily increased over time does not tell us exactly in what ways it increased, nor what this means in terms of network structure and of the sociability of students. Was this driven by the progressive disappearance of formerly heterophilic ties, or on the contrary by the apparition of many new homophilic ties? Did all students equally moved from moderately homophilic to strongly homophilic personal networks, or is the trend driven by a small number of individuals whose friendships drastically changed? And, to begin with, was socioeconomic homophily equally distributed across the population, or are some students strikingly more homophilic than others?

1.1.1. Friendship Networks in Waves 1 and 6

In order to see the frame of the relational structure more easily, we can look at the networks made of students' answers to the question: "who are the five friends that you spend the most time with at school?" (from here on referred to as "5-friends networks"). Not only are these less dense than other friendship networks, they also make it clearer which are the main groups in terms of who spends time with whom, thanks to the upward limit on students' outdegree downplaying the core-periphery dimension of the graph[166]. Both the strength of socioeconomic homophily and its evolution over time are very similar in the 5-friends and 'very good friend' networks, so the conclusions drawn from one can straightforwardly be extended to the other.

Figure 7-1 shows the 5-friends networks of waves 1 and 6 (end of 1st year – beginning of 4th year). Nodes are colored based on socioeconomic background (Socio-Academic Scores), with darker red indicating higher backgrounds. Girls are represented by circles and boys by diamonds. Additionally, the students who belong to the international section – "internationals" for short – have a star ("*") next to their name (all names have been changed for anonymization).

In wave 1, the fact that the network is homophilic in terms of socioeconomic background is visible to the naked eye. In particular, there is a small, dense group of upper-class students on the right hand on the graph: these are students from the international section. The rest of the graph, however, does not appear to be strongly homophilic: in particular the large group of girls on the left-hand part contains many nodes from both very low and very high backgrounds. Statistical analysis reveals that there is also some socioeconomic homophily among students outside of the international section but, in wave 1, it remains relatively low (cf. Table 7-1 below). Therefore, in wave 1, the separation of the internationals from the rest of the network accounts for most of the total socioeconomic homophily in the school, and for the fact that it is higher here than in other schools[167].

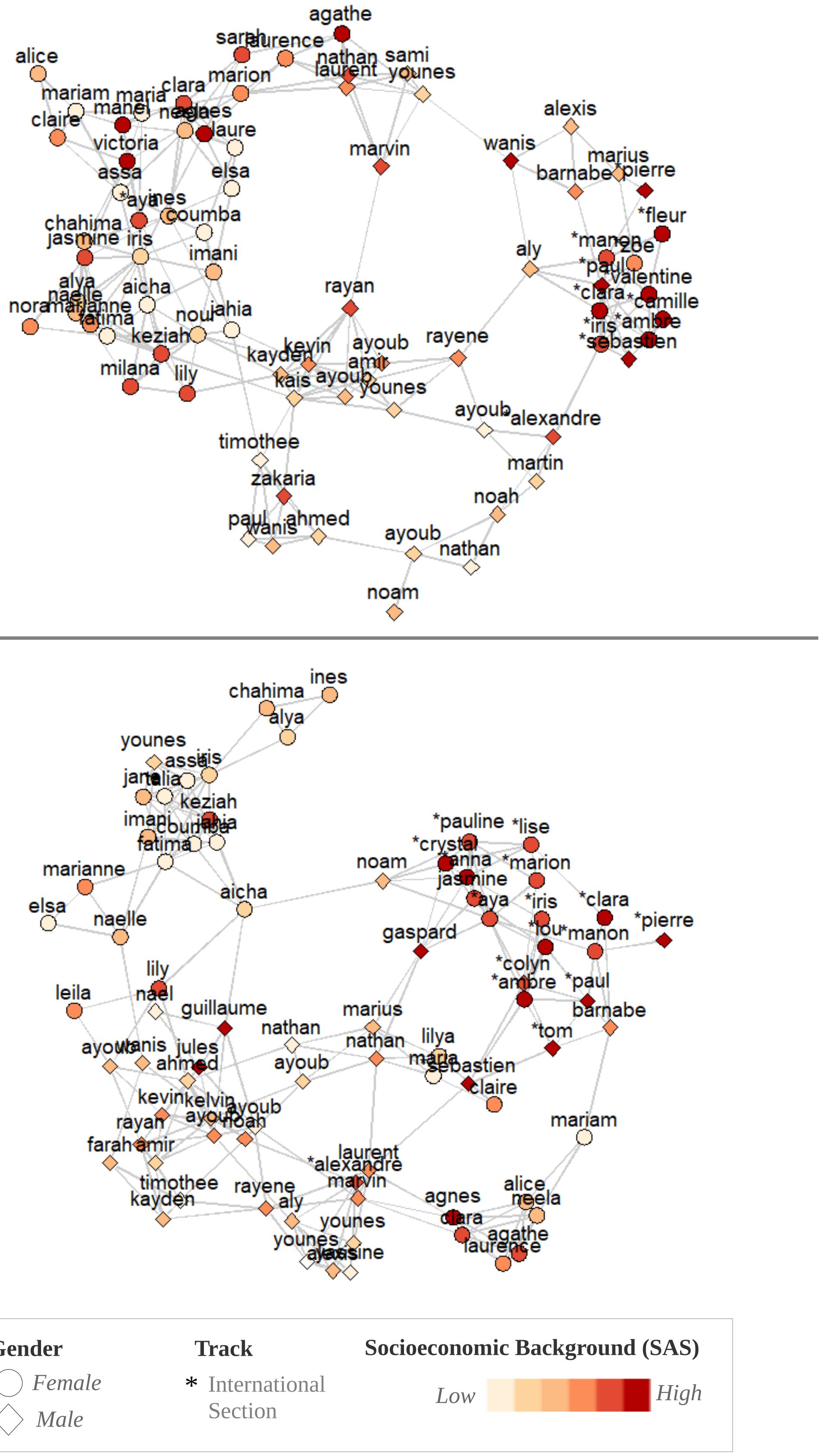
166 A common issue when representing networks with unlimited nominations is that the few individuals that get along with a lot of others represent a significant share of the total number of ties. This tends to mask the differences among less central individuals: two students may end up close to one another on the plot because they are both marginal in the network, even though they have in fact few to no relationships in common.

167 This may seem contradictory with the findings from chapter 5: contribution scores indicated an overall modest contribution of the international section to the network-level homophily (Table 5-4). However, one has to remember that contribution scores considered the impact of the model parameters associated to same-section

Two and a half years later, in wave 6, the situation changed. The group of internationals is still easily identifiable on the right of the graph, and seems to be as insulated as before – even more so in fact, as the few ties to lower-background students that existed in wave 1 have disappeared. More importantly, socioeconomic homophily now seems prevalent among non-internationals as well. Instead of the large mixed group of girls from wave 1, we can now see, on the upward left part of the graph, a dense group of girls that are almost all from lower backgrounds (except one red node, Keziah) – and, incidentally, from non-European immigrant backgrounds. Among non-international boys, things are less clear, but there also seems to be some homophilic subgroups (e.g. small clusters of working-class boys at the bottom of the graph, even though they also share a few ties with upper-class nodes above them). Thus, socioeconomic homophily appears to have particularly increased among non-international students.

selection *once other factors were accounted for*; in other words, the “pure” tendency of students to sort out their friends based on them belonging to the international section or not. Yet students from the international section also differ from “regular” ones regarding several other covariates: they live in a different neighborhood (meaning residential propinquity will contribute to section homophily), they remained in the same classroom for four years (whereas the “regular” students that are in the same classroom as the internationals tend to change every year), they are all excellent students (so academic selection will also entail section homophily), etc. All these effects were not captured by the contribution score of the international section parameters, but they are reflected in the fact that the aggregated amount of socioeconomic homophily is low when considering ties among non-internationals exclusively. In other words, once again, homophily must be differentiated from homophilic selection.

Figure 7-1: 5-Friends Networks observed at Waves 1 (left) and 6 (right) – Paris 1



To confirm that this is the case in the networks of “very good friend” nominations as well, I resort to dyad-independent ERGMs, similar to those of chapter 4 but with a slight variation. In chapter 4, socioeconomic homophily was measured through a single ERGM coefficient, which captured the change in tie odds associated to a one-point change in the absolute SAS difference between nodes. Here, I split this effect into three coefficients: one for dyads where both students belong to the international section, one for dyads where they belong to the regular section, and one for mixed dyads with one student from each section. This is done by interacting the effect of SAS difference with the matching and emission effects of the section covariate. Of course, the primary effects associated to these section covariates are also included as controls in the model (otherwise the interaction would be impossible to interpret).

Results are in Table 7-1. First, one can note that friendships among internationals never exhibit any socioeconomic homophily, which is logical given that they are all from similar backgrounds. Second, the socioeconomic homophily among mixed dyads – which can be interpreted as capturing the difference between lower- and higher-background non-internationals in their propensity to be friends with internationals – steadily increases from waves 2 to 6. However, it had first decreased between waves 1 and 2, such that the value reached in wave 6 is not significantly larger than that of wave 1 (-0.992 vs -0.852, both being extremely high values). Finally, the log-odd coefficient indicative of socioeconomic homophily among non-internationals has strongly gone up, from -0.144 in wave 1 to -0.357 in wave 6. Bringing back the (now famous) pair of fictive students, one of which is from a working-class background and the other from an upper-class background (SAS difference of 2.44), we can predict that, if they are both non-internationals, their friendship will be 1.42 times less likely than that of same-background students in wave 1. This is potentially lower than what is found in Paris 2 and Savoie 1 at the same time – granted that inter-school comparison of log-odd coefficients is always tricky. In wave 6, however, this friendship is now 2.39 times less likely than between same background students, which is much higher than what is found in other schools[168]. Therefore, it is indeed the case that socioeconomic homophily has mostly increased among non-international students between waves 1 and 6.

168 1/exp(-0.144*2.44) = 1.42 and 1/exp(-0.357*2.44) = 2.39.

Table 7-1: Socioeconomic Homophily among International-, Regular- and Mixed-Section Dyads (ERGM) – 'Very Good Friend' Nominations, Paris 1

Absdiff SAS * ...	wave 1	wave 2	wave 3	wave 4	wave 5
… Regular section	-0.144** (0.069)	-0.257*** (0.065)	-0.253*** (0.072)	-0.184** (0.073)	-0.357*** (0.086)
… International section	0.264 (0.423)	0.801** (0.372)	0.558 (0.358)	-0.058 (0.331)	0.727* (0.396)
… Mixed dyads	-0.853*** (0.199)	-0.449*** (0.128)	-0.592*** (0.179)	-0.647*** (0.181)	-0.992*** (0.184)

***: p<0.01; **: p<0.05; *: p<0.1

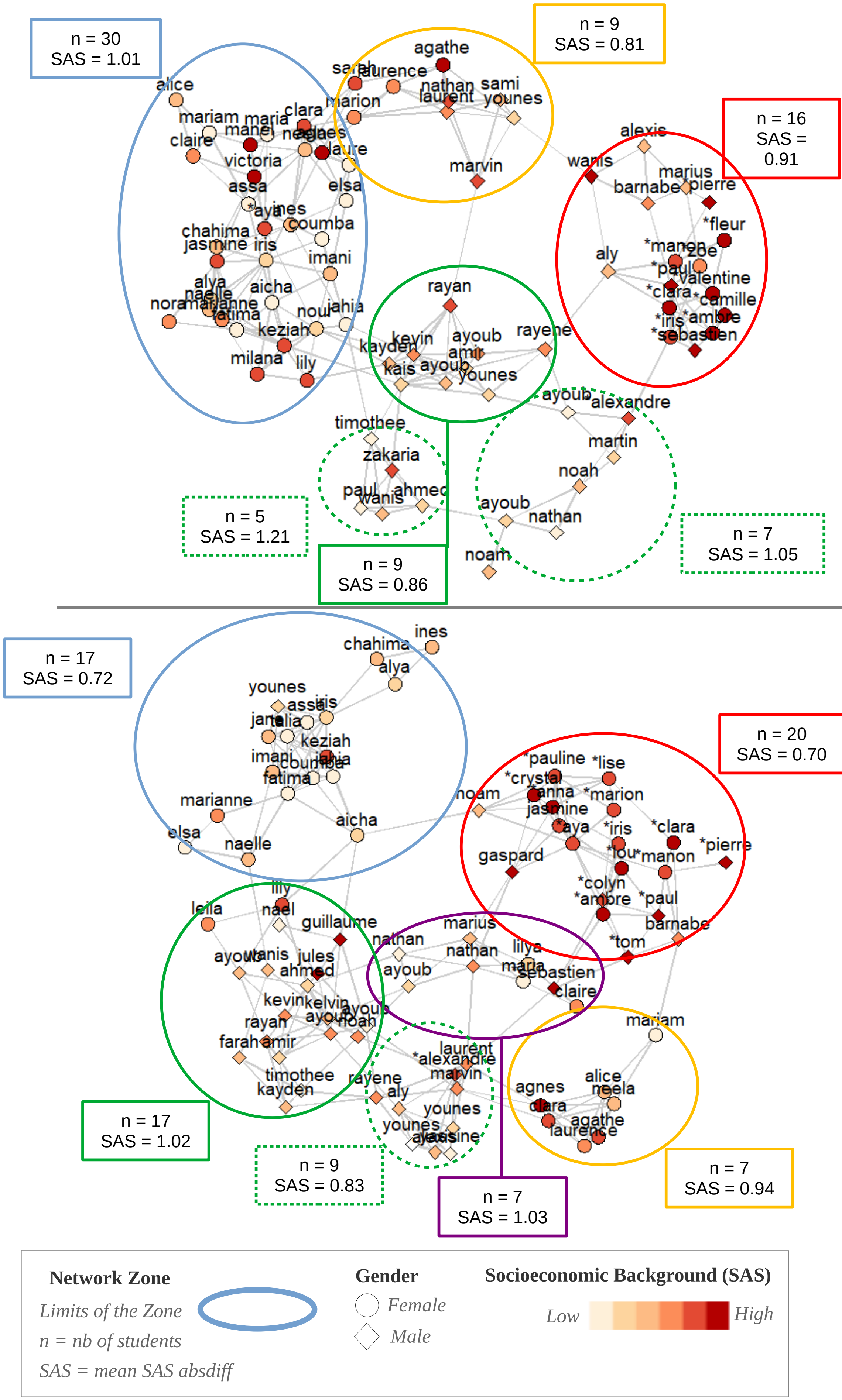
Example of interpretation: in wave 1, for two students that are not in the international section, each point of absolute SAS difference is associated to a change in the log-odd of a tie between them of -0.144. For fictive working-class (SAS=-0.96) and upper-class (SAS=1.48) students, this represents an odd-ratio of 1.42 (1/exp(-0.144*2.44)).

1.1.2. Main Zones of the Friendship Networks

Let us go further in the visual exploration of the networks. To follow the shifts in friendship groups more easily, Figure 7-2 shows the same networks as Figure 7-1, but this time with circles drawn around the main zones of the networks. To be clear, these zones are not the product of a statistical analysis, nor do they perfectly correspond to actual friendship groups as perceived by students: they are qualitative approximations meant to facilitate the reading of the graph. It is also important to remember that these zones are based on 5-friends networks: they are useful for capturing students' main groups of daily interaction, but they tend to mask the fact that many 'friend' and 'very good friend' ties also exist in-between the different zones. Conversely, not all students in a given zone are friends with one another, though they can be expected to belong to the same sociability circles.

In Figure 7-2, the numbers within the boxes indicate the number of students in the zone ("n") and the average value of the absolute SAS difference of ties within the corresponding zone ("SAS"). This is a measure of the socioeconomic homogeneity of these zones: the lower the value, the more friends within it tend to be similar in socioeconomic terms (a lower difference means that the SAS scores are close).

Figure 7-2: Visual Zones in the 5-Friends Networks in Waves 1 (left) and 6 (right) – Paris 1



Starting with the network of wave 1, six zones can be identified. First, the red circle indicates the zone of the international students. Nevertheless, we also see that a few nodes are from a middle- or lower-middle-class background and are not from the international section: Aly on the right, as well as the four boys at the top, stand slightly apart from the core made of internationals, and act as a sort of interface with the rest of the network. These are scholarly diligent boys who are in the same classroom as the internationals and who do not get along well with most other non-international boys in their classroom (cf. the discussion of Aly's case below). Second, the blue circle corresponds to the large, socially mixed zone of girls that I mentioned above. Third, the orange circle marks out a small zone of boys and girls that are not from the international section, but that are mostly from upper-class or upper-middle-class backgrounds. Finally, the three green zones are composed of non-international boys and are relatively equivalent for the purpose of our analysis (they are rather close when considering the weaker 'friend' nominations).

Considering SAS differences, the mean value over all ties in the network is 0.98, so zones with mean values inferior to this are more homogeneous than the network average. This is the case of the red, orange and central green zones. Interestingly, the red zone, that of the internationals, does not appear to be the most homogeneous one among these, mostly owing to the presence of the non-international boys mentioned above (0.91 for the red vs 0.81 and 0.86 for the orange and green respectively). On the other hand, the large zone of girls in blue is relatively mixed, with a mean difference of 1.01, as well as the two smaller zones of boys in green (1.05 and 1.21, though in these cases it mostly comes from one or two students dragging the mean upward, due to the small size of the zones)[169].

Moving on to wave 6, the group structure remains relatively similar – the colors from wave 1 have been kept to indicate continuity over time in the delimitation of the zone (i.e. most students at the core of the zones remain the same). However, there are important changes as well. Starting with the red circle, that is, the zone of the internationals, its socioeconomic homogeneity strongly increases, with an average SAS of 0.70, against 0.91 in wave 1. This is due to the lower-background non-international boys leaving the group – some of them left the school and others moved toward the green zone. There are still a few non-internationals in this group (Barnabé, Jasmine and Gaspard), but they are all from a middle-upper or upper-class background.

169 This is still less than what should be observed in perfectly non-homophilic groups: if ties were distributed at random in the whole network, the average SAS difference should be of 1.24 in wave 1 (1.28 in wave 6).

The biggest change comes from the former blue zone of wave 1, which was one of the less homogeneous parts of the network at the time. In wave 6, it turns into a highly homogeneous group of lower-background girls (SAS of 0.71, against 1.01 in wave 1), with only one upper-background (Kéziah) and one upper-middle-background (Marianne) nodes left. The zone is also smaller and denser than it used to be, with a well-bounded central group of 10/11 nodes and two small peripheral triangles (above and under the main group). Where did the former members of this zone go, particularly those from an upper-background? One of them left the school (Victoria), and others moved to the red (Aya, Jasmine), green (Neela) or orange zone (Lily) (plus Milana and Manel who missed the questionnaire in wave 6, but should be close to the small triangle made of Marianne, Naelle and Elsa at the bottom of the blue zone).

Another important change pertains to the orange zone. Although it was already relatively homogeneous in wave 1, with mostly upper-background students, it nevertheless remained connected to the blue zone (even more so when considering 'friend' and 'very good friend' nominations). In wave 6, however, the whole zone moves on the graph, and is now beneath the red zone. This shows that the girls from the orange zone have almost entirely cut ties with those from the blue zone. There is still some heterogeneity inside the orange zone (4 upper-, 2 middle-lower and one lower-background), but it is primarily structured around a cluster of 4 upper-background girls (Laurence, Agathe, Clara and Agnes), which is also connected to middle-upper-background boys from the neighboring green zone (Laurent, Marvin, Alexandre).

In the middle of the graph, we can see a purple zone that comprises 3 girls (Claire, Maria and Lilya) and 4 boys[170]. Two of the girls were formerly in the blue zone (Lilya arrived in the school in wave 2), but they acted as a kind of interface with the upper-class girls from the orange zone (especially Maria). In wave 6, they now form a small zone with a few boys as well, that stands completely apart from both the red and blue zones.

To sum up then, not only does the socioeconomic homogeneity of the zones of girls increase between waves 1 and 6, but there is also an increasing disconnection between them, further reinforcing the network-level homophily. Finally, there are also some changes within the green zones, mostly made of non-international boys, but the socioeconomic diversity of

170 Sebastien seems to be in this purple zone, but this is a visual illusion – in fact, most of his ties are attached to the red zone (since the network has much more than two dimensions, the layout algorithm usually has to make some sub-optimal placements).

these zones does not change much (they possibly increase a little). Clearly, this is not where the increase in socioeconomic homophily across waves is located.

1.1.3. Two Examples of Increasingly Homophilic Students

Among the students who moved from one zone of the network to another between waves 1 and 6, two were interviewed twice, once shortly after wave 2 and a second time one year later, after wave 4. Aly is a boy from a working-class background and Jasmine a girl from an upper-class background. Both had many heterophilic ties in wave 1, that mostly disappeared by wave 6.

Let us begin with Aly. We already encountered him in chapter 6 (section 2.2.2): he is a non-international boy with high academic ambitions (due to parental pressure), and tried to befriend students from the international section for which he was mocked by other non-international boys. In particular, Kevin, Aly's neighbor, seemed to resent him a lot, both for hanging out with internationals and for being overly zealous at school. This was made clear to me by students during the interviews at the beginning of the second year of middle school, but most of the mocking happened during the first year.

Figure 7-3 represents the ego-networks[171] of Aly at each wave. All emitted and received 'very good friend' nominations are shown, with arrows indicating their direction. Strong blue lines indicate ties that also correspond to a nomination in the 5-friends networks. As usual, node shapes indicate gender, color indicates socioeconomic background, and students from the international section have a star next to their name.

From wave 1, Aly's network appears rather mixed, with both internationals and non-internationals, as well as students from different backgrounds. This is even more true in wave 2, where he is friends with many internationals, coherent with his narrative in the interview at this time. Nevertheless, in wave 3 (end of second year), the situation seems to have changed: most internationals have disappeared (except Alexandre, who is one of the few internationals to hang out with many non-internationals), and Kevin now declares Aly as a good friend. This nomination is not reciprocated yet in wave 3, but it will be by wave 6.

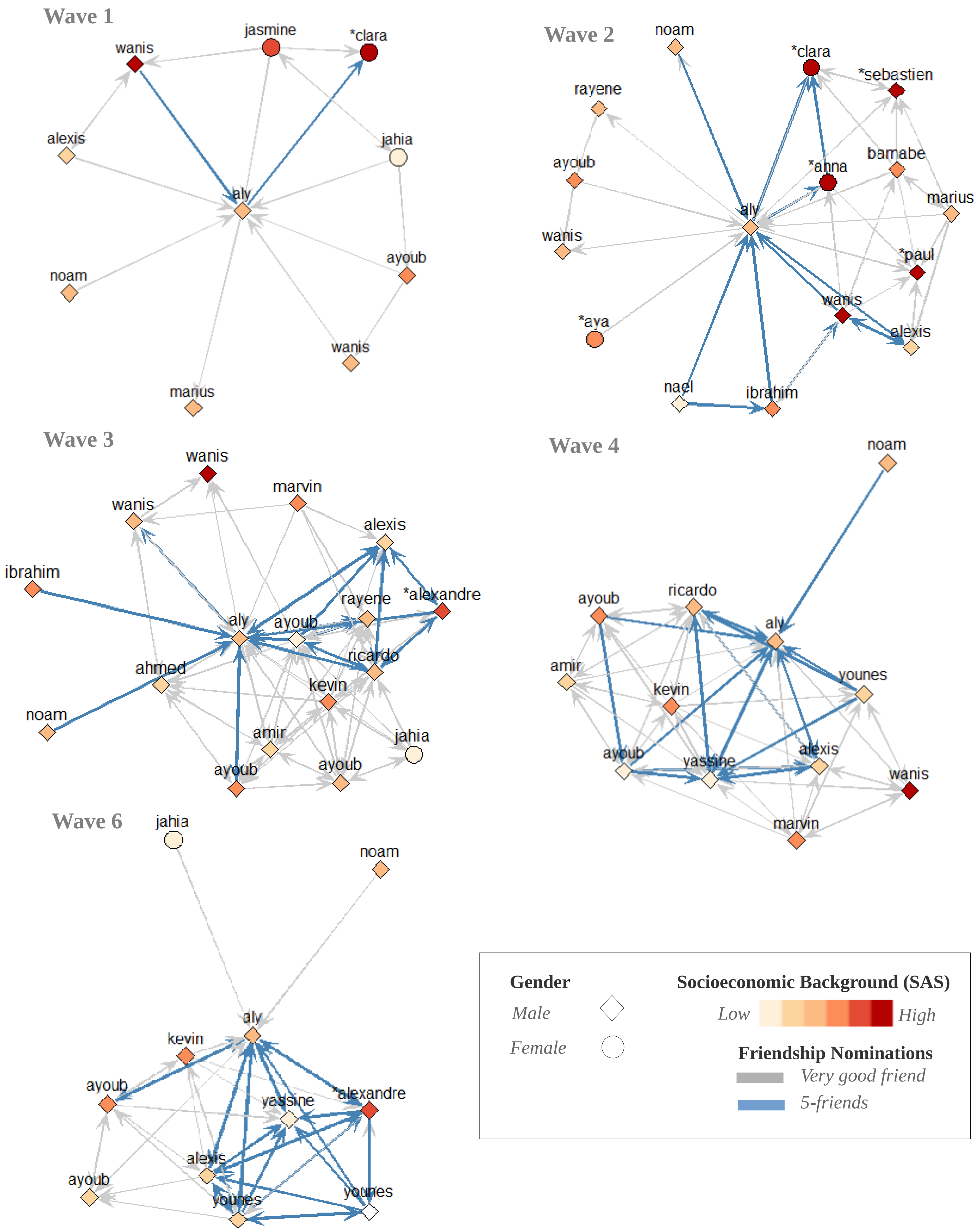
Thus, the episode of friendship with upper-class students from the international section, which seemed to matter so much to Aly in the interview of wave 2, and that created important tensions in his social circles, was in fact relatively short. Over time, he seems to

171 Also called personal network. It is a subset of the complete network that is defined from the perspective of a focal node, called *ego*. Only the other nodes to which ego is tied (called *alters*) are represented, as well as the ties among these alters.

have been drawn back to more expected friendships, with students from closer backgrounds. Note that these friends are not necessarily from a working-class background exclusively: there are many middle-class or even upper-middle-class students in his network (Aly himself being from a lower-middle-class background, rather than working-class strictly speaking). But they tend to live in Aly's neighborhood, are in the same section, or have a non-European background like him – all of which clearly set them apart from the upper-class, European-background and far-away living students of the international section. Typically, Yassine, one of Aly's closest friends in wave 4, is a new student who arrived in the school between waves 1 and 2. He moved into the same building as Aly, hence their encounter and ease in becoming friends (according to Aly's interview in wave 4). Moreover, in wave 4, Aly also became closer to friends that he met in primary school, but that he saw less during the first year of middle-school (notably Rayan and Ayoub, which he did not nominate as very good friends in wave 1). In his interview from wave 4, he simply does not mention students from the international section anymore, except for Alexandre, who is a special case among internationals (Aly actually did not name him as a friend in waves 1 and 2).

> *"A: Yassine, Alexis, Noam, hmm… in my class there is also… Alexandre, Jasmine, Chahima, and I think that's all. There may be more, but…*
>
> *Q: Ok, so these are the main ones. […] And so those that weren't in your class, did you know them last year already, or you met them this year?*
>
> *A: I knew them already. […] Yassine, I know… he was new at the end of first year, we talked a little bit, and we knew each other because he lives in my building. […] We were already friends, but we truly became friends this year.*
>
> *Q: […] And so your other friends, how did you meet them?*
>
> *A: Rayan it was in first year, we were in the same class. First, we did not talk much, but then we became friends. Ayoub I've known him since primary school. We did not talk much then, but in first year we truly became friends thanks to a video game.*
>
> *Q: Yeah. Fortnite?*
>
> *A: Yes. Everyone plays! (laugh) Rayan, I always have been very, very good friend with him ever since primary school. Hmm… Rayan [a different one], him I met him in… I became friends with him in second year. Ibrahim I've known him since primary school, we have always been friends."*
>
> **Aly, Paris 1, wave 4 (early third year)**

Figure 7-3: Ego-Networks of Aly ('Very good Friend' Nominations) – Waves 1 to 6, Paris 1



When asked about the disappearance of his friendships with students from the international section, Aly does not mention any dispute or clear split – in fact, he seems to still get along reasonably well with them. Instead, he explains that some of his friends from primary school were moved to his classroom (that of the internationals) in the second then third year, which facilitated them becoming close friends again. Retrospectively, he now explains his attempt to befriend the internationals as a consequence of his relative isolation in his first-year classroom, as he disliked the only two boys from his neighborhood that were in his classroom (both of whom harassed him together with Kevin at the time).

> *"Q: And so you're in the same class as the internationals?*
>
> *A: Yes, ever since the first year [of middle school].*
>
> *Q: Do you talk to them? You hang out with them?*
>
> *A: In first year, I hung out with them, but this year I don't. I stay with my friends. I talk to Alexandre as well. […] Sometimes I laugh with the internationals, I talk to them, but that's all. They're… my friends, but I don't hang out with them. I talk, I laugh with them, but… that's it.*
>
> *Q: […] And do you know why you stopped hanging out with them? You grew apart, a little?*
>
> *A: In first year, I mostly did it because I only had two friends from primary school [in his class], all the others I did not know them. All my friends were in other classrooms. So I started becoming friends with the internationals. In second year, I had more friends I knew. […] And it's the same this year;*
>
> *Q: Okay, and so who were your two friends in first year?*
>
> *A: Younes and Ayoub, but… they really made inappropriate jokes about me, and all. […] But it's okay now, we don't talk about it anymore, we made peace, now it's over.*
>
> *Q: So they bothered you a little. […] So your group now, and the internationals you were with in first year, these are not the same groups at all? Like, it's… it's not the same circles, let's say?*
>
> *A: […] Yeah, I'd say, like you said, me and my friends we don't spend too much time with the internationals. In fact, the internationals, they stay among themselves! And so we stay among ourselves. But sometimes we talk, and that's it.*
>
> *Q: […] Has it always been like that, or is it just this year?*
>
> *A: Yeah, in first year it was also like that. Since first year it's been like that. They prefer to stay among themselves, the internationals. […] Even so, I was with them! I was part, of… let's stay, the circle. But now I prefer to stay with my friends. Them, they always prefer to stay among them. That's how it is."*
>
> **Aly, Paris 1, wave 4 (early third year)**

In these extracts, Aly does not mention what was clear from the interview of wave 2, that is, the impact on his friendship choices of his high academic expectations (which he still holds in wave 4, though seemingly in a less salient way). Nevertheless, it also appears that the situation in wave 1/2 was largely due to chance, in the form of the quasi-random allocation of students across classrooms. As the classrooms got reshuffled, and as time passed, the “exception” of Aly’s friendship with the internationals progressively disappeared.

The second example is that of Jasmine, the child of a high school teacher. She has remained in the same classroom for four years, the same as the internationals and as Aly. An interesting thing about Jasmine is that, although she comes from an upper-class background[172] and has excellent grades herself, she has grown up in a dominantly working-class environment, attending the primary school of the neighborhood. Moreover, she has a non-European immigrant background, which is an important difference with many other upper-class students, notably the ones from the international section, and is on the contrary a common point with many working-class students of the school. Therefore, in terms of social origin, she stands somehow in-between the two main “blocks” of students, with some ground for homophilic tie formation with most of her peers depending on the traits that are considered.

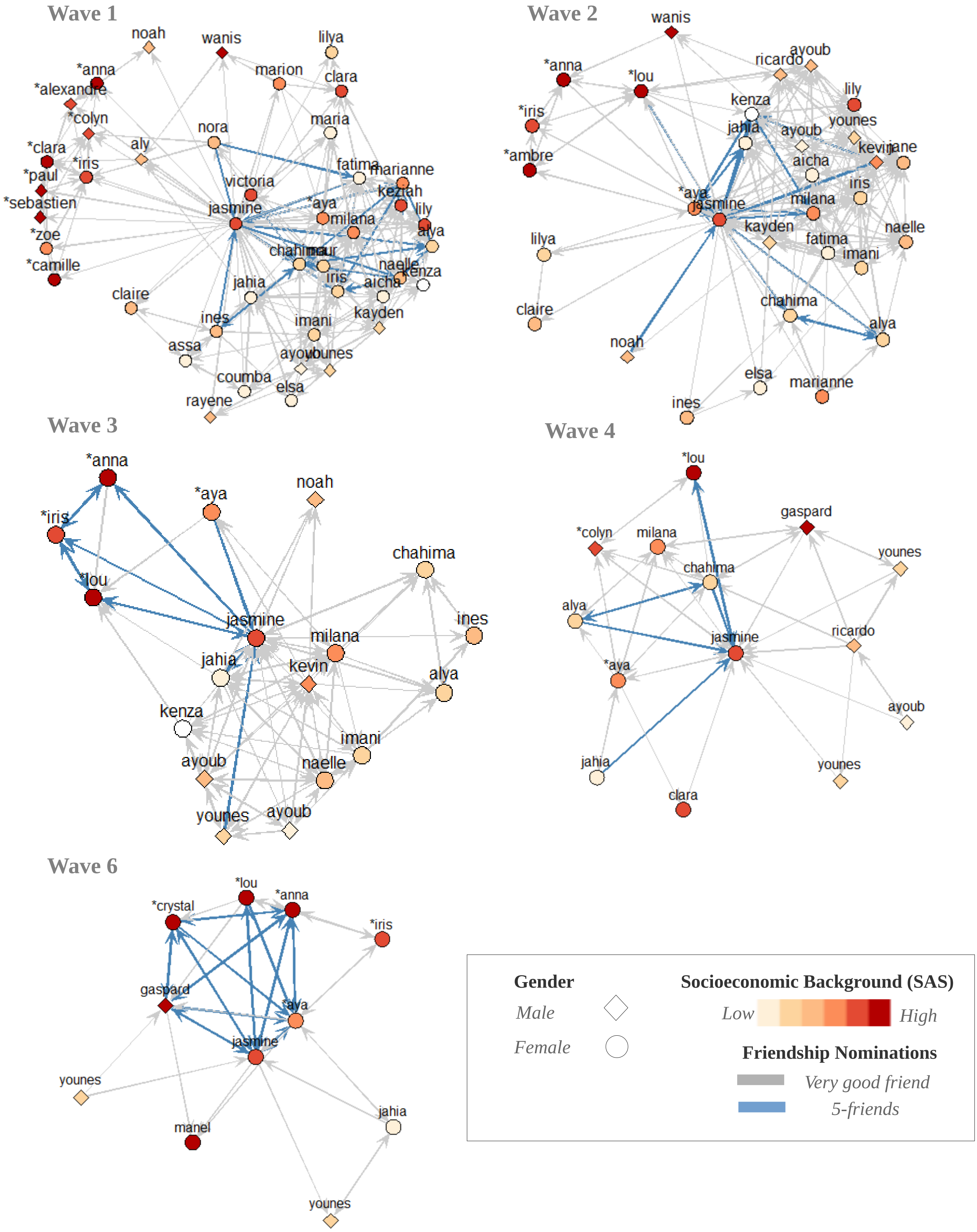
This may be one of the reasons why she is considered by many other students as one of the most popular girls of the grade-level, at least in the first two years of middle school. In her interviews, both from waves 2 and 4, she herself mentions the fact that she “talks to everyone”. In particular, this means being friends with both internationals and non-internationals, which she explicitly contrasts with the fact that most other students in her classroom are in one group or the other. Looking at her ego-networks in waves 1 and 2, she emits and receives many nominations indeed, and has a very diverse network in terms of socioeconomic (and ethnic) origins (Figure 7-4). However, in wave 1, all her closest friends, indicated by the 5-friends nominations (solid blue lines), are from the non-international section. Moreover, though she declares being very good friends with students from the international section, most of them do not reciprocate the nomination (the arrows point from Jasmine to them, but not the other way around).

Over time, the evolution of her network becomes clear: she gradually moves away from her working-class peers and toward her upper-class friends, particularly those from the international section. By wave 2, the international students start to reciprocate nominations toward her, and one of Jasmine’s closest friend becomes Lou, an international student who

172 Whether secondary school teachers should be considered middle-class or upper-class is debatable. In this case, as I constructed the scores of socioeconomic backgrounds based on the cultural capital of occupational groups (SA scores), it seems appropriate to classify teachers as upper-class, as their children are among those that have the highest school performances on average at the national level.

arrived in the school at the beginning of the second year. Wave after wave, her ego-network is then progressively reorganized around the very strong tie that she developed with Lou, even though she also keeps a few other friends. Another important trend is the overall decrease in the number of friends that Jasmine has, which is divided by more than three between waves 1 and 6. This is in part due to a general decrease in network density after wave 4; but it is also clear that Jasmine progressively lost the intermediary (and thus central) position that she had in the early waves, as her friendships with upper-class peers became more exclusive.

Figure 7-4: Ego-Networks of Jasmine ('Very good Friend' Nominations) – Waves 1 to 6, Paris 1



This evolution may have started as early as the first year of middle school, according to an interview conducted with two non-international girls, Iris and Kéziah. The interview takes place in wave 2 (early second year), at a time where the ego-network of Jasmine is still extremely mixed. According to Iris and Kéziah, Jasmine dresses in a way that is typical of upper-class tastes, and which resembles that of the international students (rock or pop style, which they call an "American" style)[173]. Moreover, they mention the fact that Milana, who has a middle-class background and comes from the same primary school, has started to experience the same changes. In the networks, Milana has a friendship trajectory somewhat similar to that of Jasmine in that she progressively moves away from the dense group of working-class and immigrant girls (blue zone in Figure 7-2). Unlike Jasmine however, she does not become friends with the internationals – as these are not in her classroom –, instead joining a small group of middle-class girls in the periphery of the blue zone (cf. Figure 7-2).

> *"A1: Milana, she changed. During the first year* [of middle school], *she wasn't like that.*
>
> *A2: She changed.*
>
> *A1: She became kind of… weird. Like Jasmine.*
>
> *A2: Yeah, exactly. […]*
>
> *Q: Jasmine and Milana? How did they change?*
>
> *A1: No, Jasmine, she has always been kind of strange, let's say. […]*
>
> *A2: I don't know how to explain.*
>
> *A1: I don't know, her attitude, her way of thinking, her way of dressing. She's totally… well, she is…*
>
> *A2: More in the American style, all that.*
>
> *A1: Yeah, she's into American things and all. […] She… she comes with these sweaters, you know, that stop before the belly button, and she has large pants pulled up until here.*
>
> *Q: Okay, so high-waist jeans.*
>
> *A1: Yes, and with big shoes.*
>
> *Q: Whereas you, you dress more…?*
>
> *A1: No, us it's normal. I mean, jeans, sneakers, a sweater and a puffa jacket* ["doudoune" in French, often indicative of a rap/street style].
>
> *Q: So Jasmine, she has always been like this?*
>
> *A1: Yeah.*
>
> *A2: No, Jasmine in primary school she wasn't like that. She changed. I've known her since kindergarten, she changed you know.*
>
> *A1: And then, Milana, at the start of the year, she was… well, jeans, sneakers and all* [here she refers to Jasmine's style of jeans and sneakers, not her own].
>
> *A2: Since they became best friends, Jasmine contaminated her.*

173 A large part of this interview extract has also been presented in chapter 6, section 3.1., to illustrate the rock/rap distinction in cultural styles.

A1: She became her best friend and, well, I don't know, she starts to become like that. […]

Q: Are they both in your classroom?

A1: No. Jasmine she's with the internationals, in class C, and Milana in class B [Iris and Kéziah are in class A].

A2: No but they dress the same, really. Well, they hang out with the internationals."

Iris and Kéziah, Paris 1, wave 2 (early second year)

Therefore, it appears that the progressive distancing of Jasmine from the working-class girls that were in her primary school was latent from the beginning of middle school, as their tastes and dressing styles started to diverge. Looking at the socioeconomic background of Jasmine, or even of her friend Milana, this seems relatively logical (in that regard, the exception would rather be Kéziah who, although from an upper-class background herself, remains in the group of working-class girls and has "lowbrow" tastes in terms of culture and dressing). It is possible that this increasing divergence is a consequence of students growing older, with fashion and music tastes becoming increasingly important in terms of social identification and group belonging as they enter adolescence (cf. chapter 6 section 3.3). Nevertheless, in the case of Jasmine, the role of the students from the international section also seems quite obvious, offering her new friendship opportunities with students with whom she shares certain tastes, and most likely contributing to developing these tastes through peer-to-peer socialization. In other words – and keeping in mind that this is a single case and should therefore be considered with care – it might be that students from the international section acted as a sort of catalyst for Jasmine's upper-class dispositions, accelerating or even triggering the gradual split from her different-background childhood friends.

1.2. Processes of Change

Having a better idea of what happened in Paris 1's networks, we may now try to understand how it happened. What are the relational processes through which this evolution came about? Did former heterophilic friends get into a dispute, or simply stop talking to one another? Was the re-organization of the network progressive, or did it happen all at once? Did it primarily occur through the disappearance of heterophilic friendships, or through the formation of new homophilic ones? Finally, did students become more homophilic in their friendship choices over time – that is, was there been an increase in homophilic selection? I first examine the difference in homophily patterns between newly formed ties and dissolved ones (1.2.1.), then look for changes in the low-order processes inferred from the generative models of chapter 5 (1.2.2).

1.2.1. Formation and Dissolution of Ties

Change in a network occurs in two ways: some ties disappear, and others are created. For parsimony's sake, both processes have been modeled jointly so far, notably in the models of chapter 5. However, they may in fact follow different principles: the factors that determine the initial apparition of a friendship are not necessarily the same as those that determine its long-term survival. In that regard, socioeconomic homophily may primarily operate at the level of friendship formation, or, on the contrary, through differential rates of survival for established relationships. In the first case, we would conclude that the increase in homophily in Paris 1 was primarily due to students reaching out to same-background peers, whereas in the second, it would point to the disappearance of heterophilic friendships. This may constitute an important hint as to how antagonistic the network evolution was: if students from different backgrounds started to get into fights or clash with one another as they grew up, one would expect dissolution processes to exhibit socioeconomic homophily (note that the opposite is not true: the fact that there has been homophily in friendship dissolution does not necessarily imply that this dissolution process has been antagonistic).

To answer this question, I resort to a variation of the ERGM, called STERGM (*Separable Temporal Exponential Random Graph Model*; Krivitsky and Handcock 2014). STERGM functions essentially like ERGM, with an optimization algorithm that looks for the set of parameters that jointly predict the correct value of network-level target statistics. Just like ERGM, it can be interpreted as a logistic regression performed at the tie level – and is in fact equivalent to such a regression for models with only dyad-independent terms. However, whereas ERGM models a network observed at a single time point, STERGM models the *transition matrices* that link two time points. Indeed, for two successive observations of a network – for example in waves 1 and 2 of the questionnaires – one can define a *formation matrix*, made up of all the ties that appeared between the two points, and a *dissolution matrix*, made up of all the ties that disappeared. Together, the formation and dissolution matrices capture the change between the two observations of the network. What STERGM does is simply to fit an ERGM model to these two matrices separately; one therefore specifies a formation and a dissolution model (which can take different parameters), and the two are estimated jointly[174].

I fit a simple dyad-independent STERGM for descriptive purposes only, similar to the ERGMs of chapter 4. Four terms are included: one for network density (baseline probability of any tie), one for socioeconomic homophily (log-odd change for one additional point of the absolute difference in Socio-Academic Scores between two nodes), and two for socioeconomic emission and reception (log-odd change for one additional point of the Socio-Academic Score of the emitter/receiver node). Each transition is modeled separately, in the network of "very good friend" nominations. Results are shown in Table 7-2.

Table 7-2: Socioeconomic Homophily in Tie Formation and Dissolution (STERGM) – 'Very Good Friend' Nominations, Paris 1

	w1 → w2	w2 → w3	w3 → w4	w4 → w6
Formation – SAS absolute difference	-0.32*** (0.09)	-0.37*** (0.09)	-0.65*** (0.10)	-0.68*** (0.11)
Dissolution – SAS absolute difference	-0.02 (0.15)	-0.49*** (0.12)	-0.20 (0.13)	-0.23 (0.15)

***: p<0.01; **: p<0.05; *: p<0.1

Note: the coefficients correspond to the model term of SAS homophily, which is the absolute difference in SA scores between the two nodes of the dyad. **For dissolution, the model considers the odds for a tie *not* to be dissolved, i.e. its odds of persisting across waves.** Therefore, like for formation, homophily is indicated by a negative coefficient: as the SAS difference increases, the odds that a tie persists decrease, meaning that the odds of dissolution increase.

Unshown model terms include the edge parameter and emission and reception effects for SAS. See Appendix 7A for the complete outputs.

Clearly, socioeconomic homophily primarily concerns tie formation: all four coefficients are large and statistically significant. There also seems to be an effect for tie dissolution, but it mostly appears in the second transition. There might be an effect as well in the third and fourth transitions, but it is not statistically significant[175] and, in any case, would be of a much smaller magnitude. Therefore, the prime reason for which socioeconomic

174 Note that this is quite different to SAOM's take on longitudinal modeling, because SAOM explicitly models the evolution of the network over time. By contrast, STERGM considers the formation and dissolution matrices almost as if they were cross-sectional data (the "trick" being to subsume the information about two time points into a single transition matrix). This comes down to the difference between SAOM and ERGM in general (including longitudinal extensions of ERGM), which is that SAOM models an evolution process whereas ERGM models the state of networks observed at certain points in time. See REFs for a discussion.

175 Separating the four transitions has the mechanical consequence of reducing statistical power. If we modeled all four transitions at once, we would find a significant effect for dissolution as well, albeit smaller than for formation (model not shown).

homophily went up over time in Paris 1 is because newly formed ties were homophilic; but heterophilic ties did not disappear at a much higher rate than homophilic ones.

Remember that this is a descriptive statement: it does not imply that homophilic *selection* operates primarily at the level of tie formation[176] (and, as we saw in chapter 5, most socioeconomic homophily is not explained by homophilic selection anyway). In any case, the fact that socioeconomic homophily mostly unveils at the level of tie formation suggests that there has not been massive fights or break-ups between friends from different social backgrounds after wave 1[177]. These friendships have disappeared at a relatively "normal" rate (i.e. not much different than that of homophilic friendships), but they have been replaced by mostly homophilic ties.

1.2.2. Changes in Low-Order Relational Processes

Let us quickly review what we have seen so far. The global increase in socioeconomic homophily over time is due in large part to its increase among non-internationals: in particular, a large, mixed group of girls in wave 1 which splits into more homogeneous subgroups. Moreover, the group of internationals also contributes to this global increase by losing some ties to lower-background non-internationals (such as Aly) and replacing them with ties toward higher-background non-internationals (such as Jasmine). For the most part, this did not entail a particularly high rate of friendship dissolution among socioeconomically dissimilar students, but newly formed ties were strongly homophilic. The question we may now ask is whether these network-level trends signal a change in the students' behavior. We saw in chapter 5 that Paris 1 is the only school where there is strong evidence for same-background selection among students. Did this selection increase, that is to say, does socioeconomic origin become an increasingly important factor for friendship selection over time? Alternatively, do the other relational processes that indirectly contribute to socioeconomic homophily change as well?

176 In principle, one could use SAOMs to distinguish between selection in tie creation and endowment (i.e. formation and dissolution). The problem is that splitting the effect of SAS similarity in two, when done within complex models such as those of chapter 5, weakens their statistical power, which in this case leaves both coefficients non-significant (model not shown).

177 Except for the possibility that these fights did occur but were counter-balanced by other processes that made the dissolution of socioeconomically homophilic friendships more likely. In other words, certain effects would act as suppressors that mask the impact of socioeconomic selection on friendship dissolution, making homophily in dissolution appear to be almost null. This is a bit far-fetched though and is not backed-up by results obtained from more complex STERGMs that include controls for gender, classroom and structural effects (not shown).

One way to test for this is to consider *time heterogeneity* in the tie-formation processes modeled by SAOM (*Stochastic Actor-Oriented Model*). The baseline assumption made by SAOM is that the value of the parameters is the same at each transition: the “rules” used for predicting the evolution of the network between waves 1 and 2 are the same for predicting its evolution between waves 2 and 3 or 3 and 4[178]. This comes from the fact that the target statistics are summed up over all transitions: for instance, the parameter for SAS similarity is the sum of the SAS similarity score of all the ties observed from waves 2 to 4. As a result, a converged model only guarantees that the simulated value of the statistic is correct over this entire sum; but it is entirely possible that the distribution of the statistic at a specific transition does not correspond to its observed value. *Time heterogeneity* is the idea that this assumption of homogeneous processes is not correct and leads to erroneous predictions of the values of the target statistics at specific waves.

A straightforward way to test for this is to include *time dummies*, that is, interaction effects between the parameters of interest and time. If the parameter of the time dummy is statistically significant, then one would conclude that there is time heterogeneity in the process of interest. However, fitting a model with all possible time dummies for all parameters is in practice impossible. Instead, one can resort to the time heterogeneity test proposed by Lospinoso et al. (2011) and implemented in the SIENA software. Based on an already estimated model, this test can check for time heterogeneity for one or several model parameters, making use of the networks and target statistics already simulated during the estimation process. If the test returns time heterogeneity, the user should then add appropriate time dummies to the specification and re-launch the estimation with those, until a model without heterogeneity is reached.

This was in fact already mentioned in chapter 5, and the models presented in Table 5-3 have been specified after accounting for time heterogeneity. As I mentioned then, the only parameter in ‘SAOM 1’ (the main estimated model) for which there was time heterogeneity in Paris 1 was the density effect (i.e. the baseline odds of any tie occurring). In other words, time heterogeneity tests do not conclude in any statistically significant change in the tie-formation processes captured by the model (except for baseline density)[179].

178 The only exception is the rate function, which always takes one parameter per transition (Ripley and Snijders 2020).

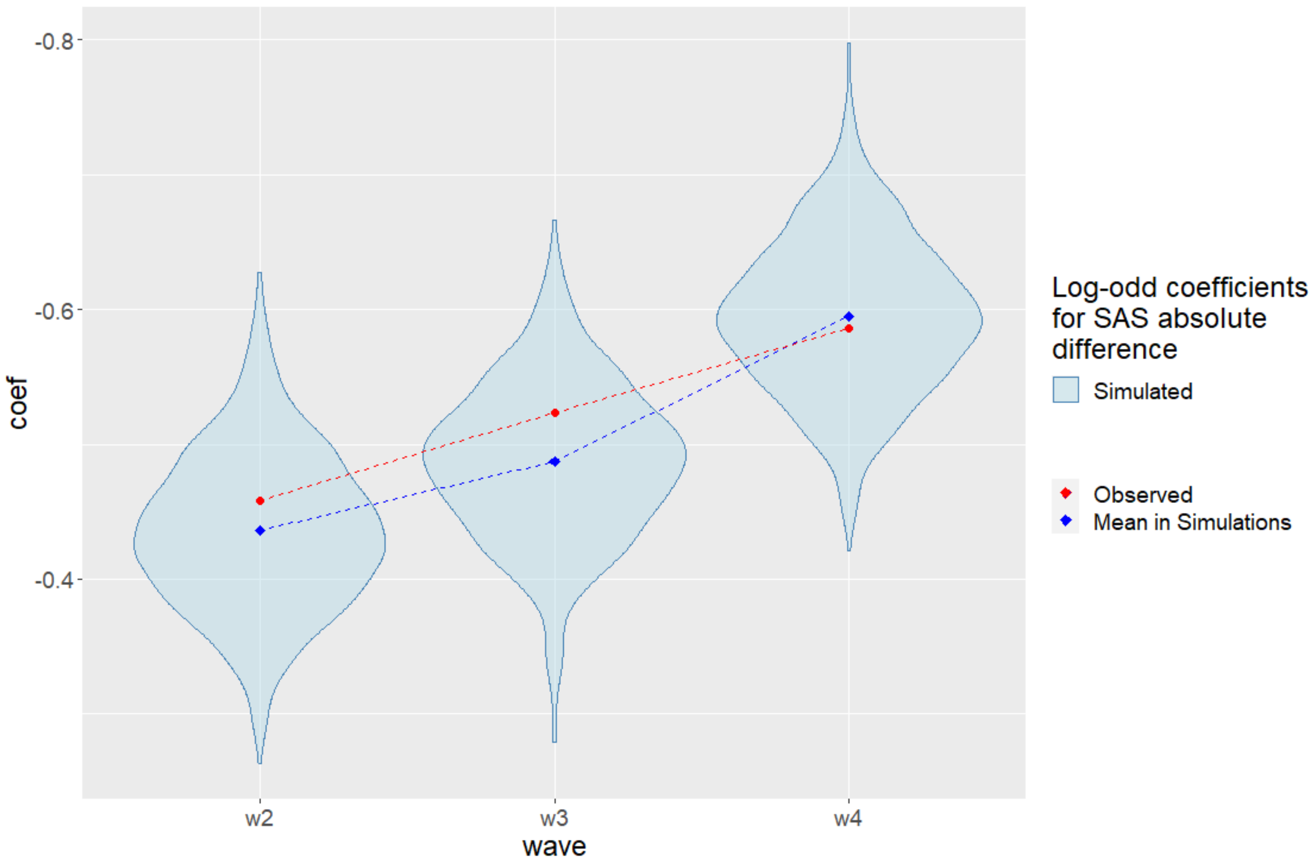
179 To confirm this, I also re-estimated the model with the inclusion of time dummy terms for the SAS homophily parameter. Coherent with what the time heterogeneity tests suggested, these parameters are small and statistically insignificant (models not shown).

To get a better sense of what this means, we can look at the level of socioeconomic homophily predicted by the model at each time point. If it is the case that socioeconomic homophily increased because some processes changed over time, then our current model – which postulates that processes are homogeneous over time – should over-estimate socioeconomic homophily in the early waves, and under-estimate it in the late ones. This is shown in Figure 7-5: 1000 networks have been simulated from the model "SAOM 1" and, for each simulated network, the log-odd coefficient indicative of SAS homophily is computed[180]. The blue shaded array represents the distributions of the simulated values; the blue points indicate the mean of this distribution; and the red points indicate the value observed in the empirical networks. Note that there is no simulated network for wave 1, since SAOM models network evolution from that first point onward.

Additionally, similar tests are performed for the transitions up to wave 6, not just wave 4. Here again, there seems to be no time heterogeneity. See Appendix 5D (Table 2).

180 Using an ERGM similar to those of chapter 4, i.e. with four parameters: density, SAS homophily, SAS emission and SAS reception.

Figure 7-5: Socioeconomic Homophily in the Networks Simulated from the Model 'SAOM 1' (Paris 1)



Note: The log-odd coefficients are taken from an ERGM model applied to each simulated networks, with four parameters (density, SAS homophily, SAS reception and SAS emission). The y-axis has an inverted scale, so that lower coefficients (indicative of more homophily) are placed upward. There are 1000 simulated networks for each wave.
The simulation process of SAOM works such that each simulated wave comes from applying the vector of model parameters to the *observed* network of the previous wave. For instance, to predict the network state in wave 3, the model simulates an evolution process from the observed network of wave 2, not from the fictive wavz 2 networks that have been predicted from the observed network of wave 1 (in other words, simulations are conducted separately period by period, they are not "chained" over several periods).

One can see that the mean value of socioeconomic homophily in the simulated networks is close to its observed value at each wave. There is a slight under-estimation of socioeconomic homophily in waves 2 and 3, but this remains well within the confidence interval. More importantly, this pattern is actually opposed to what one would expect if homophily-inducing processes had become stronger over time: it should result in an over-estimation of socioeconomic homophily in the early wave, whereas we see here that the model (slightly) under-estimates it. Altogether then, this confirms that the low-order processes driving the network evolution are roughly constant over time – or, more precisely, that we cannot reject the null hypothesis that they are constant[181].

This suggests that the increase in socioeconomic homophily observed at the aggregate level results from the fact that the impact of homophily-inducing processes is cumulative over waves: at each transition, the same processes shift the network structure a step further toward strong socioeconomic homophily. In other words, the main difference between waves 2 and 6 seems to be that homophily-inducing processes did not have the same time to unveil – which conversely suggests that the networks from the early waves had not yet reached an equilibrium in regard to socioeconomic homophily[182].

In a sense then, one could say that the observed increase in homophily was already latent in the early waves. This explanation raises another question: why was socioeconomic homophily still relatively low in wave 1? Indeed, if relational processes did not change much over time, how is it that they produced a less homophilic network structure in the past? There are three possible answers, which are not exclusive from one another. First, relational processes could have been different during the first year of middle school, prior to wave 1 (although the remarkable stability of these processes from wave 1 onward arguably goes against this idea, at least in the absence of additional information). Second, at the beginning of middle school, there may have been a strong random element to friendship formation, as many students did not know one another. Assuming some form of path dependency in existing relationships, this partially random start may have introduced a number heterophilic ties in the

181 One possible objection may be that the model used here (SAOM 1) postulates a homogeneous effect of socioeconomic similarity over all dyads. However, the findings from section 1.1. above suggest that socioeconomic selection may not be the same among non-international students and for mixed international/non-international dyads (section 1.1 showed that socioeconomic homophily differs between these dyads, so selection might as well). In this case, one could imagine that socioeconomic selection is relatively stable at the aggregate level, but that it increased among non-internationals while decreasing among mixed dyads, thus the lack of change for the aggregate effect. This is tested for in Appendix 7D: the interpretation of the fitted model is uncertain, because there are arguably over-fitting issues, but the best hint we can get is that both forms of socioeconomic selection (for non-internationals and for mixed dyads) are relatively stable over time. In fact, most socioeconomic selection seems to operate in mixed dyads overall, which further supports the idea that the presence of the international section is decisive in the emergence of socioeconomic homophily.

182 The notion of equilibrium implies that the low-order processes driving the network evolution are such that the aggregate features of the network are stable in time. A common misconception when thinking about an equilibrium is to think that it will occur under a random allocation of ties: for instance, socioeconomic homophily would be stable in time if were no socioeconomic homophilic selection. However, if low-order processes were "blind" to socioeconomic origin, then this would quickly lead the network toward a null homophily. For socioeconomic homophily to be stable in time, there needs to be homophily-inducing processes that are active in maintaining the current level of homophily, by compensating for its natural "erosion" due to the random re-allocation of certain ties. In that regard, the fact that socioeconomic homophily increases over time in Paris 1 shows that homophily-inducing processes are stronger than this, i.e. they entail more homophily than what would be needed to simply maintain a stable level.

network that then persisted for some time (this is typically what the example of Aly suggests). Finally, the friendships inherited from primary school may have been less homophilic, either because of different relational processes back in primary school (due perhaps to the large age gap and to the drastic change of context compared to middle school), or because the structure of opportunities leaned toward more heterophilic friendships (which is what may have happened in the case of Jasmine, who had a majority of working-class peers in her primary school)[183]. This conjecture is supported by the fact that, in wave 1's network, friendships between students that were in the same primary school are less homophilic in socioeconomic terms than those between students that came from different schools (see Appendix 7B). However, this should still be considered with caution, since we do not directly observe the friendship networks as they were during primary school.

In any case, the increase in socioeconomic homophily over time in Paris 1 appears as a logical consequence of the fact that, even from the early waves, the situation was extremely favorable to the formation of socioeconomically homophilic friendships. Yet again, this opens up a new question: why was this initial situation different in Paris 1 compared to the other schools? This is the subject of the next section.

2. Properties of Schools and Relation to Socioeconomic Homophily

As mentioned in the introduction of the chapter, testing the relationship between aggregate school features and the structure of friendship networks in a systematic way would demand a larger sample of schools. However, it is still possible to see in the studied schools if the alleged processes that should link certain aggregate features to specific levels of homophily are coherent with the data at hand. In order to do this, I first examine the case of Savoie 2, the least homophilic out of the four schools (2.1.). Then, I come back to the case of Paris 1, by considering the school features that may have fostered the situation described in the first part of the chapter (2.2.). Finally, I discuss the two main criteria that were used for the selection of the schools in the research design: their location (urban/rural) and sector (public/private) (2.3.).

183 Some weak primary school friendships can also become stronger at the beginning of middle school, as students fall back on the few peers that they know, potentially over-shadowing the socioeconomic homophily that existed among these primary school friendships.

2.1. Savoie 2: a Favorable Context to Socioeconomic Mixing?

2.1.1. Factors Correlated to Socioeconomic Background and Selection Bias in the School Population

Among the four schools of the sample, Savoie 2 appears as an exceptional case, exhibiting a remarkably low level of socioeconomic homophily. It is so low, in fact, that it does not appear in the evolution of the 'very good friend' network between waves 1 and 4, making the use of contribution scores to assess the impact of various processes on this homophily simply meaningless. With that said, there is still a little bit of socioeconomic homophily in the school, concentrated on the strongest friendships: the nominations of the five friends with which one spends the most time at school, as well as that of the friends seen outside, show some discernible homophily at each wave (cf. chapter 4 section 1, as well as Appendix 4C).

Why, then, is socioeconomic homophily so low? To begin with, results from other schools suggest that the lack of homophilic selection based on socioeconomic background is not, in itself, particularly surprising. Interviews revealed that students are not directly aware of their peers' backgrounds (chapter 6), and SAOMs suggested an absence or near absence of direct selection in Savoie 1 as well (chapter 5). In Paris 2, the evidence for such selection remains mixed – it is very apparent at one transition but not so much at the others – and, in any case, its magnitude ought to be modest. Rather, what really distinguishes Savoie 2 from the other schools is that, in the latter, various processes indirectly induce socioeconomic homophily. Residential segregation, academic performance, ethnic background, cultural tastes, or parental networks correlate to socioeconomic background, such that tie-formation processes associated to these elements tend to generate small or modest amounts of socioeconomic homophily. These get further amplified by transitivity processes, resulting in clear – though by no mean enormous – homophily at the aggregate level. The question, then, is why this did not occur in Savoie 2.

A straightforward explanation is that the factors that are correlated to socioeconomic background in the other schools, are not correlated to socioeconomic background (or less so) in Savoie 2. This is confirmed by Table 7-3, which shows the Pearson's coefficient of correlation for nodal and dyadic attributes in the four schools.

Table 7-3: Coefficient of Correlation of Various Nodal and Dyadic Attributes with SAS

Scores (Nodal) and SAS Absolute Differences (Dyadic)

	Nodal Attributes				Dyadic Attributes (matching or absolute difference)			
	Paris 1	Paris 2	Savoie 1	Savoie 2	Paris 1	Paris 2	Savoie 1	Savoie 2
Grades	0.53	0.20	0.45	0.33	0.20	0.03	0.13	0.08
Cultural status	0.65	0.32	0.41	0.35	0.40	0.11	0.14	0.10
Agreement (music)	-	-	-	-	0.13	0.11	0.02	0.02
Ethnic background	0.58	0.13	0.25	0.18	0.13	0.02	0.06	0.10
Classroom	-	-	-	-	0.08	0.03	0.03	0.01
Primary school	0.51	-0.17	0.43	0.15	0.16	0.00	0.04	0.00
Residential distance	-	-	-	-	0.22	0.02	0.02	-0.03
Parental ties	-	-	-	-	0.10	0.02	0.04	0.01

Note: for matching dyadic attributes (agreement, ethnic background, classroom, primary school and parental ties), coefficients have been multiplied by -1 so that a positive coefficient indicates a correlation of the two types of homophily (because for continuous attributes, the homophily statistic is larger for less similar students, whereas it is the opposition for categorical attributes).
For categorical nodal attributes (ethnic background and primary school), the value given is the square-root of the R-squared in a bivariate regression on socioeconomic background (since Pearson's R can only be computed between continuous variables).

Coefficients consistently appear lower at the dyadic level than for their nodal counterpart, which is a formal implication of converting individual scores into matrices of absolute differences.

In Table 7-3, we can see that, for instance, the correlation between students' socioeconomic origins and academic results is of 0.53 in Paris 1 and 0.45 in Savoie 1, but only 0.33 in Savoie 2 (and even less in Paris 2, but this is something I will come back to). This might not seem like a large difference but, moving from the level of individuals to that of the dyad, it becomes more pronounced: the pairwise absolute difference of SA scores and of academic achievement have a correlation of 0.20 in Paris 1, but only 0.08 in Savoie 2[184]. Similar patterns can be observed for most other variables and covariates. For residential distance, there is even a slight negative correlation in Savoie 2, meaning that different-background students even live slightly closer on average (the effect is extremely small and non-significant, though)!

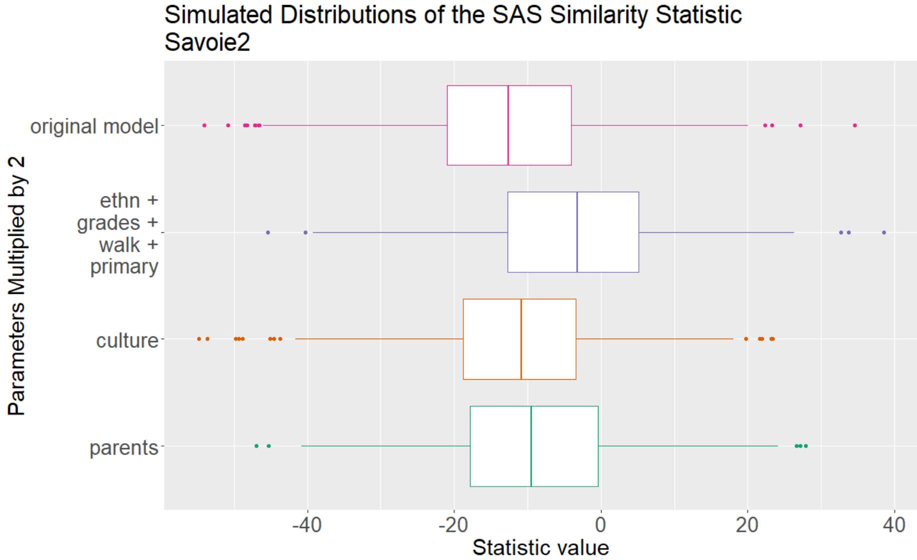
As a result, no matter how strong relational processes such as academic selection, ethnic selection or residential propinquity may be, they will induce extremely little socioeconomic homophily. To confirm that this is the case, I resort to fictive counter-factual

184 Coefficients consistently appear lower at the dyadic level than for their nodal counterpart, which is a formal implication of converting individual scores into matrices of absolute differences.

scenarios, much like what was done previously for contribution scores. However, this time, instead of setting some parameters to 0, I double the value of the parameters that may be expected to induce socioeconomic homophily (the original value of these parameters is given as a reminder in Table 7-4 below). In other words, this is a way of exploring the sensitivity of socioeconomic homophily to these low-order processes, according to the model. Results are shown in Figure 7-6. Three counter-factual scenarios are tested: one, based on Model 1, with the parameters of academic similarity, ethnic matching, residential distance and primary school matching doubled; a second, based on Model 2, with the parameters of cultural similarity doubled (both cultural agreement and cultural status); and a third, based on Model 3, with the parameter of parental ties doubled. For each scenario, the outdegree parameter is adjusted so that the simulated networks exhibit the correct density (again, just like in chapter 5). In all three cases, the simulated networks hardly exhibit any socioeconomic homophily, much like the observed ones, confirming that these relational processes simply cannot induce socioeconomic homophily given the correlational structure of students' attributes[185].

185 Another merit of this approach is to (indirectly) test for the idea of suppressors of socioeconomic homophily. Indeed, in theory, it is possible that some relational processes do induce socioeconomic homophily, but that their impact is counter-balanced by other relational processes, which act as suppressors in some way and prevent the emergence of socioeconomic homophily. If this is the case, then by increasing the effect size of the potential homophily-inducing effects, the balance between them and potential suppressors should be No.roken, and the simulated networks should exhibit some socioeconomic homophily. As we just saw, this is not the case.

Figure 7-6: Sensitivity Tests for Socioeconomic Homophily – Savoie 2



Note: the statistics are the values of the SAS similarity statistic (sum of the absolute differences of SAS scores over all ties, rescaled on (0,1) and subtracted from 1), observed in 1000 simulated networks. Values inferior to 0 signal that the statistic has decreased compared to wave 1 (SAOM considers the evolution of the network from wave 1 onward).
"Original" refer to the complete model with all parameters set to their estimated values (i.e. Model 1 in Table 7-5 below). The first scenario ('ethn + grades + walk + primary') is based on Model 1, the second ('culture') on Model 2 and the third ('parents') on Model 3. In each case, the parameters of interest are multiplied by two, the outdegree parameters is adjusted, then networks are simulated.

Table 7-4: SAOM Parameters Estimated in Savoie 2 (reminder from Table 5-3)

Effect Name	Model 1	Model 2	Model 3
transitivity (gwesp FF)	1.36*** (0.04)	1.36*** (0.04)	1.33*** (0.04)
SAS similarity	-0.05 (0.07)	-0.07 (0.07)	-0.08 (0.07)
same ethnicity	0.19*** (0.05)	0.18*** (0.05)	0.16*** (0.05)
grades similarity	0.09 (0.1)	0.09 (0.09)	0.08 (0.09)
same sex	0.33*** (0.03)	0.32*** (0.03)	0.31*** (0.03)
same classroom – current year	1.09*** (0.04)	1.09*** (0.04)	1.08*** (0.04)
same classroom – previous year	0.1*** (0.03)	0.1*** (0.03)	0.13*** (0.03)
same primary school	0.17*** (0.04)	0.16*** (0.04)	0.03 (0.04)
residential distance	-0.05 (0.03)	-0.05 (0.03)	-0.03 (0.03)
residential distance – squared	0.01** (0.004)	0.01** (0.004)	0 (0.005)
matching opinions – music	-	0.09*** (0.02)	-
matching opinions – video makers	-	-0.02 (0.02)	-
cultural status similarity	-	0.21*** (0.07)	-
parental ties (w1)	-	-	0.66*** (0.05)

***: p<0.01; **: p<0.05; *: p<0.1

Note: Standard errors of the coefficients are between parentheses. See Appendix 5B for the complete models outputs.

Why is socioeconomic background not associated in this school to the same properties as it is in the other schools (and at the national level)? To an extent, it may come from differences in the school policy and internal organization, at least for the traits that are achieved rather than ascribed. In particular, the low association of academic achievement and socioeconomic background may be seen as the result of a different pedagogy and better support of students from lower backgrounds. For instance, the fact that there are no numerical grades in Savoie 2 may contribute to this (incidentally, this may also explain why there is no evidence for academic selection in this school). Similarly, the weak association of cultural tastes and socioeconomic background may be an endogenous outcome of students' sociability: more mixed friendship groups should result in more homogeneously shared cultural tastes, to the extent that these are impacted by peer-to-peer socialization.

Nevertheless, in the case of academic performance, the literature in the sociology of education tends to agree that family socialization has a decisive role in explaining social inequalities in that regard, with the in-school context acting as an important, but probably secondary determination (Duru-Bellat 2015). The weak correlation of academic results and socioeconomic origin thus probably has to do with a specificity of students' households, not just of the school's organization or pedagogy. Moreover, for ascribed characteristics such as primary schools and ethnic background, the lack of correlation to socioeconomic origin cannot be explained by in-school socialization whatsoever.

Therefore, the most likely explanation is that there is a selection bias in the population of Savoie 2: the household that sends its children there do not exhibit the same properties as other households from similar occupational groups elsewhere in the country. Being a private school, Savoie 2 can attract families regardless of recruitment sectors, and freely choose among applicants if they are more numerous than the available slots. Although I lack explicit data on this, informal discussions with some teachers and students suggested that many parents chose this school to avoid the local public school, which has a bad reputation. Thus, the working-class households that apply (and are accepted) to Savoie 2 are probably particularly concerned about their children's school success, and/or are looking for some form of differentiation or distance from their working-class environment. Moreover, the school direction seems to be relatively demanding with students in terms of their implication in their work and disciplinary attitudes, with threats of exclusion for misbehaving students (all the points mentioned in this paragraph have been presented already in chapter 3, section 2.3.4).

In chapter 6 (section 3), I also mentioned the fact that cultural frames of perception – music and clothing – were less frequently mentioned in the interviews of Savoie 2 compared to the other schools. This may suggest that there is more cultural homogeneity in this school, thus the fact that it does not appear as an important means of distinction to most students. This is confirmed by the quantitative data: the mean of the scores of cultural status – a measure that distinguishes between lowbrow and highbrow tastes – is almost the same in the four schools, but its standard deviation is lower in Savoie 2[186]. Also in chapter 6, we saw that school attitudes were also more homogeneous in Savoie 2, with a greater adhesion to school norms overall (section 2.3.3). Altogether, this may account for the fact that Savoie 2's socioeconomic diversity appears to be, in a way, less diverse than in other schools.

2.1.2. Comparison with Paris 2

It can be seen on Table 7-5 that the association of various covariates with socioeconomic homophily is in fact not so different between Savoie 2 and Paris 2. This seems logical, given that Paris 2 is also a private school, thus potentially subject to selection biases similar to those of Savoie 2. Yet, as we saw in previous chapters, there is a moderate amount of socioeconomic homophily in Paris 2, unlike in Savoie 2. To understand why that is the case, one can consider the factors that contribute to socioeconomic homophily in Paris 2, based on the contribution scores from chapter 5. These contribution scores are presented again in Table 7-5 below.

First, ethnic background, despite its low correlation to socioeconomic background, does contribute a little to socioeconomic homophily in Paris 2 (which also depends on the details of the correlational structure of the different covariates, not just of the pairwise correlation between two covariates – thus the need for contribution scores in the first place). Second, cultural agreement is more correlated to socioeconomic similarity in Paris 2 than Savoie 2 (Table 7-5) and contributes quite notably to socioeconomic homophily in Paris 2 (contribution score of -0.20). Third, there seems to be a little bit of socioeconomic selection in Paris 2 (though it strongly depends on the waves), unlike in Savoie 2. Finally, both the density and transitivity of friendship networks are high in Paris 2 (Appendix 3B), so the aggravating effect of transitivity should be particularly strong here compared to other schools (which is indeed what the high contribution score of -0.65 suggests).

186 The means are 2.28 (Paris 1), 2.22 (Paris 2), 2.32 (Savoie 1) and 2.29 (Savoie 2). The standard deviations are, in this same order, 1.29, 1.06, 1.06 and 0.91.

To sum up then, small differences in the network structure and attribute distribution do imply an overall less favorable context for socioeconomic mixing in Paris 2 than in Savoie 2, though it probably still remains more favorable than in the public schools. In that regard, it is worth noting that contribution scores, thanks to the simulation process derived from SAOM, account for a number of features of the network structure and attribute distributions that are not visible from the simple coefficients of correlation of Table 7-3 (which is the whole reason for using contribution scores in the first place!).

Table 7-5: Contribution Scores in Paris 2 (reminder from Table 5-4)

Scenario Name	trans	SAS	ethn	grades	sex	class-room	resid	cult	parents
Parameter(s) set to 0	transitivity	SAS similarity	same ethnicity	grades similarity	same sex	same current and past classroom	same primary school + residential distance + resid. dist. squared	cultural status similarity + music agreement + youtube agreement	parental ties on wave 1
Predicted % change in SAS homophily	-0.65	-0.28	-0.08	-0.03	-0.03	-0.01	-0.02	-0.20	-0.06

2.2. Paris 1: Implications of the International Section

One straightforward way to characterize the situation in Paris 1 is to present it as the opposite of Savoie 2. Indeed, as can be seen from Table 7-3, the association of socioeconomic background with other individual and dyadic attributes is much higher in this school than in the other ones. It comes as no surprise then, that the related relational processes induce more socioeconomic homophily than in other schools, as the contribution scores reveal (cf. chapter 5, Table 5-4). Even if students exhibited no propensity to same-background selection whatsoever, socioeconomic segregation in Paris 1's networks would most likely be higher than in the other schools.

To a large extent, this can be attributed to the presence of the international section. We saw in section 1.1. of the present chapter that socioeconomic homophily appears much lower among non-international students, particularly in the early waves. In that regard, there are at least three ways in which the presence of the international section may foster particularly high levels of homophily.

First, internationals tend to be from higher backgrounds than most other upper-class students, both in Paris 1 and other schools. This has a double consequence: not only is the difference in socioeconomic background between these students and working-class ones particularly large, the correlation between students' backgrounds and other important attributes is also reinforced. That is, students from the international section have particularly highbrow tastes and leisure, they are particularly good at school, they live particularly far away from other students, and are particularly homogeneous in terms of ethnic background (native or European). This can be seen from Table 7-6: it presents the correlations between socioeconomic background and various nodal and dyadic attributes, just like Table 7-3 did, but this time also presenting these correlations for the subgroup of non-international students in Paris 1. All correlations appear lower than for the entire population of the school and, for most of them, lower than what can be found in Savoie 1. Therefore, among non-international students, the inducing impact of other relational processes (ethnic matching, academic selection, residential propinquity, etc.) on socioeconomic homophily should be less pronounced than for the entire cohort of Paris 1, and not necessarily higher than what is found in other schools[187].

Table 7-6: Coefficient of Correlation of Nodal and Dyadic Attributes with SAS Scores (Nodal) and SAS Absolute Differences (Dyadic)

	Nodal Attributes					Dyadic Attributes (matching or absolute difference)				
	Paris 1	**Paris1 except intern.**	Paris 2	Savoie 1	Savoie 2	Paris 1	**Paris1 except intern.**	Paris 2	Savoie 1	Savoie 2
Grades	0.53	**0.31**	0.20	0.45	0.33	0.20	**0.03**	0.03	0.13	0.08
Cultural status	0.65	**0.43**	0.32	0.41	0.35	0.40	**0.23**	0.11	0.14	0.10
Agreement (music)	-		-	-	-	0.13	**0.02**	0.11	0.02	0.02
Ethnic background	0.58	**0.46**	0.13	0.25	0.18	0.13	**0.05**	0.02	0.06	0.10
Classroom	-		-	-	-	0.08	**-0.02**	0.03	0.03	0.01
Primary school	0.51	**0.42**	-0.17	0.43	0.15	0.16	**0.15**	0.00	0.04	0.00
Residential distance	-		-	-	-	0.22	**0.10**	0.02	0.02	-0.03
Parental ties	-		-	-	-	0.10	**0.05**	0.02	0.04	0.01

Note: for matching dyadic attributes (agreement, ethnic background, classroom, primary school and parental ties), coefficients have been multiplied by -1 so that a positive coefficient indicates a correlation of the two types of homophily (because for continuous attributes, the homophily statistic is larger for less similar students, whereas it is the opposition for categorical attributes).

187 Unfortunately, estimating contribution scores only for non-international students would be problematic from a methodological point of view (not only would the smaller number of students reduce the statistical power of the SAOM, it would also be problematic to exclude all the triangles that involve both international and non-international nodes, as these may be important in explaining certain tie configurations among non-international ones – e.g. two non-international students meeting through a common friend that is in the international section).

For categorical nodal attributes (ethnic background and primary school), the value given is the square-root of the R-squared in a bivariate regression on socioeconomic background (since Pearson's R can only be computed between continuous variables).
Coefficients consistently appear lower at the dyadic level than for their nodal counterpart, which is a formal implication of converting individual scores into matrices of absolute differences.

Second, one particular feature of Paris 1 is that socioeconomic distribution is much more polarized than in the other schools. Indeed, it is the school with the smallest proportion of middle-class students (defined as children of artisans, small shop-owners and intermediate occupations). Without taking the international section into account, Paris 1 is the school with the lowest average socioeconomic background of the sample (SAS scores of 0.09, 0.23, 0.17 and 0.26 on average for Paris 1, Paris 2, Savoie 1 and Savoie 2 respectively). This is due to a large group of working-class students (about half of the sample), with middle- and upper-class students being present in smaller proportions (about a quarter of the sample each). However, since students from the international section are all from an upper- or middle-upper background, they drag the school's mean upward, from 0.09 to 0.44 SAS points on average, even though they only represent about 15% of the cohort. Adding this group of high-background students to the "local" population results in an approximately even numbers of students coming from upper- and working-class households, but far fewer from a middle-class background.

This polarization may make heterophilic friendship formation more difficult: students from middle-class backgrounds are structurally closer socially to all other students and may thus act as a bridge between more distant occupational groups. If this is the case, then a socioeconomic setting such as that of Paris 1 would be less favorable to heterophilic bonding than that of other schools, where middle-class households are more represented. Additionally, in Paris 1, the effect of socioeconomic distance on friendship odds is not linear but quadratic: same-background selection seems to operate only among sufficiently distant students (cf. the additional SAOM in Appendix 7C, similar to 'SAOM 1' but with a quadratic effect introduced for the SAS similarity parameter). This supports the idea that a situation in which there are large numbers of socioeconomically distant students, with few intermediate groups, may favor the emergence of high levels of homophily.

Table 7-7: Number of Students per PCS of the Reference Parent and by Section – Paris 1 (all waves)

	Paris 1 – Regular Section	Paris 1 – International Section
PCS 1 (farmers)	0 (0.00)	0 (0.00)
PCS 2 (independent craftsmen, shop owners, company owners)	8 (0.10)	0 (0.00)
PCS 3 (managers, executives, professionals)	14 (0.17)	19 (0.90)
PCS 4 (intermediate occupations)	10 (0.12)	2 (0.10)
PCS 5 (employees)	30 (0.36)	0 (0.00)
PCS 6 (workers)	6 (0.7)	0 (0.00)
PCS 8 (inactive)	2 (0.2)	0 (0.00)
No information	13 (0.16)	0 (0.00)
Total	**83 (1)**	**21 (1)**

Note: column percentages are between parentheses. The parent of reference is the one that has custody of the child or, when both do, that has the highest PCS. When the PCS of one parent was ambiguous, the other was used.

Finally, the international section constitutes a well-identified group of students, standing symbolically apart from the rest of the student body. Not only can this labeling create a barrier to friendship formation in itself, it may also act as a sort of anchor point for students' perception of their peers' backgrounds. By this, I mean that some differences in behaviors between upper- and working-class students may, in the other schools, remain unperceived by students, by lack of appropriate mental categories to make sense of several unrelated behaviors. As we saw in chapter 6, occupational groups are not, by themselves, something that students are aware of or care about, so if behavioral differences between these occupational groups are subtle enough, they may simply turn out to be inconsequential. However, should these differences largely overlap with the distinction between internationals and non-internationals, students may become more aware of them. Interestingly, this may in turn make certain perception frames more salient even among non-international students. In other words, the presence of a dense and homogeneous group of upper-class students may act as a contextual moderator which increases the salience of certain distinction criteria among other students as well.

Thus, the homophily of the international section may entail socioeconomic homophily above and beyond the direct effect that was captured by contribution scores (i.e. the increased likelihood of close-background ties that mechanically result from section-based selection). Such potential effects cannot be captured by quantitative models because they imply feedback processes between different relational processes that go beyond the correlational structure of the attributes: from the point of view of the multivariate SAOMs, this increased salience of socioeconomic attributes due to the presence of the international section should, if it exists, be captured by the SAS similarity parameter, not by the parameter of section homophily. Here, one touches the limits of the model-based dissection of relational processes, since most of these processes are eventually dependent on the contextual moderators entailed by the entire relational configuration of each school. With that said, the two examples presented in section 1.1.3. based on the interviews lend some support to this hypothesis. We saw above that certain girl students perceived the clothing tastes of Jasmine – an upper-class non-international girl – as similar to those of students from the international section, which seemingly reinforced their feeling of distance toward her. As for Aly, the working-class boy who befriended internationals during his first year of middle school, his attempt to look for academically-oriented peers was judged particularly harshly by a few of his working-class peers, who disliked the international section and resented Aly for hanging out with them. Would Jasmine's tastes have appeared less divisive without the international section, and would Aly's neighbors have more easily accepted him becoming friends with upper-class students outside of the international section? These two cases alone are far from conclusive, but the hypothesis appears plausible at the very least.

2.3. Location and Schooling Sector of the Schools

To close this chapter, let us come back to the two main criteria that distinguishes between the four schools, and that were presented in chapter 3 (section 2.2). Paris 1 and Savoie 1 are public schools, that have no tuition fees and are bound to a specific recruitment sector (except for the international section in Paris 1). On the contrary, Paris 2 and Savoie 2 are private schools, with fees that can prove significant for working-class households (several hundred euros per year). This also gives them the power of student selection, through a procedure that seems particularly selective in Savoie 2. Moreover, Paris 1 and Paris 2 are located in the Parisian agglomeration, in urban neighborhoods. This is associated with a large share of students with immigrant backgrounds, and also implies that travel between students' homes is quite straightforward. Savoie 1 and Savoie 2 are located in rural (Savoie 1) or semi-rural (Savoie 2) areas, where transportation by individual car is more important, and where students from immigrant backgrounds are in a minority. The occupations found in the different zones are also different: cultural upper-classes (e.g. teacher and academics) and care and service workers (e.g. childminders and cooks) are more represented in Parisian schools, while economic upper-classes (e.g. manager) and industrial and agricultural workers (e.g. plant operator) are more represented in Savoyard ones.

These two axes of differentiation correspond to a rather clear pattern in the sample: public schools exhibit more socioeconomic homophily than private ones (Paris 1 compared to Paris 2, Savoie 1 compared to Savoie 2) and Parisian schools more than Savoyard ones (Paris 1 compared to Savoie 1, Paris 2 compared to Savoie 2).

2.3.1. The Public / Private Distinction

The data analyzed so far does lend some support to the idea that belonging to the private sector contributes to having weaker socioeconomic homophily at the level of the school. There are clear hints of selection biases at work in the private schools that result in a weaker association of socioeconomic background with the various intervening factors that could foster homophily (residential distance, ethnic background, academic performance, and cultural tastes). This likely comes from the fact that the working-class households that choose to send their children to private schools, despite the significant financial effort this implies, are particularly invested in their children's school success, which may lead them to actively push their child toward certain academic attitudes and cultural practices (e.g. inciting them to read more). Moreover, registration to a private school may also signal some unobserved socioeconomic properties of households that are poorly captured by the declared occupation. For instance, in Paris 2, certain immigrant parents may have experienced a downward mobility upon coming to France and may have qualifications and experience that their current positions would not imply. It could also be that the working-class families that can afford private schooling have other financial resources, such as help from other family members (e.g. the grand-parents of the child may have higher socioeconomic positions). Furthermore, in the case of Savoie 2, one of the reasons often mentioned by students for their coming to this school is that the public school of the town has a bad reputation. Parents that are aware of this are probably better integrated in the local information networks, they may have received advice from primary school teachers, or they may have been incited to send their children to the private school by friends from higher socioeconomic backgrounds – in a word, they may have more social capital.

From a methodological point of view, this constitutes a warning against an overly nominal interpretation of the variables built from occupational information: these do not capture all the dimensions that are relevant to the definition of socioeconomic groups, such that the same value of the measure may indicate different social conditions depending on the local context. To an extent, it is fair to say that the occupational situations of the students' households do not really have the same "content" in the different schools. With that said, when comparing Paris 2 to Savoie 1, there is some evidence for homophilic selection based on socioeconomic background in the former (which is private) but not the latter (which is public). Similarly, the inducing impact of cultural similarity on socioeconomic homophily is higher in Paris 2 than in Savoie 1 (chapter 5, Table 5-4). Therefore, one should not be too quick to conclude that social distances are necessarily less "substantial" in private schools; arguably, this seems to be a property of Savoie 2 more so than Paris 2.

Another property that, in our sample, is associated with the schooling sector, is the integration of the middle school into a school complex with a primary and a high school. In principle, this could increase students' familiarity with the school and with some of their classmates that they have known for a long time, thus acting as a contextual moderator that reduces socioeconomic selection. Nevertheless, the only school for which there is a negative interaction effect between socioeconomic homophily and primary school homophily is Paris 1 (Table 7-8). In Paris 2 and Savoie 2, there is no evidence that students coming from the same primary school exhibit less socioeconomic homophily among each other, which is what we would expect if the school complex reduced social distances[188]. Therefore, there is no clear support for the hypothesis that the integration of the middle school into a school complex – particularly its connection to a primary school – reduces socioeconomic homophily.

188 Of course, it is still possible to argue that the effect of the school complex as a contextual moderator extends to all students of the middle school, including those that were not part of the primary school, in which case this could not be detected by within-school models (e.g. because students coming from the primary school have little socioeconomic homophily among them, it favors a form of culture of social opening and ends up reducing homophily among newcomers as well). This sounds like a somewhat far-fetched hypothesis, though, and is in any case impossible to explore with the data at hand.

Table 7-8: Socioeconomic Homophily and Primary School Homophily (ERGM) – 'Very Good Friend' Nominations, Wave 1

	Paris 1	Paris 2	Savoie 1	Savoie 2
SAS absolute difference	-0.378*** (0.071)	-0.163*** (0.024)	-0.199*** (0.048)	-0.109*** (0.037)
Same primary school	0.809*** (0.184)	0.953*** (0.068)	2.043*** (0.110)	1.307*** (0.097)
SAS abs. diff. * same primary school	0.305** (0.139)	-0.115** (0.048)	-0.089 (0.091)	-0.062 (0.071)

***: p<0.01; **: p<0.05; *: p<0.1

Note: Unshown model terms include the edge parameter, emission and reception effects for SAS and emission and reception effects for the former primary school. See Appendix 7B for the complete outputs.

Finally, a plausible hypothesis is that in-school socialization is stronger in private schools, which should tend to increase the dispositional similarity of students from different backgrounds. Indeed, schools in the catholic private sector tend to have a strong school identity, often backed-up by a relatively long history and by religious values. This school identity supports, and is supported by, a dense network of relationships between parents and teachers. Da Costa and van Zanten (2011) write: "*the adhesion to certain values […], the stability of headmasters and teachers, and the interconnections among the school community, turn many catholic schools into small or large 'domestic cities', in which the distinction between the public space of the school and the private space of the family is blurred*" (§16, I translate). They further report that, in many catholic schools, teachers proudly declare to be "*not colleagues, but a family*" (*ibid*). A similar argument has been made by Coleman and Hoffer (1987) that private schools are characterized by high levels of social capital (which, according to them, would explain their better educational outcomes).

This "family" aspect is something that I very much observed in Savoie 2, as the teachers and headmaster strongly insisted on the local anchoring and tradition of the school (for instance, some teachers are former students of the school; cf. the presentation of Savoie 2 in chapter 3, section 2.3.4). However, the extent to which this effectively increases in-school socialization in terms of academic attitudes, or even cultural tastes (by bolstering peer-to-peer socialization), is up for debate; I did not directly test for this in the present study (arguably, this would demand an in-depth ethnographic design, in addition to the network data and interviews that I have collected). Still, this may be one of the reasons that there seems to be more dispositional homogeneity in Savoie 2 than in Paris 2: the school staff from Paris 2 did not emphasize this social capital aspect of the school community nearly as much as they did in Savoie 2, which might indicate an overall stronger local anchoring – and thus in-school socialization – in Savoie 2.

2.3.2. The Urban / Rural Distinction

Three main properties that are associated with the schools' geographical locations can be identified as the following: the residential structure of the recruitment sector, the ethnic composition of the student body, and the exact type of occupations found.

The residential structure has a clear impact on youths' sociability: Savoie 1, the school with the most spread-out housing, is also the one where the place of residence and former primary school have the strongest effect on friendship formation. Consequently, it is the only school where residential propinquity fosters socioeconomic homophily (cf. the contribution scores in chapter 5, Table 5-4). With that said, the main pattern that we observe is that urban schools are more homophilic than rural ones, not the other way around, so this arguably remains a minor determinant of socioeconomic homophily. Overall, imposed foci and spatial constraints did not appear as strong inducers: disposition-driven selection and network endogenous processes contribute more to socioeconomic homophily than propinquity. Of course, this is only true for the four studied schools, and other mixed schools with higher levels of residential segregation may exhibit different patterns.

In all four schools, students tend to select friends from similar ethnic backgrounds. However, in Paris 1 and 2, the inducing effect this has on socioeconomic homophily remains relatively weak (cf. the contribution scores of chapter 5, Table 5-4). It is in Savoie 1 that this inducing effect is the strongest; the reason for this is not clear, though, and may pertain to the exact structure of the ethnic groups there[189]. In any case, this goes against the idea that greater ethnic diversity mechanically entails more socioeconomic homophily at the school level. There is no evidence that the greater socioeconomic homophily of Parisian schools can be explained by their ethnic composition.

Neither ethnic composition nor residential structure can account for the higher socioeconomic homophily found in the Parisian schools. Is it due to the type of

189 For instance, since immigrant students are fewer in Savoie 1, the same amount of ethnic selection may result in denser cliques, subject to stronger transitivity processes, which eventually create dense groups of friends that are socioeconomically very similar. This is difficult to know for sure, though – here we see the limits of the simulation method, which can make it hard to figure out exactly what happens inside this "black box".

socioeconomic groups found in the different locations, then? Unfortunately, the sample is too small to enable subgroup analyses that would assess socioeconomic homophily and/or selection for specific occupational groups. What can be noted is that Paris 1 and 2 are the two schools where there is some evidence for homophilic selection based on socioeconomic origin. Particularly clear is the case of Paris 1, the evidence was more mixed for Paris 2 (the SAOM coefficient exhibits time heterogeneity and is only significant for certain transitions; cf. chapter 5, Table 5-3). However, it can also be noted that selection based on cultural similarity – which conceptually pertains to internalized dispositions linked to one's background – was particularly salient in Paris 2. Therefore, there are reasons to believe that disposition-based selection based directly or indirectly (cultural tastes) on socioeconomic origin is more present in the Parisian schools. By contrast, the socioeconomic homophily in Savoie 1 is mostly explained by residential propinquity and ethnic selection – and there is simply not much homophily to be explained in Savoie 2. The hypothesis that the substantial "content" of occupational groups is different depending on the location – in other words, that the cultural and dispositional distance between upper-class and working-class groups is larger in urban zones – is one possible explanation for this observation.

This could come from the type of occupations represented in the different contexts. For instance, certain parents in Savoie 1 and 2 work in large industrial factories, as workers, technicians, or engineers. These occupations may experience more socioeconomic mixing at work – as workers and engineers may meet and talk – than, say, garbage collectors and university scholars do[190], which are the types of occupations found in Parisian schools. Alternatively, this could also come from a difference in contextual moderators across urban and rural contexts, even for the same occupations (for adults) or backgrounds (for students). For example, working-class students in Parisian schools may more easily identify with a "street" or "gangsta" style, as the culture it is based on is primarily urban (rap artists rarely film their clips in a field...). Moreover, as I mentioned in chapter 3, the town in which Paris 2 is located is quite famous for its working-class and immigrant population (e.g. it is often mentioned in the media under this label), so it may make students more prone to take on corresponding roles. In the same vein, upper-class households in the Parisian context may have more opportunities to participate in highbrow cultural activities (visiting museums, going to concerts, etc.) or to frequent dense and homogeneous upper-class sociability circles. For all these reasons, certain dispositions that are associated with a particular socioeconomic

190 Except sociologists, of course. We are very keen on socioeconomic mixing.

position may be either more or less salient depending on whether these positions are in Paris or in Savoie, eventually entailing different intensities of homophilic selection among students. This remains a very broad hypothesis, though, and further research would be needed to refine it and test it.

Conclusion

The sharp increase in socioeconomic homophily observed over time in Paris 1 can be explained without needing to assume that tie-formation processes changed during the period. This suggests that the networks of the early waves had not yet reached an equilibrium in terms of socioeconomic homophily, and that many “unlikely” friendships persisted for some time, but were eventually replaced by more “expected” ones. Consequently, this increase in homophily is probably not the result of some external shock or change of context in Paris 1 – and this includes students’ aging. It should rather be attributed to the initial conditions found in the school, which were already particularly favorable to the formation of homophilic friendships.

Being a public school, Paris 1 exhibited moderate to strong correlations between socioeconomic background and other factors relevant to tie formation. Crucially, socioeconomic diversity there largely came from the presence of the international section, with several direct and indirect consequences. This entailed a much more “radical” form of mixing than in the other schools, with a well-bound group of particularly high-background students that sharply differ from the majority of the school’s population regarding most attributes relevant to friendship formation (academic attitudes, place of residence, ethnic background, cultural and leisure tastes, etc.). It also created a polarized distribution of socioeconomic origins, with few students from middle-class backgrounds. Finally, this group of international students was singled out by an institutional label, making identification and distinction much easier for all students. Altogether, this not only resulted in a clear split between internationals and most other students but may have also fostered socioeconomic homophily among the students from the regular track. The international section seemingly served as an “anchor” for non-international upper-class students, some of whom progressively shifted toward them in the friendship networks and away from lower-background non-international peers. Additionally, the presence of the internationals may have acted as a contextual moderator that made various traits related to socioeconomic origin more salient in students’ perceptive frames.

In previous research conducted for my master's thesis, I studied socioeconomic homophily in another Parisian school relatively similar to Paris 1: it was a small public school, with many students from working-class and upper-class backgrounds but relatively few from middle-class ones, and with a strong geographic split, as students from different backgrounds lived in distinct neighborhoods (Chabot 2017). The research design being somewhat different, a strict comparison with Paris 1 is difficult – there was no longitudinal data in this former study, and the name generators used for creating networks were also different. Still, I found levels of socioeconomic homophily that, though noticeable, were clearly lower than what is observed in Paris 1 in the later waves (but probably comparable to that of the first wave). This suggests that not all Parisian and public schools necessarily exhibit high levels of socioeconomic homophily. Additionally, one interesting property of this former school was that the administration had an intentional policy of distributing different-background students across classrooms: in particular, students choosing the rare language options of Latin and Chinese – upper-background ones for the most part – were split between two classrooms and mixed with other students, instead of being grouped into a single classroom[191]. Since I did not conduct a formal comparison of the two schools, this should be considered with care; but it does support the idea that the management of special tracks strongly affects levels of socioeconomic homophily. This is yet another reason to believe that the strong relational segregation observed in Paris 1 pertains in large part to the presence and management of the international section.

With that said, the relative failure of Paris 1 in terms of promoting socioeconomic mixing among students – assuming that this is a desirable objective, of course – should be evaluated in regard to the higher ambition, so to speak, that this objective represents there: the school hosts the most socioeconomic diversity out of the sample, meaning that student backgrounds in Paris 1 were simply more varied than they were in other schools. By contrast, the low socioeconomic homophily found in Savoie 2 is strongly facilitated by the fact that different-background students are in fact less distant in this particular context, most likely due to a selection bias in the type of working-class families that apply and/or get accepted at the school.

191 This school policy was not necessarily meant to directly favor socioeconomic mixing, but was adopted mostly for disciplinary reasons (dispatching the "good" students across classrooms so as to not end up with an exceedingly "difficult" classrooms, i.e. that without any "good" student). See the master's thesis for details, notably chapter 1 section 2 (Chabot 2017).

As for Paris 2 and Savoie 1, they appear as intermediate cases. Interestingly, they exhibit very similar amounts of socioeconomic homophily, but the processes through which they emerge are partly different. Paris 2, like Savoie 2, exhibits low correlations between socioeconomic origin and other attributes, probably also due to it being a private school (though from what I could tell, the selection biases at work seemed less pronounced than in Savoie 2). Apart from transitivity (arguably a catalyst of homophily more than a source strictly speaking), the first driver of socioeconomic homophily there seems to be dispositional selection, either based on socioeconomic origin or on cultural similarity. This may be linked to its Parisian location, although there is no direct evidence for this in the data. Savoie 1, on the other hand, exhibits lower dispositional selection, but more indirect inducement of socioeconomic homophily through compositional effects, notably linked to ethnic background, place of residence and parental networks.

Chapter 8: The Role of Disposition-Based Selection in Socioeconomic Homophily: Evidence from a Three-Week Summer Camp

This chapter is adapted from an article co-written with Marion Hoffman (PhD researcher, ETH Zürich), which is currently undergoing revision in view of its publication in a peer-reviewed journal. Empirical treatments have been left unchanged, but the text of the article has been partly re-written to fit the overall narrative of the PhD. All figures and tables are taken from the original article[192].

***

Chapters 5 and 6 aimed at disentangling the various relational processes that jointly generate homophilic network structures in students' friendships – through quantitative modeling in chapter 5, and through the qualitative investigation of students' mental categories in chapter 6. Furthermore, the comparison of the different schools proposed in chapter 7 let us explore the ways in which these low-order processes may relate to school-level properties. While these approaches have (hopefully!) proven informative in several regards, they have important limitations. More precisely, one can identify three methodological challenges that are posed to the study of socioeconomic homophily among adolescents.

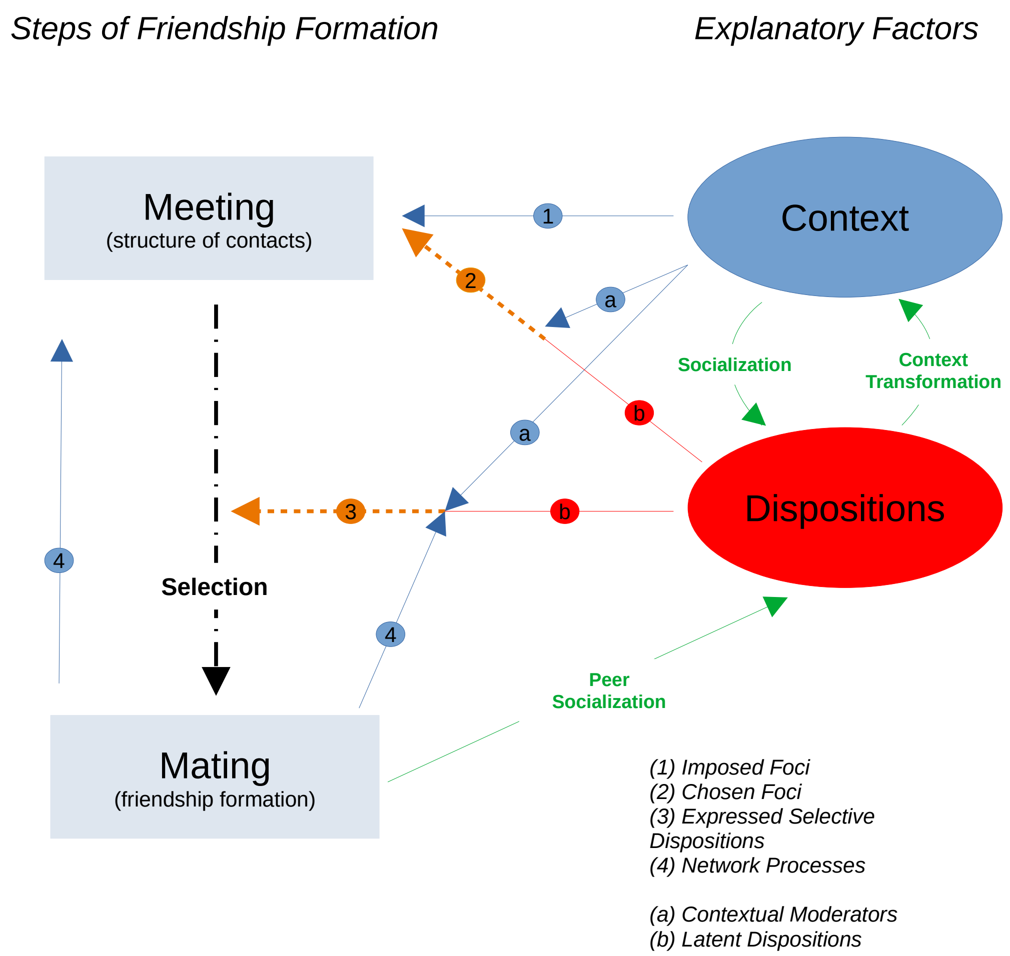
First, despite school settings being particularly convenient for the collection of network data, information about students' social lives necessarily remains incomplete. Discrete nominations tend to mask heterogeneity in the intensity and subjective content of relationships; and part of the foci through which these relationships are formed cannot be observed by the analyst, in particular those that stretch outside of school (e.g. going to the same park or mall, having common friends that do not attend the same school, etc.). Although

192 The pronoun "we" is used in sections that come from the article and thus were written by both authors, and "I" in those that have been significantly re-written for my PhD (introduction and conclusion in particular). In the latter case, the interpretations are to be considered my own, and do not engage Marion Hoffman.
If one wishes to cite results from this chapter, please also refer to the original article once published (Hoffman & Chabot, forthcoming).
Both authors contributed equally to the original article. Contributions are distributed as follows:
- T. Chabot: theoretical framework, research questions, study design and data collection.
- M. Hoffman: conception of the statistical models, data analysis, conception and editing of the figures and tables.
- Shared: interpretation of the results, drafting, revision.

I tried to mitigate these issues through, respectively, the use of different levels of nominations ("friend", "very good friend", "5 friends you spend the most time with", etc) and of geographic information about students' out-of-school foci (place of residence and primary school), it should be clear that a significant portion of students' social lives remains poorly captured by these measures. This may confound results in various ways. For instance, one may over-estimate selection processes because of unobserved foci; or, on the contrary, under-estimate these processes because of unobserved heterogeneity in friendship ties (stronger relationships may be more homophilic).

Second, many of the relational processes at work are so deeply intertwined that it is very difficult to separate them empirically. Quantitative modeling offers a partial answer to this issue, but only in cases where a given covariate can unambiguously be associated to a single process (e.g. residential proximity as a measure of propinquity). In other cases, contact opportunities and homophilic selection are essentially indistinguishable – this was notably the case of special tracks. Moreover, though quantitative models may separate different processes over a relatively short time period (here a year and a half), numerous feedback loops exist between most processes in the long-run. For example, students' selective dispositions toward socioeconomically similar peers may be inherited from past sociability contexts that were segregated, or they may get reinforced within homogeneous friendship clusters, themselves the product of propinquity or endogenous network processes. This is illustrated by Figure 2-5 below, which is taken from chapter 2: feedback processes eventually blur the distinction between the different classes of explanations. This is a difficulty inherent to most phenomena in the social sciences, but it remains extremely challenging for empirical studies. It was handled in chapter 5 by making the (necessary) assumption that students' dispositions and out-of-school contexts could reasonably be treated as constant over the span of the study, but this assumption should ultimately be seen as false.

Figure 2-5 (Reminder): Extended Model integrating Feedback Processes



Third, one class of explanatory factors could not be investigated based on the in-school samples, namely contextual moderators. Remember that selection processes pertain to students' dispositions, in the sense that they depend on their internalized traits – preferences, habits, personality, biological endowments etc. However, the expression of these traits also depends on contextual solicitations. As Lahire (2005) points out, one never gets to observe such a thing as a "pure" disposition: behaviors always emerge through the encounter of a set of dispositions and of a given environment. On Figure 2-5 above, this is illustrated by the fact that expressed dispositions (orange arrows) are functions of both latent dispositions (red arrows) and contextual moderators (blue arrows). For instance, a student may have a latent inclination toward racist prejudices (e.g. having heard racist comments from her parents), but this inclination may then be heightened, or on the contrary toned down, by contextual incentives (e.g. whether other students around her also make racist comments). From a methodological perspective, the issue is that disentangling contextual moderators from latent dispositions necessarily demands to observe individuals across different contexts. By construction, in-school networks take place in a single context, that of the school, which means that contextual moderators are essentially constant[193].

The omission of contextual moderators may be particularly problematic in cases where it hides potential transfers among estimated effects. One can think of the following example, already mentioned in chapter 5. Imagine that, in a school, classrooms are strongly segregated by socioeconomic background, such that classroom propinquity accounts entirely for the observed socioeconomic homophily. Estimating multivariate models or computing contribution scores would only tell us that, in the observed networks, socioeconomic homophily can be attributed to the distribution of students across classrooms. However, it may be the case that students did have a preference for same-background peers, but that they did not need to express this preference, as the classroom distribution alone provided them with satisfactory friendships. Should the school administration decide to redistribute students across classrooms, individuals may now perform homophilic selection to compensate for this, in which case the impact of reshuffling classrooms on socioeconomic homophily may be null. Once again, this is a common problem in the social sciences, which pertains to the fact that the assumption of independent statistical effects often proves ultimately incorrect (which is not to say that it is useless): the impact of a given variable depends on the entire configuration of variables surrounding it (Passeron [1991] 2013).

To be clear, these limitations are common to most in-school studies, and they do not necessarily invalidate the results presented before. Rather, they define their scope of validity: for instance, we found transitivity to bolster initial homophilic inclinations *in the context of socioeconomically mixed schools and over the timespan covered by the study* exclusively. With that said, the next step should be to extend these results to other settings, and to assess their claim to a higher-level of generality. This means looking for research designs able to bypass, or at least to mitigate, these three limitations.

In this chapter, I propose shifting the focus from school settings to a very different sociability context: a three-week summer camp organized by a medical association during the holidays in France. The data comes from a separate study, unrelated to the main in-school sample. As we shall see, this particular camp exhibits several important properties that make it

193 Fine-grained qualitative investigation may be able to document small-scale variations in contextual moderators even within a single school (e.g. whether a student behaves differently depending on the group of peers she is with at a given moment). However, this seems difficult with the data at hand.

possible to narrow down the effect of disposition-driven selection processes in youths' friendships. In doing so, it partly answers the challenges raised above, and offers a valuable complement to the main study.

Compared with previous studies on adolescent sociability, this approach is characterized by two key features. First, we argue that this camp in particular offers exceptional settings for the study of homophilic selection based on socioeconomic background, because most other processes that could elicit socioeconomic homophily are either structurally absent or can be accounted for in a straightforward way. The reason for this is that the camp is largely insulated from attendees' usual lives, it also exhibits a particular enrollment process that tends to favor socioeconomic diversity and to reduce selection biases in the population. In that regard, our design is partly inspired by the rationale of natural experiments, though with important differences. Second, we introduce a new class of statistical models, which is a longitudinal extension of the recently developed Exponential Random Partition Model (ERPM; Hoffman, Block, and Snijders 2020). ERPM is designed for discrete-group partitions, rather than network data, which opens for novel ways of modeling relational processes. It also makes it possible to study new types of relational data – in this case, meal-sharing patterns among adolescents – that usual network models would struggle with.

Section 1 offers a general presentation of the study. Section 2 then presents the data in more detail, and section 3 the methods used – the ERPM in particular. Results are then presented in section 4 and are discussed in section 5. The chapter's conclusion put these results in perspective with that of the previous PhD chapters.

1. Presentation of the Study

Data comes from a three-week summer camp organized in the French Alps in the summer of 2019. Some data was also collected in a similar camp two years before, in 2017, but is it far less detailed than that of 2019. Therefore, only results from the 2019 camp are presented in the body of the text; the 2017 camp is used for robustness checks, with results presented in Appendix 8F. Overall, results are very similar across both camps.

The camp is a holiday resort that gathered 60 teenagers (age 10 to 14) coming from all over France, such that none of them lives in the same place during the school year. We collected sociometric data on their sociability during the entire duration of the camp, in order to measure socioeconomic homophily among them.

1.1. General Presentation of the camp

1.1.1. A medical camp

The camp is organized by a medical institution dedicated to helping individuals that are affected by a certain rare, chronic disease. Its aim is to create a secure environment for children with this disease to enjoy their holidays, but also to learn more about the disease and how to handle it (therapeutic education). Therefore, all the adolescents at the camp suffer from this pathology. While the pathology most certainly impacts the lives of these adolescents, it does not impair their mobility nor cognitive function. As a result, attendees at the camp come from all over France. In the general population, this pathology also happens to be evenly spread across socioeconomic backgrounds. Moreover, the financial cost of the stay is entirely taken care of by French social security. Finally, all camps are multi-activity ones, and are not organized around a specific theme. These are major differences with most French summer camps, which tend to be either dedicated to one activity (e.g. sailing, horse-riding…) or to be organized by local institutions (e.g. town councils), and therefore gather attendees from a single geographical area (Rouff-Fiorenzi 2019).

1.1.2. Previous camps and lasting friendships

Given the chronic nature of the disease, many adolescents attend these camps every summer. There are several such camps organized each summer in different, far-away facilities; although families can ask for a specific one, the organizing institution distributes attendees across facilities depending on available spots as well as some internal criteria[194].

Moreover, each year, there are more demands than available spots, such that some families cannot be registered (priority is given to those that come for the first time). As a result, among the 60 adolescents of the camp, 16.5% of all dyads had met in at one of the camps in the past, and 2.% declared that they were already friends before (see data section below). While none of them met in person during the school year (living in different parts of France), they could stay in contact through social media.

194 Informal discussion with members of the institution suggest that they try to distribute adolescents who may need particular attention and effort from the camp staff evenly across facilities (e.g. unique medical conditions or known family issues).

1.1.3. Pedagogical orientation

The animation team of this particular camp (not of the entire institution) adhered to a liberal pedagogical approach, partly inspired by so-called "alternative" education theories (notably those of the "new education" movement, originally inspired by anarchist political thinking; Monchaux 2011). In practice, this means that, unlike in many other camps, adults were not the ones that decided which activities the youths should attend; nor did they decide on any form of group or team organization. Instead, activities, games and outings were agreed upon between adults and youths – often on the initiative of the latter – and only concerned individuals that volunteered to take part in them. Staying in one's room, playing outside or simply hanging around the facility was always an option for the adolescents, such "unstructured" time eventually making for a large part of the three weeks. The composition of bedrooms was also left to youths' choice on the first day; and several times during the stay some of the teenagers chose to switch bedrooms, to sleep in the same dorms as newly formed friends or following a fallout with roommates.

1.2. Isolating the effect of expressed dispositions

These particular settings provide an exceptional opportunity to study the impact of homophilic selection on adolescents' sociability. Because of the camp's medical nature, attendees were effectively cut off from their usual contexts of sociability: school, parents, siblings, neighborhood, leisure activities, etc. Thus, the only part of their “normal” life that could still impact what was going on during the camp was what they brought with them; that is, their psychological dispositions (including explicit preferences as well as habits, tastes, ways of behaving etc.)[195]. Moreover, the staff's liberal pedagogy implied that there were no imposed foci during the camp that could bias structural opportunities among adolescents. Under these conditions, the level of freedom that they enjoyed in organizing their sociability was much higher than that found in school-year contexts.

Finally, the fact that the camp creates a relatively holistic environment – teenagers live with one another for three weeks without interruption – means that there were fewer unobserved foci or network ties than in most contexts, which is typically an issue when

195 There is also the issue of the material belongings they brought with them: most notably clothing, electronic devices (phones or tablets) and pocket money. However, while these can act as signals of one's socioeconomic background, their impact on relationships is necessarily mediated by individuals' dispositions (clothing preferences, sense of fashion etc.). For instance, while pocket money can determine access to certain sociability contexts during the school year (e.g. going to the movies with classmates), this was not the case in the camp (activities were paid for by the staff for all youths equally, and pocket money was only meant to buy food or souvenirs)

modeling network dependencies (e.g. transitivity effects can sometimes reflect common participation into unobserved foci, rather than a real mechanism of triadic closure). This makes it easier to model such dependencies and, therefore, to distinguish it from other relational processes.

In these regards, our approach is inspired by that of natural experiments (Dunning 2012), in the sense that we look for "exceptional" settings that shed light on a particular process or mechanisms also occurring in "normal" settings (i.e. homophilic selection based on socioeconomic background). Nevertheless, unlike true natural experiments, we lack a control group not exposed to the same treatment (in this case, attending the summer camp). In a sense, the amount of socioeconomic homophily found by other studies in school settings act as an implicit reference for comparison, but this lacks the rigor of experimental comparison and does not allow for strict causal inference about the treatment's effect.

1.3. Selection biases

A key concern for a design such as ours is that of potential selection biases in the studied population[196]. The socioeconomic composition of the camp is of course important but is not the most problematic aspect; since this composition is known, it can be considered in the interpretation of the results. Rather, the question is whether attendees are selected on their openness, or on the contrary, reluctance to socioeconomic mixing, given their own socioeconomic backgrounds. The question extends to all the traits or attitudes that could make homophilic bonding prevalent (cultural preferences, habits, worldviews etc.).

The medical nature of the camp provides a certain amount of protection against these selection effects: enrollment is usually initiated by a doctor’s recommendation and, even though families can refuse participation, their decision should be primarily based on medical considerations (it is hard to imagine that many families would disregard their doctor’s advice and not give their child an opportunity to learn about their disease simply because they are wary of socioeconomic mixing). As for the pedagogical orientation of the animation staff, it varies from year to year and parents are not informed of it upon registering. More problematic is the fact that adolescents that have attended similar camps in previous years are more likely to try and come back the next summer. Over time, this should select for individuals that had a positive experience in previous years, relative to those who did not, which might itself be

196 Note that the term “selection” used here, which pertains to the expression “selection biases”, is completely different from the concept of selection used in my model of friendship formation – they just happen to be homonyms.

related to one's attitude toward socioeconomic mixing (if the camps are mixed, adolescents adverse to mixing are more likely to dislike their vacation).

The reasons that adolescents declared as to why they came to the camp are discussed in Appendix 8B. Some of the analyses that we shall present below are also replicated on sub-samples to further investigate selection biases (if teenagers are selected on homophilic attitudes then homophily should be lower among individuals that attend a camp for the first time). Altogether, we find no evidence that individuals coming to the camp were selected on their openness to socioeconomic mixing, though we cannot definitely disprove it either. Moreover, this selection, if it exists, should be indirect (mediation through the friendships of previous summers), as well as being interfered with by the various elements of the admission procedure that are not under the adolescents' nor their families' control (being afflicted by the disease, having a doctor that knows of these camps and that recommends them, allocation by the institution).

Therefore, we expect potential selection effects toward "pro-mixing" attitudes to be less pronounced in the camp than what is typically found in other mixed contexts, and notably schools, where avoidance of socioeconomic mixing by certain families is frequent and strategic (van Zanten 2015). In that regard, it is worth noting that the relevant reference population to which that of the camp should be compared is not a random draw from the general French population, as there is no reason to expect any sociability context to be nationally representative. Rather, it should be compared with other mixed contexts in which socioeconomic homophily has been found, such as mixed schools or neighborhoods.

1.4. Measuring social relations

In the literature, social relations among adolescents are typically measured through friendship nominations obtained by a name generator, either within a questionnaire or interview. There are, however, important limitations to such data.

First, it is subject to a number of declarative biases (Furman 1996; Marsden 1990), linked to the inherently subjective nature of concepts such as "friend" or "liking". Typically, individuals may tend to over-evaluate their number of friends so as to improve their positive self-image; or, more problematic for the present study, the definition of what a friend is may vary with socioeconomic background (as Bidart 1991 found for adults). To mitigate those issues, one solution is to use a combination of self-reports and behavioral measures (Baumeister, Vohs, and Funder 2007). Second, the discrete nature of these nominations means we lose information as per the strength of relationships, both above and under a (subjectively determined) threshold. In order to account for the strength of relationships, one can instead measure repeated interactions over time, as stronger relations should generally imply more interactions.

Last but not least, there may be certain relational processes that pertain to group structures in social interactions, and that are not appropriately captured by friendship nominations collected at the level of the dyad. Indeed, many sociability settings are organized as discrete groups, either because of physical constraints or endogenous relational processes. We call group dependencies or group endogenous mechanisms – by analogy with network endogenous mechanisms – the structural constraints that are implied by such settings. For example, an adolescent may need to choose between two groups, while two of her friends are present in one group each (whereas in a network, she could maintain both ties at once); the size of groups may be limited by some external feature (e.g. the number of players allowed for a particular board game); or individuals may have a preference for certain group sizes or structures, for various reasons (e.g. looking for small groups to share secrets, but large groups to organize a party). In classical network modeling, such constraints are partly accounted for by network endogenous terms such as transitivity, which should effectively summarize group dependencies over a large number of distinct interactions. Nevertheless, we argue that group dependencies sometimes imply higher-order dependencies than the ones observed in networks, and therefore, that explicitly modeling them provides additional leverage to understanding complex endogenous mechanisms in social interactions.

For these reasons, we collected two types of relational data. First, teenagers were asked to nominate their friends through written questionnaires, twice during the stay. From there, we construct two directed, complete friendship networks. Second, we systematically noted who ate with whom during meals, twice a day (lunch and dinner) for the whole camp. This results in 38 separate observations of co-seating arrangements, that is, 38 partitions. This is an objective measure of social interaction, which neutralizes differences in subjective perceptions among adolescents as per the nature of friendships; and it is data organized in partition (i.e. group) rather than network form, therefore allowing us to directly model endogenous group dependencies.

1.5. Research Questions

To summarize, there are two reasons why this case-study offers strong insights into the relational processes underlying homophily. First, several potential inducers of socioeconomic homophily are structurally absent from the camp, leaving a smaller range of potential explanations to explore. Second, we were able to collect particularly detailed and extensive data about teenagers' sociability over these three weeks, which lets us model their relationships in novel ways, as well as reducing the amount of missing information about their sociability (unobserved foci).

To understand how this impacts our ability to isolate selection processes, it is necessary to come back to the theoretical framework introduced in chapter 2, and to the four classes of relational processes that were identified and that are susceptible to eliciting socioeconomic homophily. These were:

1 *imposed foci:* segregated meeting contexts over which individuals have no control (e.g. school or parental networks)

2 *chosen foci*: segregated meeting contexts in which individuals can self-select (e.g. activities or games)

3 *expressed selective dispositions:* an individual's propensity to select certain friends rather than others, given the structure of contacts among them (e.g. looking for friends with certain musical tastes). Homophilic selection is one particular form of these processes.

4 *endogenous network/group mechanisms* (e.g. aggravation of homophily through transitivity)

We can now assess which of these should be operative during the camp. Processes pertaining to point (1) (*imposed foci*) are simply absent here, because structural opportunities were equalized at the beginning of the camp. Surely, there can still be unequal meeting opportunities among teenagers during the stay, but those can only be the endogenous product of their sociability.

Similarly, some usual inducers of homophily that pertain to point (2) (*chosen foci*) are absent from the camp (e.g. activity clubs). As for the chosen foci that do emerge during the camp, we are able to account for them using data on teenagers' activities and bedrooms (see data section below). Moreover, point (5) (*endogenous network mechanisms*) is something that we are able to model explicitly, both for friendship networks and co-sitting during meals (see method section below). Its impact (or lack of) on socioeconomic homophily can therefore be isolated relatively precisely.

The processes relating to point (3) (*expressed selective dispositions*) should be operative during the camp, and they are our main focus. Furthermore, as was discussed in chapter 2, expressed dispositions can be split into two component parts: *latent* or *intrinsic dispositions* – that is, the (unobserved) inner-state of individuals – and *contextual moderators* – all the elements from the context through which latent dispositions are “filtered” in producing effective behaviors. These moderators include the composition of the camp’s population (e.g. how many same-background peers were available for each individual), as well as the general context in which that sociability took place (the facility, the animation team, the overall “climate” of the stay, etc.). Given our aim is to gain knowledge about school-year processes and not only about this particular camp, ideally we would like to separate these two components. Nevertheless, this is especially difficult from a methodological point of view: since we are looking at a single case-study, there is no variance on these moderators, such that their effect cannot be isolated. We are therefore unable to distinguish between “raw” individual preferences, and contextual moderators: we only get to observe the expression of preferences that has already been “filtered” by contextual elements.

With that said, one key feature of the camp is the fact that, as we mentioned, teenagers enjoyed a great deal of freedom in organizing their sociability. Under these conditions, they should have been able to express their relational inclinations in a relatively unconstrained way: contextual moderators most likely affected this expression, but it seems unlikely that they would entirely over-write it (e.g. forcing the formation of a friendship between two highly dissimilar individuals that also have strong latent homophilic inclinations). In other words, if it is the case that teenagers came to the camp with a strong preference for socioeconomically similar peers, then one would expect at least some level of socioeconomic homophily to be apparent in their friendships or co-sittings. Conversely, if they came to the camp with no such preference whatsoever, then it seems unlikely that context alone would generate high levels of homophily.

Altogether then, the descriptive amount of socioeconomic homophily that will be observed during the camp can be seen as the joint product of processes pertaining to points (2) to (4). Our first research question is therefore whether there was such homophily at all.

RQ1: in the absence of socioeconomically segregated imposed foci, is socioeconomic homophily observed during the camp?

Furthermore, we will use appropriate modeling tools to assess the effects of *chosen foci* (2) and *network endogenous mechanisms* (4), such that homophily estimates will closely reflect the effect of *expressed dispositions (3).* Note that our primary interest is in these expressed preferences; network mechanisms and chosen foci are controlled for rather than specifically investigated, because our design is not particularly well-suited to illustrating the effect they may have in other contexts (network mechanisms mostly work as an aggravating mechanism, meaning that their impact on homophily depends on the initial amount of homophily induced by other factors; and many important chosen foci found in school-year contexts are absent from the camp). This leads us to our second research question:

RQ2: is there evidence for expressed selective dispositions acting as inducers of socioeconomic homophily during the camp?

Finally, although further separating latent preferences and contextual moderators is not possible with the data at hand, we will provide informed suggestions in the discussion section, based on the literature as well as on exploratory treatments from our data.

2. Data

In this section, we provide a description of the data used in this study, including the individual attributes of the adolescents participating in the camp and the two measures of friendship (i.e., reported friendships and meal sharing behaviors) used in our analyses.

2.1. Individual attributes

Gender, age, and parental occupations were collected through questionnaires for each participant. Among 60 participants, 25 were girls and 35 were boys, all of them aged from 10 to 14 (M=12.48, SD=1.14). The declared occupation of their parents was first coded into the International Standard Classification of Occupation score (ISCO-08; International Labor Office 2012) and further translated to the International Socio-Economic Indexes of occupational status (ISEI; Ganzeboom et al. 1992), using the R package *ISCO08ConveRsions*[197].

197 The choice was made to use ISCO/ISEI instead of Socio-Academic Scores (SAS), unlike in the rest of the dissertation. As explained in chapter 3, SAS are meant to order occupational groups based on the expected school performance of their children; the rationale being that students' inherited cultural or academic capital is most impacting on their relationships (relative to economic capital in particular). This seemed appropriate for in-school samples, but the situation might be different during the summer camp (less solicitations from

For each participant, the average value of parents' ISEIs was computed in order to capture the influence of both parents' occupation equally, as well as correcting for potential errors in participants' self-reports, following the approach of Thaning and Hällsten (2020). Alternatively, the maximal value of parents' ISEIs was used for robustness checks, following the argument of Erikson (1984) that the living conditions in a household can be better represented by the highest SES of the parents. The average obtained scores range from 15.49 to 85.55 (Mean=55.30, Median=58.69, SD=20.72), covering the whole spectrum of socioeconomic backgrounds since the individual ISEI variable takes values between 10 (for execution workers) and 90 (for managers and professionals) (Ganzeboom 2019). However, we observe an over-representation of high ISEI scores compared to a national sample (see Appendix 8B on selection biases).

2.2. Friendships anterior to the stay

During the stay, participants were asked whether they had already participated in camps organized by the same association in the past, and whether they were friends with individuals they had met at those previous camps (these medical camps are organized every year). From the piece of information, we construct a network of previous encounters, and from the second, a network of friendships anterior to the camp. Note that these are retrospective declarations, thus potentially subject to recollection biases (previous friendships in particular). Altogether, 613 directed dyads (out of the 3540 of our camp) had already met in the past (up to 4 years before this camp), of which 101 corresponded to a friendship nomination. Put into percentages, 16.5% of adolescents pairs who got a chance to meet in the past, and 2.8% of all participants pairs, indicated that they had been friends before the beginning of the camp. Among the 60 participants, 38 reported having at least one friend from previous camps.

As mentioned before (section 1.3), these past friendships can impact the probability of coming back to the camp from one year to the next. In particular, 14 participants declared that they had chosen this camp specifically in order to meet with past friends. These older participants were more often from upper-class backgrounds, which may contribute to explaining the over-representation of high ISEI scores in the sample. Nevertheless, we find no

adults that would “activate” or reveal individuals’ cultural capital, as teachers do at school), thus the use of a more standard measure such as ISEI. Unlike what was found for the in-school sample, there was no marked difference in levels of socioeconomic homophily when using either measure (in the in-school sample, homophily was weaker when using ISEI, cf. Appendix 3D). The correlation between the two measures is of 0.89 among the 60 attendees of the camp (0.86 in the in-school sample).

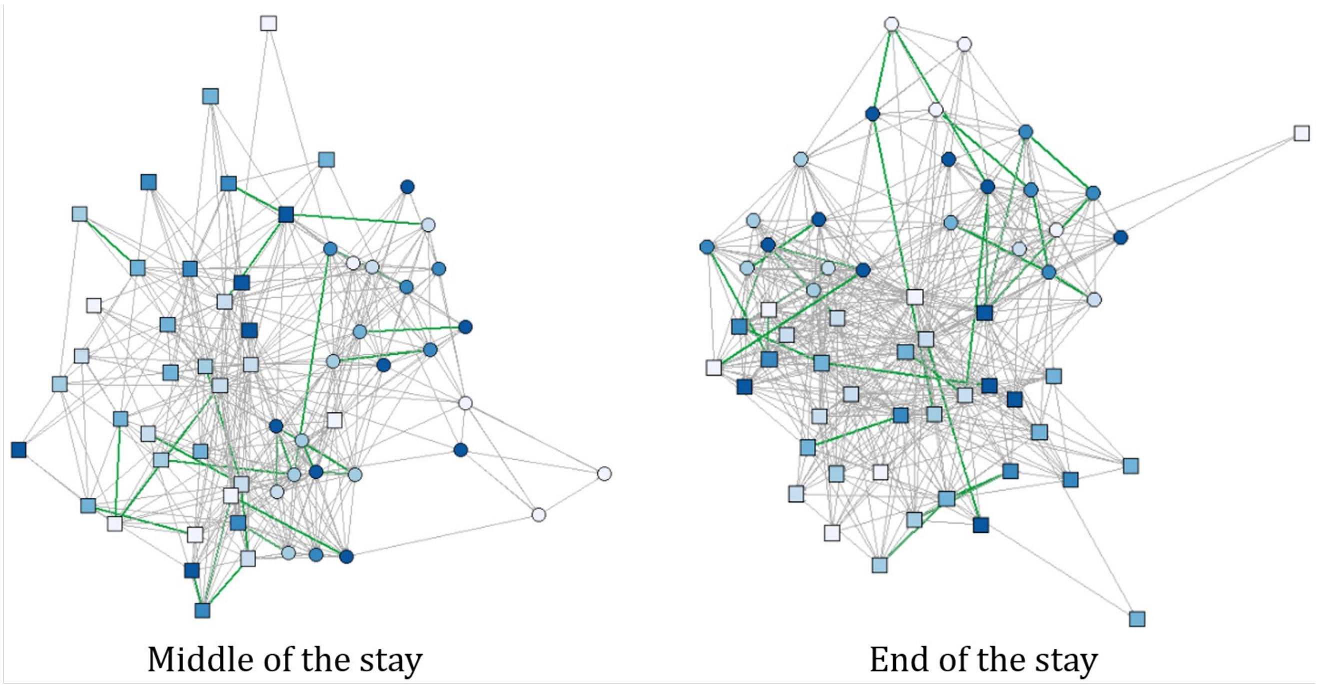
evidence that they were particularly well-disposed toward socioeconomic mixing – perhaps even the opposite, in fact (see Appendix 8B).

2.3. Friendship networks during the stay

Friendship networks constitute our subjective measurement of social relations during the stay. Participants filled in two paper questionnaires in which they indicated for each other participant whether they considered them as a friend at the moment they were asked. The first questionnaire provides a friendship network corresponding to the middle of the stay (9th and 10th day) and the second a network corresponding to the end of the stay (19th and the 20th day). All participants answered, except one who left the camp before its end and is missing in the second network.

Figure 8-1 shows a representation of these two friendship networks. In the middle of the stay, the density of the network (16.3%) is already rather high, as is the average in-degree with a value of 9.62. The density in the network collected at the end of the stay slightly increases to 20.2%, and the average in-degree reaches the value of 11.73.

Figure 8-1: Friendship Networks during the Stay



Note: Friendship networks observed in the middle of the stay (one week and a half into the camp) and at the end of the stay (three weeks into the camp). Node shapes indicate gender (circle = girl, square = boy) and node color indicates average ISEI scores of parents (darker shade = higher score). Green edges denote friendships anterior to the camp.

In terms of attributes, all networks appear segregated in terms of gender with a proportion of same gender ties above 70%. Age differences between friends are relatively low, around 1. Regarding SES, the average ISEI difference in friendships is 22.95 in the middle of the stay and 22.84 in the last network.

2.4. Meal partitions

In addition to the subjective measures of friendship, an observer recorded for each lunch and dinner (in total, 38 meals) the composition of the groups eating together, which provides for each meal a partition of the present participants. Camp organizers also ate at the same place but were not included in the data. The number of individuals present varies from one meal to another, ranging from 27 to 60 and averaging around 49 (small groups of teenagers sometimes ate outside during out-of-facility activities, in which case they were not part of the observed partition for that meal). We observe on average 13.5 groups per meal, with an average size close to 4 and an average number of participants eating alone around 3. The groups seem homogeneous in terms of gender and age: the proportion of same gender pairs in the groups is above 90% on average and the age difference in these pairs is around 1. Regarding ISEI scores, we observe an average difference of 21.54 and an average intra-class correlation of 0.34.

Unsurprisingly, friendships and table partitions are linked. Friends reported in the middle of the stay ate together 8.2 times on average during the whole camp, and friends reported at the end of the stay 7.3 times (vs. less than 0.9 times for non-friends). Reversely, the proportion of co-eating pairs in which at least one individual has indicated the other as a friend is slightly above 50% in all meals, for both friendship networks. As a reference, the proportion of such pairs among all possible co-seating ties is only 32% and 39% for the two networks respectively.

2.5. Rooms and activities

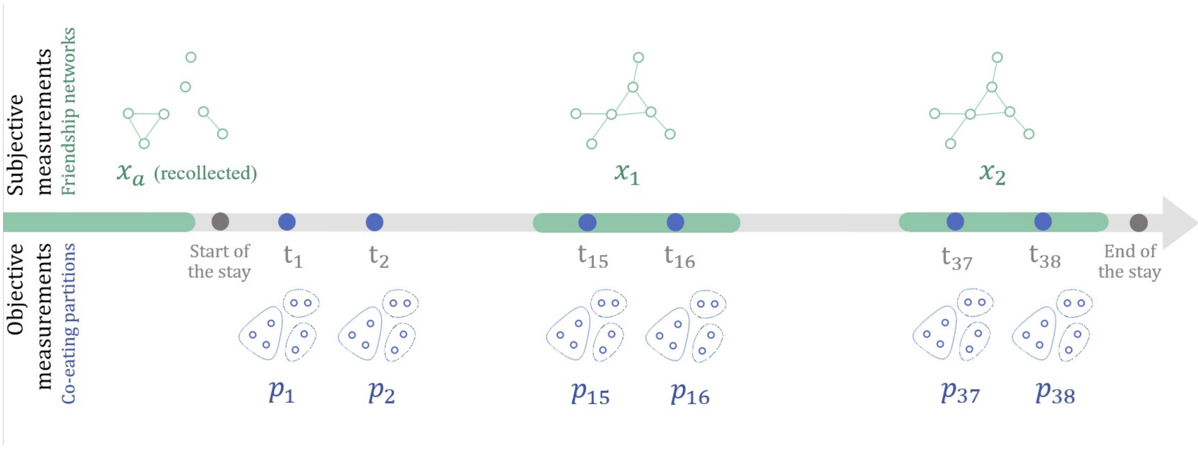
Upon arrival at the camp, the adolescents divided themselves into 15 rooms, with sizes ranging from 2 to 7. They were free to form rooms as they wished – to do so, they typically relied on the affinities formed during the first moments of the camp, the car trip up to the facility, or (for some of them) their acquaintances from previous years. Thus, rooms are a type of chosen foci, although with a strong random element as well, given how little information teenagers had about one another at this point. These rooms were not mixed in gender, and rather homogeneous in terms of age. Their composition slightly changed during the stay (5 pairs of participants switched rooms and 6 individual participants transferred to another room). Moreover, each participant could participate in two paid activities among eleven proposed in the course of the stay (each typically lasting half a day), in addition to the many in-camp activities. Bedroom compositions and attendance to the out-camp activities were systematically collected.

3. Methods

3.1. Notations

We define x_a as the network of anterior friendships, i.e., friendships prior to the camp, x_1 the friendship network reported in the middle of the stay, and x_2 the friendship network reported at the end. For each meal at time *t*, with *t* between 1 and *T*=38, the set of groups eating together is defined as the partition p_t of the present participants. Figure 8-2 summarizes the timeline of the collection of networks and co-eating partitions. Blue points indicate the times at which the adolescents had meals, while green bars represent the periods during which the ties in the networks x_a, x_1 and x_2 were collected (sometimes overlapping with several meals).

Figure 8-2: Timeline of the Network and Partition Observations Collected during the Camp



3.2. Permutation Tests

To answer our first research question, we need to assess the presence of homophily in the observed networks and partitions. For this, we define the statistics that capture similarity in ties and groups as well as the procedure to compare the observed levels of similarity to reference distributions.

In the case of networks, similarity between friends in terms of ISEI is described by the average absolute differences in ISEI scores among friendship ties. In the partitions, we use the average absolute score differences for co-seating pairs to assess dyadic homophily inside the groups, as well as the sum of the range of ISEI scores in each group to assess group homogeneity. In both cases, these statistics highly depend on the distribution of attributes in the sample and on the structural characteristics of the networks and partitions (e.g., density in networks, group sizes in partitions).

We resort to permutation tests to test whether the observed values of these statistics would be plausible under a reference distribution that controls for the attributes present in the sample and the network and partition structures. Specifically, our test compares the observed statistics to the statistics obtained for networks and partitions in which nodes are randomly permuted, similarly to the logic of the Quadratic Assignment Procedure (QAP; Krackhardt 1988). The comparison of network or partition statistics to a null distribution follows the same procedure as Conditionial Uniform Graph Tests (Anderson, Butts, and Carley 1999; Wasserman, Faust, and Urbana-Champaign 1994). Non-parametric two-sided p-values are calculated as the proportions of permuted networks or partitions for which the test statistics are greater than or equal to the observed statistics.

Importantly, these tests act as marginal tests and do not take into account the distribution of other variables, which might bias the results. Moreover, these tests control for dependencies in ties and groups, but do not model them explicitly. To fully understand the factors underlying the formation of ties and groups, and to identify the role of the different homophily-inducing processes identified above, we resort to more elaborate statistical models described in the following two sections. To make an analogy to the analysis of independent single observations, the permutation tests serve a similar purpose as classic correlation tests to assess the association of two variables, while the models described below can be considered equivalent to running a multivariate regression model.

3.3. Network modeling: the Stochastic Actor-Oriented Model

To understand the factors underlying the formation of friendships, and therefore answer our second research question regarding the role of homophilic preferences, we model the networks x_a, x_1 and x_2 using the Stochastic Actor-Oriented Model (SAOM). This is the same class of models that was used in chapter 5 of the PhD. As a reminder, SAOM is a model of network evolution, that takes the first network observation (here x_a) for granted and tries to predict the odds of tie's creation, maintenance, or deletion from that point onward. This is done through a simulation process where nodes (actors) are "asked" by the algorithm to adjust their outgoing ties at fictive "mini-steps", artificially reconstituted in-between empirical observations. The evolution of the network is described by a set of aggregated target statistics (network-level counts of specific low-order tie configurations), and the model parameters give the conditional (logged) odds for the corresponding low-order configurations to be observed at a given mini-step (always from the perspective of a focal node). The aim of the algorithm is to find parameter values that jointly predict the correct values of the target statistics, which is done by iterative trial-and-error (trying out initial parameters, simulating networks, assessing the error compared to empirical target statistics, adjusting the parameters, etc., until convergence is reached).

Here, we describe how the processes related to homophily identified in Section 1.5 are translated into SAOM effects (i.e. target statistics and corresponding parameters), in three different model specifications (Models 1, 2, and 3)[198].

To capture *endogenous network mechanisms*, all models include three types of structural effects that depend on tie configurations in the network. These effects represent general tendencies commonly observed in social networks and reproduce sociologically reasonable dependencies among ties. First, they include the tendency towards reciprocating ties. Second, they include the tendency of closing transitive triads (i.e., forming the tie $i \rightarrow j$ when $i \rightarrow k \rightarrow j$), to represent individuals' propensity to befriend the friend of a friend; of forming a reciprocated transitive triad (i.e., forming $i \rightarrow j$ when $i \rightarrow k \rightarrow j$ and $j \rightarrow i$), to represent the interaction between reciprocity and transitivity (Block 2015); and of forming triads of the form $i \rightarrow j$ with $i \rightarrow k \leftarrow j$, to represent individuals' propensity to befriend an actor who has nominated one common friend (this is essentially another form of transitivity). These triadic effects use the definition of geometrically weighted edgewise shared partners effects (Hunter 2007; Snijders et al. 2006). Third, the dynamics of in-degrees and out-degrees are captured by effects representing the dispersion of in- and out-degrees and the co-variance between in- and out-degrees. One can note that degree effects offer some protection against omitted variables that might influence degree distributions (Ripley and Snijders 2020), while triadic effects can

198 Most effects mentioned here were also present in the SAOM from chapter 5, with the exception of the second type of transitivity effect ($i \rightarrow j$ when $i \rightarrow k \leftarrow j$). On the other hand, 3-cycles ($i \rightarrow j$ when $i \leftarrow k- \leftarrow j$) were sometimes used in chapter 5, but not here. As usual, structural effects were chosen so as to improve the model's fit as much as possible, which may therefore result in slightly different specifications depending on the networks.

partially control for homophily on unobserved attributes or meeting opportunities (Block 2018).

To answer our second research question, we add to all three models an effect for the tendency toward ties among dyads with similar ISEI scores (ISEI homophily effect). This effect is meant to capture *expressed homophilic dispositions* toward socioeconomically similar others (itself the joint product of latent dispositions and contextual moderators), given proper control for other homophily-inducing mechanisms in the rest of the model. This effect is calculated as the sum of centered similarity scores between a central actor *i* and potential friends *j*. We also estimate separately a variation of this specification of homophilic preferences in Model 2, where we add an endowment effect (Snijders, van de Bunt, and Steglich 2010) for the similarity effect, to test whether similar ties are particularly likely to be either maintained or created[199].

The influence of *chosen foci* is modeled by including effects for the tendency towards ties between individuals sharing a room or participating in the same paid activity, as these are the foci whose composition are known. In the case of activities, this effect is only included in the second period since they occurred approximately in the middle of the stay. As the assignment into room and activities can be endogenous to the formation of friendships, these effects are only added in Model 3, after examining the parameters of the model without them.

From the four homophily-inducing processes proposed in Section 1.5, we therefore capture three of them in the model. The influence of *imposed foci* is considered null, for reasons explained previously. Moreover, one can note that we do not include the effect of all chosen foci, since we did not have the composition of the games or activities that were organized every day on camp. However, given the importance of rooms and paid activities, we can expect their effects to be relatively indicative of the general effect of chosen foci in the camp. Moreover, the inclusion of endogenous triadic effects provides some confidence that

199 One important difference between the models estimated here and those of chapter 5 is that, in the camp's models, we do not try to distinguish between direct homophilic selection based on socioeconomic origin, and homophilic selection based on correlated attributes that induce socioeconomic homophily through compositional effects (e.g. selection based on ethnicity, cultural tastes or academic results, as I did in chapter 5). The reason is that our theoretical interest here is to separate foci, network processes and homophilic selection; what this homophilic selection exactly consists of (i.e. what traits students perceive and operate selection on) is a different matter. Consequently, in the rest of the chapter, "socioeconomic selection" should be understood as selection pertaining to any trait derived or associated to socioeconomic origin, not necessarily a direct preference for same-background friends. Typically, similarity in cultural tastes may be one key mediator of socioeconomic selection during the camp (students with the same interests in music or sport will likely be attracted toward one another).

the effect related to expressed dispositions does not capture the effect of meeting opportunities.

Other effects are included to control for confounding effects. We model the tendency toward more incoming or outcoming ties for individuals with high ISEI scores, to control for popularity effects of certain ISEI values. We also test the effect of gender and age, following their particular salience in adolescent relations. We include effects for the tendency toward ties between same gender or similar aged individuals, as well as the tendencies to observe incoming and outcoming ties for boys or for older individuals.

Finally, we include time heterogeneity effects because the two periods we model are highly different in terms of the number of ties created. These effects are essentially interaction dummies for the effects whose parameters might change over periods, following the approach of Lospinoso and colleagues (Lospinoso et al. 2011).

3.4. Partition modeling: the longitudinal Exponential Random Partition Model

Following a similar logic, we want to understand the mechanisms underlying the composition of the groups formed during meals and identify the ones related to SES homophily. The statistical model used for this purpose is a longitudinal extension of the partition model developed by Hoffman et al. (2020). This model allows us to model social processes underlying the formation of non-overlapping groups, such as the ones we observe during the camp meals, while taking into account the dependencies between group memberships. The partition model defines exponential family distributions for cross-sectional partitions, and social mechanisms are modeled as sufficient statistics of these partitions. Because of the theoretical similarities with the Exponential Random Graph Model (Lusher et al. 2013; Robins et al. 2007), this model is also referred to as Exponential Random Partition Model (ERPM). However, because the dependencies between groups in a partition and ties in a network are essentially different, major differences exist in the specification and interpretation of sufficient statistics in ERGMs and ERPMs. Most importantly, partition statistics should be defined as counts of group statistics, rather than dyadic tie statistics (for more detail on the definitions and properties of the model, see (Hoffman et al. 2020).

3.4.1. Model Definition

At each time point$t \in \{1, ... T\}$, we define a random partition$P_t$over the set $\mathcal{P}_t$of all possible partitions of the individuals present at this specific meal. We pose β the parameter vector associated to the sufficient statistics vector *u* that contains statistics of the partition P_t. Some statistics in *u* therefore depend on both the current modeled partition P_t and the previous observations p_1 up to p_{t-1}, except in the case *t = 1*. For convenience, we set the statistics depending on previous observations as zero in the model at time *t = 1*. The associated probability distribution at time *t* is expressed by the joint form:

$$1 \leqslant t \leqslant T : Pr_\beta\left(P_t = p_t \vee P_{t-1} = p_{t-1}\right) = \frac{\exp\left(\beta^T u\left(P_t, p_{t-1}, \ldots p_1\right)\right)}{\sum \exp\left(\beta^T u\left(\widetilde{P}, p_{t-1}, \ldots, p_1\right)\right)} \quad (1)$$

The novelty added to the original model here is that the probability distribution defined at time *t* depends explicitly on the previously observed partitions, from time 1 up to *t* – 1. The probability distribution for the whole sequence of meals is then simply the product of all these distributions, for *t* varying from 1 to *T*. By doing so, we turn the model into a longitudinal model and consider that each new meal represents a new partitioning of the adolescents influenced by the previous meals.

It is important to note that the sets of individuals present at a meal varied from one meal to another. The presence and absence of nodes is not explicitly modeled, as it was impossible to assess whether an individual skipped a meal for endogenous reasons (e.g., not wanting to socialize with others), or exogenous reasons (e.g., being engaged in an activity or ill). The statistics calculated from the previous observed partition p_{t-1} in the model at time *t* are only calculated for the individuals at time *t* (i.e., the information on absent individuals is not used).

The computational time needed for the estimation of a longitudinal ERPM is high at this time – in part because of the nature of meal partitions, and in part due to it being a new class of models whose algorithm can be further optimized[200]. As a result, fitting a well-converged model on all meals (i.e. 38 observations) takes a high amount of computing power. For this reason, we present in the text a model estimated on a subset of meals, which are the dinners collected every three days during the camp (seven observations). As shown later, socioeconomic homophily in this subsample is comparable to the general sample[201].

200 Regarding the first point, the simulation of the Markov chains requires an important number of steps, because transitions between discrete partitions imply large statistics changes, making the space of partitions more difficult to explore than is typically the case for friendship networks. As for the second point, the current code has not been optimized for high amounts of calculations (notably, ERPM is implemented as R code at the time, but could be made more efficient in another computing language).

201 Note that the complete model (i.e. on all meals and with good convergence) is still under development, and will hopefully be achieved for the forthcoming publication of the article of which this chapter is an adapta-

3.4.2. Model Specification

As with the SAOM, the processes we aim to model are represented by effects, that are in essence the sufficient statistics *u* of Equation 1. Again, we list the effects used for this study and describe how they relate to processes potentially responsible for socioeconomic homophily in three different model specifications.

Structural effects are included in all three models to reproduce the group size distribution and control for *endogenous group mechanisms*. As proposed in the original definition of the ERPM (Hoffman et al. 2020), we use the count of groups in the partition as a general tendency to form groups. We also include the count of isolates and the sum of squared sizes to model the proportion of isolates and sizes dispersion. Table constraints are not modeled because the observed group sizes never reached the maximal size of tables, and qualitative observations also indicated that the adolescents were free to re-organize the tables and chairs to their liking. To limit computational costs in estimating the model, we simply limited group sizes to 9, which is the maximal observed size.

Expressed dispositions toward similar socioeconomic backgrounds are represented in all three models by two effects, one representing dyadic similarity and the other group homogeneity[202]. The dyadic similarity effect represents the tendency of individuals to be similar to all other group members in a pair-wise manner, while the group homogeneity effect captures the tendency for groups where everyone is close to the same attribute value, based on the idea that both dyadic and group processes might be important to consider. We model homophily at the dyadic level using the count of absolute differences in co-seating ties, and group-level homophily using the range of ISEI scores within groups. Similarly, to the endowment effect in the SAOM, we include in Models 2 and 3 en effect that tests whether ties between similar individuals are more likely to be repeated: this effect is an interaction between the count of absolute differences and the number of previously shared meals (see below).

Additionally, we include effects meant to capture the longitudinal dynamics of the meals. First, we add to all three models an effect representing the influence of previous meals, in order to capture the tendency to re-conduct the same group configurations over time. The statistic counts the sum of meals shared in the past (from time 1 to *t*-1) among the individuals

tion. The uncertainty associated to using a smaller subset of meals will be taken into consideration in the interpretation of the results (see below).

202 Same remark as above for the SAOM: selection toward same-background peers include selection on correlated attributes (ethnicity, tastes, etc.).

in the same groups. Moreover, in order to get a specification comparable to that of the SAOM (which took the friendships anterior to the camp as a starting point), we add an effect counting the friendship ties anterior to the camp in group meals (i.e. how many nodes who were friends prior to the camp ate together at a given meal). Finally, in Models 2 and 3, we add the pseudo endowment effect mentioned earlier to check whether homophilic selection may differ in the formation and maintenance of relationships.

Regarding *chosen foci*, the effect of rooms is represented by the count of roommates in the same groups. The effect of activities could not be added to the ERPM, since the exact date and time of the different activities was not available. The effect of rooms is only added in Model 3, similar to what we did in the SAOM.

Finally, we include additional effects to control for the influence of other attributes. First, we control for the popularity of certain ISEI scores by including interactions between the sum of ISEIs in a group and the group size. Second, we add homophily, emission and reception effects for age and gender, with homophily being considered at the dyadic level and the group-level in a similar way as for the ISEI scores.

Table 8-1 details the effects presented above, as we use here a novel type of model. In terms of notations, the letter G is used to denote groups within the partitions and the symbol # defines the size, or cardinality, of a set (which can be a group or a partition). For a given group $G \in P_t$, the subset of G only including individuals with the attribute a equal to a value A is noted as $G_{a=A}$. The maximum (or minimum) value of an attribute within a group G is $max_G(a)$ (or $min_G(a)$).

Table 8-1: Definition of Effect Terms Used in the Partition Model

Effect	**Sufficient Statistic** $u(P_t, p_1, \ldots, p_{t-1})$
Tendency to form groups	$\# P_t$
Tendency to form isolates	$\sum_{G \in P_t} \lvert \# G_t = 1 \rvert$
Dispersion of group sizes	$\sum_{G \in P_t} \# G^2$
Dyadic homophily for a binary attribute *b*	$\sum_{G \in P_t} \sum_{i,j \in G; i \neq j} \lvert b_i = b_j \rvert$
Group homogeneity for a binary attribute *b*	$\sum_{G \in P_t} \lvert \# G_{b=0} = \# G \text{ or } \# G_{b=1} = \# G \rvert$
Popularity for a binary attribute *b*	$\sum_{G \in P_t} \# G \# G_{b=1}$
Dyadic homophily for a continuous attribute *c*	$\sum_{G \in P_t} \sum_{i,j \in G; i \neq j} \lvert c_i - c_j \rvert$
Group homogeneity for a continuous attribute *c*	$\sum_{G \in P_t} max_G(c) - min_G(c)$
Popularity for a continuous attribute *c*	$\sum_{G \in P_t} \# G \sum_{i \in G} c_i$
Influence of all previous co-sittings	$\sum_{G \in P_t} \sum_{i,j \in G; i \neq j} \sum_{t' < t} \lvert p_{t'}(i) == p_{t'}(j) \rvert$
Influence of anterior friendships	$\sum_{G \in P_t} \sum_{i,j \in G; i \neq j} \lvert x_a(i,j) == 1 \rvert$
Previous co-sittings X dyadic homophily	$\sum_{G \in P_t} \sum_{i,j \in G; i \neq j} \sum_{t' < t} \lvert p_{t'}(i) == p_{t'}(j) \rvert \lvert c_i -$
Influence of sharing rooms	$\sum_{G \in P_t} \sum_{i,j \in G; i \neq j} \lvert p_r(i) = p_r(j) \rvert$

4. Results

Results are organized as follows. We answer the first research question by testing for the presence of socioeconomic homophily in teenagers' social relations during the camp, through both subjective (i.e. friendship) and objective (i.e. meal partitions) measures. We then turn to the second research question regarding expressed homophilic dispositions. Finally, we investigate the link between anterior friendships and those formed during the camp.

4.1. Socioeconomic homophily in the camp

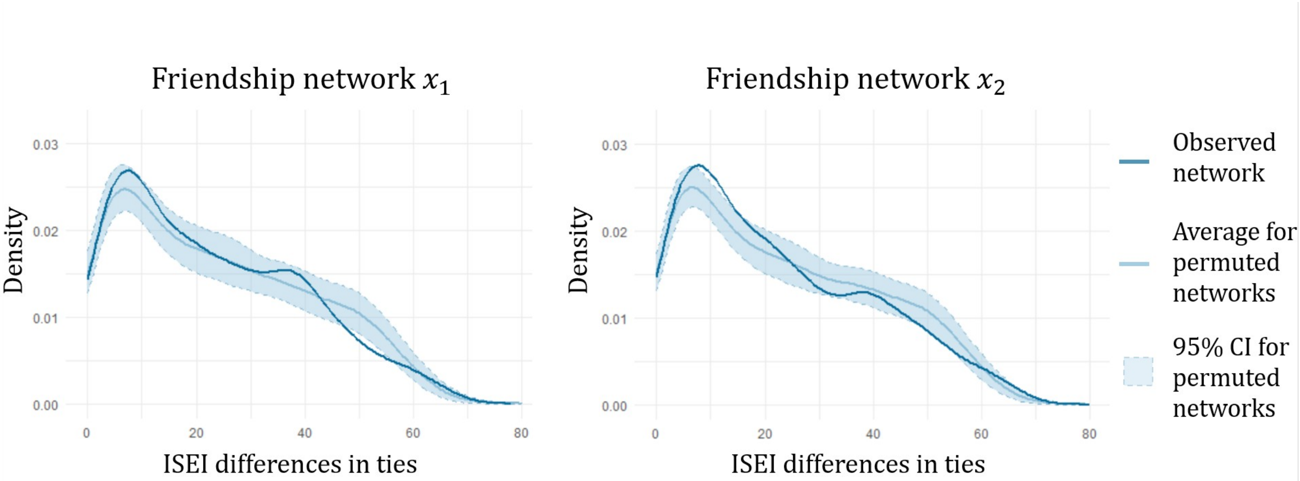
In this section, we address our first research question and assess the presence of socioeconomic homophily in the reported friendship networks and the observed meal partitions, using permutation tests.

4.1.1. Friendship Networks

To test whether the friendship networks x_1 and x_2 contain homophily in terms of socioeconomic origin, we compare the observed average absolute difference in ISEI scores between friends to the same average computed for randomly permuted networks. This permutation test suggests that ISEI differences in both networks are lower than expected at random, conditioning on the network structure, but that these differences are not statistically significant (for x_1, observed value: 23, average in permutations: 23.8, p-value: 0.37; for x_2, observed value: 22.9, average in permutations: 23.9, p-value: 0.24). Complete tests are presented in Appendix 8C.

Figure 8-3 further shows the density distributions of ISEI differences in the ties of x_1 and x_2, compared with their expected distributions under random permutations. Visually, it appears that the distribution of ISEI differences closely approaches the one expected at random, with minor fluctuations. All in all, we find no conclusive evidence for socioeconomic homophily in friendship networks during the camp.

Figure 8-3: Permutation Tests for Socioeconomic Homophily in Friendship Networks



Note: Density functions for ISEI differences in the ties reported in x_1 (left) and x_2 (right). Dark blue lines correspond to observed densities. Light blue lines correspond to average densities in permutations and dotted lines around them indicate the 95% confidence interval.

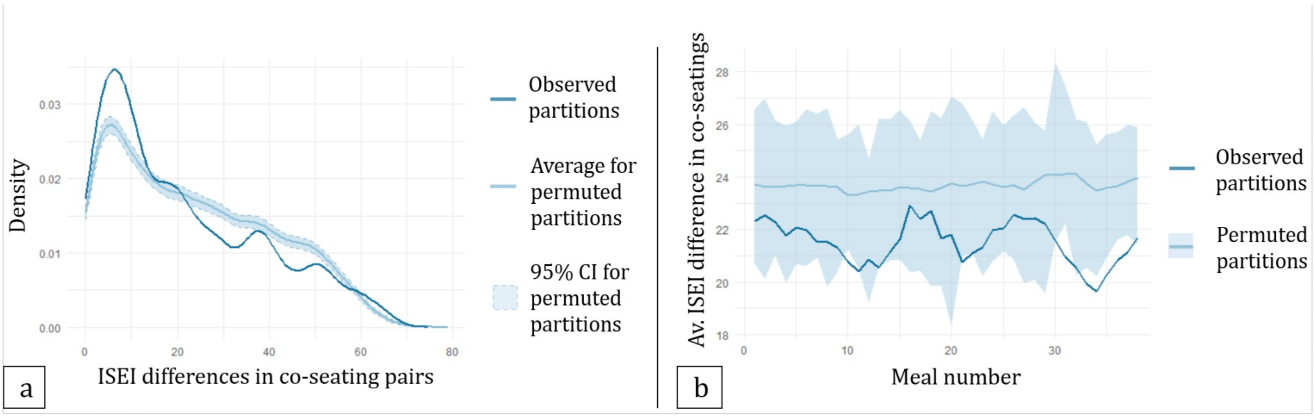
4.1.2. Meal Partitions

For each meal, we compare the observed average absolute difference in ISEI scores for co-sitting ties, to this same average in permuted partitions (average value observed per meal: 21.5, average per permutation and per meal: 22.4, average p-value: 0.31). Among all p-values calculated, 6 of them (out of 38 meals, i.e. 16%) are below the 0.05 significance threshold, meaning that ISEI differences in co-sitting ties are only lower than expected at random, conditioning on group structure, for these 6 meals. We also carry out this test for the sum of ISEI score ranges in groups (average value observed per meal: 30.5, average per permutation and per meal: 31.5, average p-value: 0.48). We find only four meals (10%) with p-values lower than 0.05.

The same test is used across all meals at the same time. This time, we find an average absolute difference in ISEI scores significantly lower than expected at random (observed value: 21.4, average per permutation: 23.7, p-value: 0.04), but the same test is not significant for average score ranges (observed value: 29.8, average per permutation: 31.2, p-value: 0.34). Groups are therefore (a little bit) more homogeneous overall than expected at random, with some variation across meals. Furthermore, we apply this same test to the subset of meals that will be used for the estimation of the ERPM (dinner every three days for a total of seven meals). The results are virtually the same then for all the meals (observed ISEI absolute difference: 21.4, average in permutation: 23.8, p-value: 0.034). Once more, complete tests are presented in Appendix 8C.

Figure 8-4 presents the density functions for ISEI differences in meal-sharing averaged over the entire camp, as well as the evolution over time of the average ISEI difference per meal. Figure 8-4a shows that the average density of ISEI differences over all meals fluctuates around the average observed in the permutations, with a higher density for very low differences (i.e. strongly homophilic co-sittings). As for potential variation over time in homophily, Figure 8-4b shows that the sum of ISEI differences in each observed meal is in general lower than the same sum calculated for the random permutations (~10% lower on average) but this difference varies across the stay. It becomes significantly lower than the 95% confidence interval in a few meals toward the end of the stay but remains within the 95% confidence interval for most meals.

Figure 8-4: Permutation Tests for Socioeconomic Homophily in Meal Partitions



Note: Left: average density function for ISEI differences in co-sitting ties over all meals, for observations (dark blue line) and permutations (light blue line). Dotted lines indicate the 95% confidence interval for permutations. Right: sum of ISEI differences for each meal in observed co-sitting ties (dark blue line) and permutations (light blue line). The 95% confidence interval for permutations is indicated by the light blue area. Curves are smoothed by averaging over 5 meals.

To better understand the magnitude of the homophily in the partitions, we can compare the average difference in observed co-seating ties to the theoretical minimum value of this difference. If we maintain the group structure but optimize the sittings in order to minimize ISEI differences, we get a value of 2.9. Remember that the observed value is 21.4, and the average value under permutations (i.e. assuming no homophily whatsoever) 23.7. The difference between observed and permuted values is therefore approximately 10% of what it could be in the most extreme case[203]. Furthermore, we can quantify the over-representation of ties with very low differences by comparing the proportion of ties with a ISEI difference below 20 (i.e. strongly homophilic) in the permutations (47.8%) to the one in the observed data (56.6%). All in all, these numbers indicate that the magnitude of homophily in the data can be considered relatively small.

The answer to the first research question is therefore different for the two measures of social relations. Although we find no evidence for socioeconomic homophily in friendship networks, we do find evidence for homophily in the groups sharing meals, albeit a relatively small one.

203 (23.7-21.4)/(23.7-2.9) ~= 0.1106.

4.2. Expressed Dispositions toward Similar Socioeconomic Backgrounds

While the previous analyses aimed to detect the presence of socioeconomic homophily as an outcome, we now turn to identifying the role of different homophily-inducing processes in the evolution of friendships and meal partitions. Even though permutation tests showed no homophily in friendship networks, various processes could have acted as suppressors, compensating or over-shadowing the effect of expressed dispositions. As for meal partitions, the observed homophily is not, in itself, indicative of homophilic selection on the part of individuals. Therefore, we now use network and partition models to test for the presence of expressed dispositions for similar socioeconomic backgrounds, controlling for other simultaneous processes.

4.2.1. Preliminary note on chosen foci

As outlined in Section 2, chosen foci can be important contributors to homophily. This could be the case in the camp if bedrooms and activities were somehow segregated in terms of socioeconomic backgrounds. However, we find no evidence of such a segregation. Regarding bedrooms, permutation tests indicate that the average absolute differences in ISEI scores among roommates is not significantly lower than in randomly permuted partitions (observed value: 23.1, average in permutations: 23.7, p-value: 0.68). Performing a similar test by permuting nodes in the bipartite network of the paid activities leads to the same result: the average difference in ISEI scores between the participants to the same activities is not significantly lower than expected in randomly permuted networks (observed value: 22.6, average in permutations: 23.8, p-value: 0.12). From these tests, we cannot conclude that the composition of bedrooms and of paid activities was segregated in terms of socioeconomic origins during the camp (note that observed values are still slightly inferior to permutation ones, albeit not significantly so, which means there might still be some slight segregation effect that we lack the statistical power to confirm). Therefore, it seems very unlikely that they could act as inducers of socioeconomic homophily. Nevertheless, because they could potentially act as a suppressing effect for homophilic dispositions, we still include them in the analyses. Since the process of forming rooms and activities is potentially endogenous to the formation of friendships or group tables, we include the effects of rooms and activities in separate models.

4.2.2. Friendship Networks

Table 8-2 presents the estimated parameters related to the ISEI variable for the SAOMs described in Section 5.3. Model 1 tests whether individuals were more likely to be tied to others with similar ISEI scores, while Model 2 tests whether individuals were both more likely to create and to maintain ties with similar others (endowment effect). The effect of chosen foci is added separately in Model 3, because of the endogeneity issue mentioned above. The other model parameters that are not presented here control for endogenous network mechanisms and effects related to other individual attributes. These parameters can be found in Appendix 8D, along with time heterogeneity tests and goodness of fit plots.

Table 8-2: Results for the SAOM Estimation on Friendship Networks

	Model 1		**Model 2**		**Model 3**	
Effect	est.	s.e.	est.	s.e.	est.	s.e.
ISEI receiver	-0.004	(0.0027)	-0.004	(0.003)	-0.004	(0.0024)
ISEI sender	-0.021	(0.013)	-0.021 *	(0.010)	-0.023 *	(0.010)
ISEI similarity	0.24	(0.20)	0.11	(0.27)	0.31	(0.20)
ISEI similarity (endowment)	-	-	0.99	(1.10)	-	-
Same bedroom	-	-	-	-	1.32 ***	(0.24)
Same activities (period 2)	-	-	-	-	0.29	(0.18)

*p<0.1; **p<0.05; ***p<0.01

Note: Omitted structural effects are density, reciprocity, transitivity (gwesp), transitivity (gwesp) * reciprocity, triads $i \rightarrow k \leftarrow j$ (gwesp), indegree popularity, outdegree popularity (squared), indegree activity. Omitted attribute effects are ego, alter and similarity effects for age and gender. Time dummies are added in period 2 for density, indegree activity and gender sender.

In both Models 1 and 2, we find no significant effect of the similarity in ISEI scores in explaining the observed ties. In Model 2, we see that controlling for the interaction between homophily and maintaining ties (i.e., using an endowment effect) does not yield any significant parameter either.
Only one term related to ISEI scores turns significant in Model 2, which shows a potential tendency for individuals to nominate more alters with lower scores. Moreover, Model 3 indicates that controlling for chosen foci does not yield a significant effect for ISEI similarity either, confirming that these foci neither reinforced nor suppressed the effect of expressed dispositions. All in all, we find no evidence for expressed homophilic dispositions toward same-background peers in the friendships reported during the stay.

4.2.3. Meal partitions

We now turn to the estimation of the ERPMs presented in Section 5.4. Estimated parameters for three different models are presented in Table 8-3. Model 1 includes endogenous group effects, attribute-related effects, the repetition effect and the effect of past friendships. In Model 2, we add the interaction between the repetition of previous co-sittings and ISEI homophily (similar to the endowment effect in the SAOM). Finally, Model 3 contains the effect of bedrooms. Goodness of fit plots are in Appendix 8E.

Table 8-3: Results for the ERPM Estimation on Meal Partitions

	Model 1		**Model 2**		**Model 3**	
Effect	est.	s.e.	est.	s.e.	est.	s.e.
Number groups	-6.39***	(1.11)	-6.26***	(1.07)	-5.07***	(1.09)
Sum squared sizes	-0.099	(0.33)	-0.075	(0.34)	-0.081	(0.33)
Number isolates	4.14***	(0.87)	4.05***	(0.85)	3.48***	(0.86)
Girls popularity	0.054	(0.046)	0.045	(0.046)	0.070	(0.045)
Same gender ties	-0.023*	(0.094)	-0.22*	(0.097)	-0.24**	(0.091)
Same gender groups	3.49***	(0.62)	3.46***	(0.61)	3.14***	(0.60)
Higher age popularity	0.0012	(0.024)	0.0021	(0.025)	0.0045	(0.024)
Age differences	-0.075	(0.059)	-0.072	(0.059)	-0.070	(0.054)
Age ranges	-0.48*	(0.24)	-0.48*	(0.24)	-0.49*	(0.23)
Higher ISEI popularity	-0.0010	(0.0017)	-0.0011	(0.0016)	-0.00061	(0.0016)
ISEI differences	0.00012	(0.0033)	-0.0020	(0.0037)	-0.0039	(0.0037)
ISEI ranges	-0.018	(0.013)	-0.017	(0.013)	-0.0011	(0.013)
Number meals	0.11***	(0.0094)	0.092***	(0.019)	0.076***	(0.017)
Anterior friendship ties	0.32	(0.17)	0.35*	(0.17)	0.079	(0.17)
Number meals X ISEI diff	-	-	0.00078	(0.00075)	0.00030	(0.00068)
Same bedroom ties	-	-	-	-	0.66***	(0.095)
Max. abs. convergence ratio	0.69		0.75		0.66	

*p<0.1; **p<0.05; ***p<0.01

Because of the novelty of the longitudinal ERPM framework, we briefly describe the values of parameters not related to ISEI scores. First, the structural effects can be interpreted as follows. The negative parameter for number of groups in all models suggests a tendency to form less groups (in other words, larger groups) than what would be expected on average in a random partition. The positive parameter for number of isolates indicates a tendency to observe isolated individuals.

Turning to gender effects, we find no significant parameter for the tendency of groups including more girls to be smaller, but we do find a significant tendency toward the formation of homogeneous groups in terms of gender. The parameter 3.49 in Model 1 can be interpreted as a log-odds ratio and indicates that a fully homogeneous group (containing either only girls or only boys) is about 33 times more likely than a mixed group[204].

Moreover, the purely dyadic tendency to observe co-seating ties of the same sex is also
statistically significant in all models and the direction of its parameter suggests that groups with fewer same-gender ties are more likely, controlling for the previous tendency to form non-mixed groups (though the effect is much smaller than that of group homogeneity). This indicates that if a group is mixed, it tends to be balanced rather than containing a majority of either girls or boys[205].

Similarly, we find a significant tendency toward groups with low age ranges in the three models, but no significant parameter for the absolute difference in ages within groups (net of the group's range). For example, the parameters of Model 1 indicate that a group where everyone has the same age is almost seven times more likely than a group with a range of 4[206]. However, a group with children aged 10, 11, 12, and 13 is equally as likely as a group with two children aged 10 and two others aged 13.

Regarding time-related effects, we see in each of the three models a significant impact from the number of meals shared in the past, meaning that the more attendees ate together up to a certain point, the more likely they were to eat together again after that point. In other words, teenagers tend to establish constant meal partners. Additionally, we see in Model 2 that friends from past camps kept eating together: having one more pair of previous friends eating

204 exp(3.49) = 32.79. Mixed groups are very rare indeed! Note that this is partly due to how group meals were coded: a *group* is not the same as a *table*, because different groups of teenagers could eat at the same table (identification of the groups was evaluated qualitatively by the observer, based on e.g. whether the teenagers arrived or left the refectory at the same time, kept seats for one another if some arrived later, or talked together while eating). When boys and girls ate at the same table, which was less rare (though not frequent either), it was generally due to two clearly distinct groups occupying the same table (which itself tells a lot about how cross-gender relations are mostly managed and organized through groups at this age – but this is off-topic here).

205 Which confirms what was mentioned in the previous note, i.e. that cross-gender ties are usually due to groups of boys and groups of girls coming into contact as groups.

206 exp(-0*0.48) / exp(-4*0.48) = 6.82.

together makes a partition 1.4 times more likely, other statistics being constant[207]. However, this effect is not statistically significant in the two other models. In Model 1, the effect size is virtually the same, meaning this probably has to do with statistical power (cf. our previous discussion of statistical power and convergence issues, section 3.4.2). By contrast, in Model 3, the effect size is greatly reduced, following the introduction of the effect of bedrooms; this indicates that, for the most part, former friends kept eating together only to the extent that they were in the same bedroom.

As regards our main variable of interest, the effect linking ISEI scores and group size ('Higher ISEI popularity') is not statistically significant, which means adolescents do not eat in larger or smaller groups depending on their socioeconomic background. More importantly, the two effects meant to capture the selection of same-background peers are the dyadic homophily effect ('ISEI differences') and the group range effect ('ISEI ranges'). We see that none of these are statistically significant in any of the models, and they have very small effect sizes as well. Moreover, in Model 2, the interaction of the difference effect with the number of meals shared in the past (pseudo endowment effect) is not significant either. We therefore find no evidence that the total socioeconomic distance in groups (captured by the difference effect) and the range of extremes in the group (captured by the range effect) impacted the selection of meal partners during the camp, once other factors are accounted for.

Finally, in Model 3, we find a strong and significant effect of room composition, but its addition does not affect our main result regarding the expression of homophilic dispositions. We can nevertheless notice a small decrease in the parameter capturing the number of same gender groups, probably due to the bedrooms being non-mixed.

Note that in these models, the maximal convergence ratio is relatively higher than what would be expected for better known models (for ERGMs, convergence ratios should be kept below 0.1). This is due to ERPM being a new class of model, such that the estimation algorithm is not yet optimized as well as that of established methods (running time is also longer for ERPM at this stage, which makes it harder to explore the parameter space to find a better-converged model in a reasonable time). Consequently, the significance tests should be interpreted with caution, as they might not be fully reliable. However, a number of points indicate that our results are robust concerning socioeconomic selection specifically. First, several estimations of the model were run and yielded similar results, with slightly different parameter values and convergence ratios. Second, the statistics that are not perfectly

207 $\exp(0.35) = 1.42$

converged are only related to the number of anterior friendships in the groups and the number of meals shared in the past, but statistics related to sizes and attributes (including the ISEI scores) show good convergence. Third, even if the effects related to homophilic selection were to turn significant with better converged models and/or more statistical power, their magnitude would likely be very small. Details about the expected statistics in each model, convergence ratios, and some goodness of fit statistics related to sizes and ISEI values are in Appendix 8E.

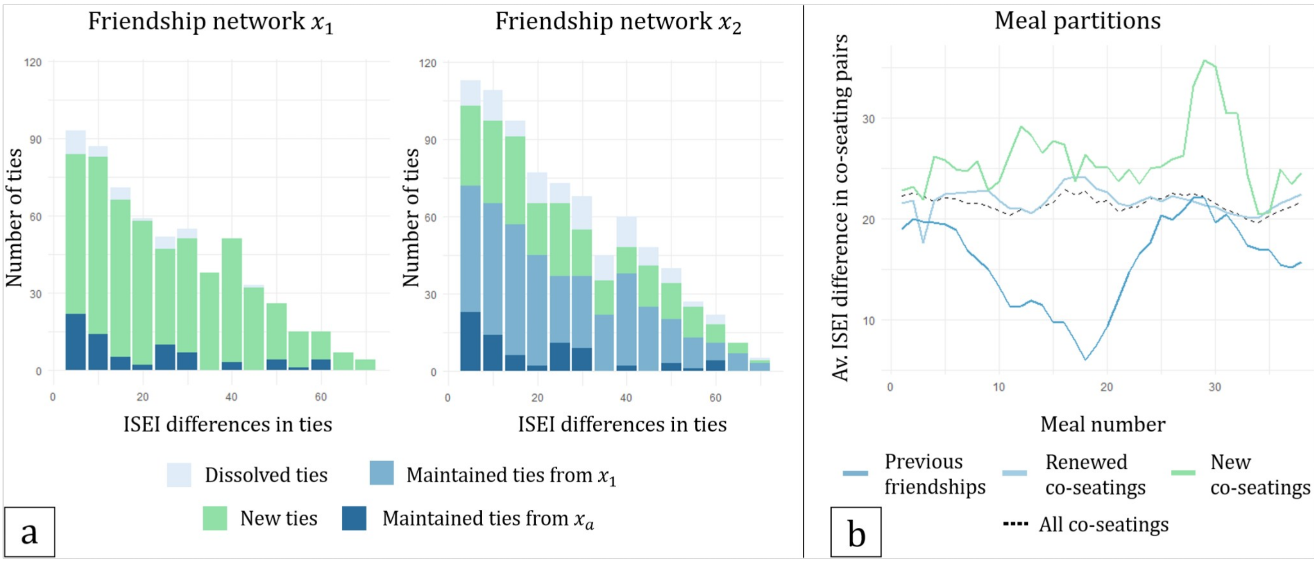
To conclude, the answer to our second research question is the same for both measures of social relations. There is no evidence for expressed dispositions toward socioeconomically similar peers, either in declared friendships or meal partitions. With that said, we also saw that the observed levels of socioeconomic homophily (as an outcome) differed between the two measures. For the friendship networks, there was no evidence for homophily, which makes it relatively logical that there would not be homophilic selection either. For meal partitions, however, we did observe a relatively clear (though moderate in size) homophily. The absence of an estimated selection effect in the ERPM therefore suggests that this homophily is explained by another process beyond disposition-based selection. In that regard, the most likely inducer is the role of the former friendships inherited from past camps, which is the subject of the next section.

4.2.4. The special case of friendships anterior to the stay

The network x_a of prior friendships is essentially the starting point of all relations formed during the camp. To test whether it exhibits some homophily, we use a similar permutation test as the one used for x_1 and x_2 in Section 6.1. This permutation test suggests that ISEI differences in x_a are significantly lower than expected at random, conditioning on the network structure (observed value: 16.4, average in permutations: 23.8, p-value: 0.01). Moreover, it seems that this homophily cannot be explained by past meeting opportunities alone: there is homophily in past friendships even when accounting for the structure of co-attendance to the same camps in the past (see Appendix 8A).

This observation is somewhat surprising, since we find no socioeconomic homophily in the friendships reported during the camp: we could have expected the homophily in prior friendships to carry on into later friendship networks. On the other hand, this homophily of anterior friendships may contribute to explaining the homophily observed in the meal partitions, given that we found no hint of direct homophilic selection. In the following analyses, we therefore try to understand how this initial homophily was reducedin the friendships developed during the stay, and whether it played a role in the homophily found in the meal observations.

Figure 8-5: Evolution of Anterior Friendships in the Networks and Partitions during the Stay



Left: number of maintained, newly formed and dissolved ties in x_1 and x_2, depending on their ISEI difference.
Right: average ISEI difference in co-sitting pairs for each meal depending on the type of tie. “Previous friendships” indicate dyads that were already friends prior to the camp, “renewed co-sittings” indicate dyads that were not prior friends but who already ate together at an earlier meal (relative to the meal at which the co-sitting is considered), and “new co-sittings” indicate dyads that were not prior friends and who eat together for the first time at the considered meal. Curves are smoothed by averaging over 5 meals.

Figure 8-5a (left) shows the distribution of ISEI differences for the friendship ties that were either maintained, dropped, or newly formed in$x_1$andx_2. We first observe that anterior friendships (in dark blue) are mostly preserved in both in-camp networks, indicating that homophily did not disappear because these friendships did. Another explanation could be that newly formed ties (in green) had relatively high ISEI differences, while dropped ties (in grey) had relatively low differences – which would essentially indicate a heterophilic pattern of friendship formation (newly formed ties are less homophilic than deleted ties). However, this tendency is hardly visible in Figure 8-5a. Moreover, the parameters of Model 2 presented in Table 8-2 above do not show any evidence for heterophilic dispositions in the creation of ties, nor in the dropping of ties.

To sum up, we find no evidence that anterior friendships dissolved at a high rate, nor that newly formed ties were particularly heterophilic. Therefore, it seems that the homophily initially observed in $x_a$did not disappear as much as it was "diluted" in the high number of newly formed ties observed in$x_1$andx_2. New ties did not exhibit any homophily (though no heterophily either), such that anterior friendships were simply too few compared to these new ties to entail significant homophily at the aggregated level.

As for co-sitting ties, Figure 8-5b (right) shows that ties between previous friends have lower ISEI differences on average than other ties, and particularly so in comparison to new co-sitting ties. To check whether there is still socioeconomic homophily in the co-sittings beyond that due to prior friendships, we conduct a permutation test on all the meal partitions at once – similar to that of section 4.1.2 (Figure 8-4a) –, but this time excluding the dyads of individuals that were already friends prior to the camp. The statistic tested is the average absolute difference in ISEI scores over all co-sitting ties observed during the camp. For the full sample, the observed value was 21.4, and the average obtained in permutation 23.7 (p-value: 0.04), indicating socioeconomic homophily. This time, excluding the dyads of prior friends, we find an observed value equal to 21.4 and an average in permutation of 24.0 (p-value: 0.20). Therefore, the gap between the observed and permuted partitions has shrunk a little bit and, more importantly, is not statistically significant anymore. This suggests that a good part of the socioeconomic homophily found in the meal partitions was due to the homophily of the prior friends.

Altogether then, prior friendships, which were strongly homophilic, have mostly carried over to both friendship networks and meal partitions. However, the resulting socioeconomic homophily was very slight in the friendship networks and seems to have gotten "diluted" in the many newly formed, non-homophilic ties. By contrast, in the meal partitions, this seems to have been sufficient for a discernible homophily to emerge at the aggregate level. Why is the homophily-inducing effect of prior friendships more pronounced in the meal partitions than in the friendship networks? This is difficult to know for sure, but one explanation may be that the endogenous group processes at work in the meal partitions have a stronger aggravating impact on an initial slight homophilic imbalance (i.e. due to prior friendships), compared to the endogenous network processes that underlie the friendship networks. This would also explain why the estimated ERPMs manage to reach the correct amount of socioeconomic homophily without needing to postulate an effect of homophilic selection (since ERPM accounts for these group processes). Alternatively, there may be a slight effect of homophilic selection in the meal partitions, and not in the friendship networks, that the longitudinal ERPM failed to detect by lack of statistical power, but that could account for a small excess of socioeconomic homophily beyond the mechanical consequence of prior friendships.

Finally, a perhaps simpler explanation may be that there is also a slight socioeconomic homophily in friendship networks, however, with only two observations we simply lacked the statistical power to confirm this, unlike in the meal partitions for which we have much more data. Looking back on Figure 8-3a (section 4.1.1), one can see that the distribution of ISEI differences in friendships correspond to a slight socioeconomic homophily (more low-difference ties and less high-difference ones than the average of permutations), albeit non-significant; and, comparing it with Figure 8-4a, that the confidence interval in the permuted networks is much larger than for meal partitions.

5. Discussion

Adolescents' declared friendships exhibited no socioeconomic homophily during the camp. This result did not change when accounting for chosen foci (bedrooms and activities, which were themselves not segregated in socioeconomic terms) nor network and group endogenous mechanisms. This is even more noteworthy given the state of initial friendships inherited from past camps, which were distinctly homophilic; meaning that the camp's relational structure was initially tilted toward socioeconomic homophily. Yet this early homophily had disappeared by the middle of the stay, being essentially diluted into newly formed non-homophilic ties. Therefore, we find no evidence of expressed dispositions toward socioeconomically similar peers in students' subjective reports of their social relations.

On the contrary, socioeconomic homophily was observed in meal-sharing among participants, though it appeared to be of a moderate, if not small magnitude. However, most, if not all of this homophily can be accounted for by the inertia of past friendships inherited from previous camps. Therefore, there is no clear evidence that teenagers tended to select same-background peers as meal partners – and, even if they did, the magnitude of this effect ought to be extremely small. Together with the fact that chosen foci (bedrooms and activities) were not socioeconomically segregated, and that adolescents enjoyed a large amount of freedom in managing their sociability during the stay, with little formal constraints or direct interventions from adults, this suggests that there was little to no expressed dispositions toward same-background peers in teenagers' sitting behaviors.

Additional analyses conducted in a similar camp that took place two years earlier, though less in-depth because of data limitation, suggest a similar pattern, that is, no effect of expressed dispositions in meal-sharing (or a very slight one; see Appendix 8F). This supports the idea that our main result is not just accidental or one-of-a-kind but could be observed quite robustly over several camps that share the same contextual properties. Moreover, in this 2017 camp, past friendships were not socioeconomically homophilic, unlike in the 2019 camp, and the meal partitions during the camp exhibited no clear homophily either – thus supporting the idea that the homophily of 2019 is explained by the friendships from past camps.

Altogether, the main conclusion to be drawn from these findings is that expressed homophilic dispositions appear extremely weak during the camp, if not null, having little to no discernible impact on friendship networks and meal partitions (and taking into account issues of statistical power that might lead us to miss a very small effect). We now discuss potential interpretations of this finding (5.1), before turning to the implications it may have for school-year contexts (5.2.)[208].

5.1. Latent Dispositions and Contextual Moderators

As we just mentioned, disposition-driven selection toward socioeconomically similar peers appears very weak during the camp, if not null. Does this mean that adolescents coming to the camp had no latent dispositions toward socioeconomic homophily? Not necessarily. A first explanation for the weak selection found during the stay might pertain to its short duration. To the extent that dispositions are not just expressed at the level of friendship formation, but also through different rates of survival of relationships – as early friends gradually get to know one another (Tuma and Hallinan 1979) – the relational structure after three weeks might not have yet reached an equilibrium regarding socioeconomic homophily.

This might also explain why we observed socioeconomic homophily in past friendships: friendships among socioeconomically similar peers might be more resilient across summers, either through the ability to stay in contact during the year, recollection bias from one year to the next as to who was a friend, or the willingness and ability to coordinate

208 Note that the entire discussion of this section is meant to rely only on the literature and on the results obtained in the camp (as it was written with the purpose of becoming a stand-alone article). The connection to the rest of the PhD and to the four schools from my main sample are discussed in the conclusion of the chapter.

to attend the same camp the following summer. Nevertheless, this hypothesis cannot account for the fact that there was no socioeconomic homophily in past friendships for the 2017 camp (Appendix 8F). More generally, it is difficult to guess what happened in past campsas we could not observe them directly. In addition, not all the adolescents at the 2019 camp had attended previous camps, so the sample size for these treatments is particularly small (19 out of 60 adolescents had never been to a camp before).

Beyond a potential lack of time for homophilic processes to develop, another explanation for this weak homophilic selection may lie in the contextual moderators operating in the camp. We mentioned already the fact that individuals' homophilic dispositions could be heightened, or on the contrary toned down, by contextual moderators. In that sense, while individuals can be assumed to have certain latent dispositions, what we effectively get to observe in empirical contexts are expressed dispositions, already filtered through contextual elements.

In that regard, the camp's short duration, insulation from attendees' usual life, and overall positive climate, might encourage some adolescents to drift away from the social roles that they are used to, using the camp as an opportunity to "be someone else" (figuratively speaking). Indeed, teenagers are typically going through a period of self-definition and search of personal identity (Dubet and Martuccelli 2014), which may imply the experimentation of different behaviors in different contexts. In that regard, the stay's positive climate (most of the teenagers declared that they enjoyed their vacation at the end of the three weeks), as well as the fact that the camp is short and has no direct repercussion on the rest of their lives, might make it more reassuring to perform such experimentation compared to the context of a schoolyear. This could also be facilitated by the medical nature of the camp, which entails strong emotional investment for many attendees and is one of the few times of the year where they can share with others their experience of the disease. As a result, youths might be more willing to adjust their behavior to whatever social group they are part of during the stay, and to experiment different sociability practices, which may in turn result in less homophily.

Furthermore, the camp is also characterized by the special role of the animation staff. Indeed, though not directly constraining adolescents' relational choices, they deliberately attempted to mitigate conflicts or rivalries among them – for e.g. by acting as mediators after a quarrel, or by purposely spending more time with teenagers that seemed to have trouble interacting with peers. In doing so, adults may have facilitated positive contact among dissimilar individuals, essentially "greasing the wheels" of the camp's relational structure.

Further supporting this interpretation, McFarland et al. (2014) have argued that homophily is generally weaker in schools with a positive climate (in the sense of increased students' well-being), as individuals feel less threatened and defensive in their pairing choices.

Having observed a single camp, we cannot rigorously isolate the effect of these potential moderators, which also means that we cannot be sure about the state of adolescents' latent dispositions toward same-background peers. These may be nonexistent, or they may have been inhibited in this particular context. Nevertheless, if attendees had strong, conscious and systematic inclinations toward same-background peers –e.g. inter-group prejudices, discrimination, or a drastic discrepancy in tastes, values and interests – then we would expect those to be manifested regardless of moderators, especially considering that one key property of this camp is the great autonomy that the teenagers enjoyed in organizing their social lives. That adolescents from different backgrounds could name one another as friends and eat together seems to indicate that these latent dispositions are not rigid nor definitive obstacles to positive social interaction; the very fact that they could be toned down by contextual moderators is an indication of their relative plasticity. In other words, our results preclude the idea that the inclination toward socioeconomically similar peers is ubiquitous and insurmountable, though it may still exist in a more subtle or situational form.

5.2. Implications for school-year contexts

The aim of our design was to find a context that made it easier to isolate the effect of adolescents' dispositions, compared to school-year situations where many intertwined factors operate simultaneously. Assuming the adolescents in our sample are similar enough to those found in other mixed contexts, what would our results imply for the socioeconomic homophily typically observed at school? Two (non-exclusive) interpretations are possible. First, adolescents' intrinsic dispositions toward same-background peers may be weaker than what in-school studies seemed to suggest, which would point to the role of imposed foci and adult intervention in explaining the homophily found in these contexts. Second, these intrinsic dispositions may be highly sensitive to contextual moderators, such that they would be weakly expressed in the camp's context, but much more saliently expressed in certain schools.

Supporting the first hypothesis, we know that different-background peers typically live in different neighborhoods, attend different curricular options at school, or participate in socioeconomically segregated activities. Moreover, parental networks and inter-generational closure, as well as (some) parents' active avoidance of socioeconomic mixing, can also impact youths' social lives. Finally, network endogenous mechanisms are generally expected to act as aggravating factors of homophilic patterns. Therefore, even a modest initial amount of homophily, induced by structural opportunities, might then "snowball" into significantly higher levels. Altogether, this may be enough to account for the socioeconomic homophily found in several contexts, without needing to postulate much homophilic dispositions – if any – on the part of individual students.

As for contextual moderators, ethnographic studies suggest that the academic competition typically found at school heightens tensions and rivalries among students with different school goals and ethos (Cousin and Felouzis 2002), and thus, on average, between students from different socioeconomic backgrounds. Additionally, it is likely that schoolteachers generally do not act as mediators and facilitators in peer-relations as much as the camp's animation staff did. In the French context particularly, adults typically believe that they are not to meddle into students' sociability (Rayou and Zanten 2004); and many students do not trust their educators or sometimes antagonize them due to student-teacher conflicts (Merle 2015), which can only impede teachers' efforts to regulate peer-to-peer relations.

6. Conclusion on the Summer Camp Study

In this study, we have looked at socioeconomic homophily in a context that is only rarely explored, that of summer camps for teenagers. The unique nature of the chosen camps – medicalized and with a liberal pedagogy – let us study homophilic dispositions under ideal settings: structural opportunities were equalized at the start of the stay, and most contextual factors believed to induce socioeconomic segregation during school year (school, parents, residential segregation, or formal activities) were removed. Furthermore, the collection of both subjective and objective data allowed for a thorough modeling of adolescents' relations with one another. In particular, we extended the existing cross-sectional ERPM framework to a longitudinal specification; to the best of our knowledge, this is the first-time repeated group interactions are modeled as a set of time-dependent partitions, which opens the door for novel ways to study homophily and other relational processes. The fact that we observed no socioeconomic homophily in reported friendships, but found some in table partitions, also points to the added value of observational measures and of group partitions.

This design, however, is not without limits. First, we could not separate internalized dispositions from contextual moderators, due to the fact that we observed but a single context. This monographic approach, together with the uniqueness of the chosen camp and the small size of our sample, also limits the possibility for generalization; under different conditions, adolescents' sociability may imply widely different social processes. Moreover, our sample, although socioeconomically diverse, was characterized by an over-representation of upper-background adolescents. Different socioeconomic compositions might have different implications in terms of relational processes and homophilic attitudes (for example one might expect socioeconomic distance to be more salient in case of a polarized distribution, i.e. many lower- and upper-backgrounds but few middle-class ones).

Altogether, the studied adolescents exhibited high levels of socioeconomic mixing, as declared friendships were not significantly more likely among same-background respondents, and as the socioeconomic homophily that we did observe in meal-sharing behaviors remained relatively weak. Moreover, this latter homophily was largely a consequence of homophilic friendships from previous camps carrying over to the studied one, rather than to homophilic processes occurring within the camp itself – as to the reason why these past friendships were homophilic, this is something that cannot be answered with the data at hand. Therefore, the present study suggests that socioeconomic mixing in teenagers' peer groups is, at the very least, possible. The observed adolescents did not appear to have strong and consistent intrinsic dispositions toward same background peers; which in turn suggests that structural opportunities, as well as contextual moderators, are essential in explaining the amount of socioeconomic homophily that may emerge at school or in other contexts.

General Chapter Conclusion

The study design presented throughout the chapter was meant to answer two of the three methodological challenges that I mentioned in the opening of the chapter; and it offers valuable hints as per the third one as well. To reiterate, these challenges were (a) the large amount of missing information about students' sociability and meeting opportunities in most in-school studies (b) multiple, closely intertwined relational processes that feed back into one another over time and jointly generate homophilic network structures and (c) the impossibility of isolating the effect of contextual moderators when considering a single context of sociability. Challenge (a) could be avoided by considering a holistic context with few unobserved chosen foci, over a short time span, and with fine-grained information about the day-to-day evolution of teenagers' relational behaviors. Regarding (b) the structural absence of many potential inducers of homophily in the camp made it much easier to distinguish between the remaining processes. As for (c), the camp itself was a single context, thus making it impossible to rigorously assess the impact of contextual moderators. Nevertheless, an indicative comparison of the results found during the stay with those reported in the previous chapters of the dissertation suggest that these moderators may play an important role in heightening homophilic inclinations in certain school contexts, and, conversely, of toning them down during the camp.

To see this, it is necessary to try and bring together the results obtained in the camp with those presented in previous chapters and based on my in-school sample. To start with, it is important to consider that because the samples are different, there is no guarantee that the pool of latent dispositions in the population is the same: teenagers from the camp and those from the four studied schools may have different traits and be either more or less inclined to select socioeconomically similar peers. Still, for the sake of the argument, let us assume that they are somehow comparable. Following the reasoning of section 5.3. above, the weak or null homophilic dispositions that could be inferred from the camp's data suggested that there are two (non-exclusive) options as per schoolyear homophily: a strong role of imposed foci, and/or of contextual moderators that heighten the impact of latent dispositions. Now considering the results from chapter 5 and, in particular, the contribution scores presented in Table 5-4, it is possible to see which of these options appears most plausible in the four studied schools. Indeed, if imposed foci are decisive in inducing socioeconomic homophily, then the contribution scores associated to these foci (classrooms, place of residence and parental networks) should be relatively large. On the other hand, if contextual moderators play a key role in heightening students' homophilic dispositions, then this should be reflected in the contribution scores that capture the impact of attribute-based selection (academic achievement, ethnic background, socioeconomic background, and cultural tastes).

Looking back on Table 5-4, the contribution scores suggest that both types of factors have some inducing power, with different magnitudes depending on the school. Nevertheless, altogether, scores appear higher for dispositional processes. For example, direct socioeconomic selection accounts for about 30% of socioeconomic homophily in Paris 1 and 2, and same-ethnic selection for about 18% in Savoie 1. By contrast, propinquity processes never account for more than 14% of the total homophily (residential propinquity in Savoie 1). Therefore, the multivariate models used in chapter 5 seem to indicate that attribute-based selection plays a larger role overall than imposed foci in inducing socioeconomic homophily in the studied schools. This suggests that, in school contexts, students do perform some selection of similar peers and, thus, that expressed homophilic dispositions are at work. This is especially true in Paris 1 and 2. As we saw in chapter 7 – for Savoie 1, the role of imposed foci appears more important, though there also seems to be some homophilic selection on properties correlated to socioeconomic origin (ethnicity in particular[209]).

To account for the fact that this same type of disposition-driven selection appeared extremely weak during the summer camp, it is tempting to conclude that contextual moderators indeed play a key role in toning down and/or heightening dispositions in different sociability settings. As mentioned in 5.3. above, academic competition, parental influence, adult regulation of teenager conflicts, and possibly the short duration of the camp, appear to me as likely moderators that could explain this difference.

Once again, this line of reasoning should be seen as speculative, because there is no way to be sure that the populations compared in the two studies are similar enough in terms of their latent dispositions. It is also an issue that the exact type of contextual moderation at work cannot be pointed down. In particular, the influence of time stands conceptually apart from other types of factors such as adult influence, as this is not really a contextual feature strictly speaking – though I am personally inclined to believe that the short duration of the stay is not the decisive factor that would explain its low socioeconomic homophily[210]. Nevertheless, this appears to be a promising hypothesis to explore, and future research should be specifically targeted at studying the impact of contextual factors on the expression of homophilic dispositions among teenagers. This demands us to examine teenagers' peer relations in a large diversity of situations. Summer camps, activity clubs, public spaces (gardens, streets, sport fields...) or even online virtual contexts (e.g. video games) have been understudied so far compared to the school and classroom; yet disentangling psychological from environmental factors demands variation in the observed contexts. Beyond separate studies in different environments, an ideal design would be to follow adolescents attending different contexts, for instance comparing ego-networks at school and during the summer for the same individuals.

209 Note that the summer camp studied in this chapter, despite a fair amount of socioeconomic mixing, was rather homogeneous in ethnic terms (there were some non-native adolescents, but they were a clear minority).

210 Anterior friendships from past camps were socioeconomically homophilic in 2019, but not in 2017. More generally, most of social theory would generally expect homophilic dispositions to decrease in the long run in a mixed population, as individuals get to know one another, develop shared experiences, and eventually socialize each other into similarity (see Flache et al. 2017 for a discussion of assimilative influence in sociological and psychological theory).

Conclusion

In all the studied contexts – the four schools and the summer camp –, some socioeconomic homophily can be found. Its magnitude, however, varies drastically, from barely noticeable effects (summer camp and Savoie 2) to a pronounced relational segregation (Paris 1). It also depends on the type of ties, being generally stronger for more intimate relationships; and it is fairly stable over time, except in one school, Paris 1, where it sharply increases all along middle school years. The driving mechanisms behind these homophilic patterns are multiple, with no single factor generating high levels of homophily on its own. There can be a slight or moderate inclination of students toward socioeconomically similar peers, which is pronounced to varying degrees across contexts, and that pertains to unconscious attraction, cultural similarity or daily micro-foci more than to conscious preferences or outright discrimination. A similar – though seemingly more conscious – inclination exists toward peers with similar attributes correlated to socioeconomic origin, like ethnicity and academic results. Imposed foci also slightly favor homophilic ties on average, due to the socioeconomic segregation of classrooms, neighborhoods, and parental networks. Taken separately, all these inducers are modest, but they mutually reinforce one another, and get further amplified by network and group processes, notably transitivity. Finally, contextual moderators on the one hand, and the structure of the population of adolescents on the other, strongly affect some of these inducing processes. Expressed dispositions toward socioeconomically similar peers, in particular, are either nonexistent or extremely slight in at least two of the studied contexts (Savoie 2 and summer camp), and possibly in a third one as well (Savoie 1 prior to the covid). Altogether then, the emergence of positive ties among different-background adolescents appears neither systematic nor impossible, and schools or summer camps have some latitude in promoting organizational contexts which are either more, or less favorable to socioeconomic mixing.

The above paragraph gives a condensed answer to the five research questions that were raised in the introduction of the thesis. Each of those, however, deserves to be examined in more details. As the reader might remember, they were the following:

(1) *Given a situation of socioeconomic diversity, to what extent is there socioeconomic homophily in teenagers' friendships?*

(2) *Does the strength of socioeconomic homophily vary over time and/or across contexts?*

(3) *Does the strength of socioeconomic homophily depend on the type of relationship that is considered?*

(4) *What are the relational processes through which socioeconomic homophily appears among youth?*

(5) *Do the relational processes identified in response to RQ4 vary over time and/or across schools, and can this account for the trends observed in response to RQ2?*

The next four sections are therefore dedicated to each of these points in turn (RQs 1 and 2 are part of the same section). Moreover, these questions were themselves derived from the two key interests that originally motivated this entire study. First, socioeconomic homophily among teenagers plays an important role in the reproduction of social inequalities, by aggravating schooling inequalities and by contributing to the transmission of class-specific dispositions and resources – in that regard, relational segregation very much resembles spatial segregation. Second, it offers important insights into the processes underlying socioeconomic homophily at large and can therefore help us to understand its emergence and persistence among adults and in wider society. Sections 5 and 6 come back to these two points, in light of what we learned throughout the chapters. Finally, the last section discusses the limits of the present work and attempts to identify promising leads for future research.

1. The Impact of Socioeconomic Origin on Adolescent Sociability (RQs 1 and 2)

1.1. Strength and Evolution in Time of Socioeconomic Homophily

The existence of socioeconomic homophily in diverse French schools is clear. However, its strength appears moderate overall, with the salient exception of Paris 1. In three out of the four schools, as well as in the summer camp, cross-background friendships are by no means uncommon and, at the level of weak ties at least, a dense network of positive relationships connect the different socioeconomic groups. The magnitude of socioeconomic homophily is generally comparable to that of ethnic and academic homophilies, though weaker than gender homophily. Its most notable feature, however, is that it strongly varies across contexts, ranging from almost complete segregation (Paris 1) to extremely slight effects (Savoie 2 and summer camp). If anything, this supports the idea that the processes underlying this homophily are strongly subject to contextual moderation, which is something that I will come back to in a moment. Altogether then, socioeconomic origin does not appear as an insurmountable obstacle to friendship formation among adolescents, although this will strongly depend on the conditions under which mixing occurs.

As for time trends, socioeconomic homophily remains stable over time in all the studied schools but Paris 1, where it increases. The case of the summer camp is more complex: the weak homophily observed during the stay is stable over the three weeks, but former ties inherited from past summers are far more homophilic than in-camp ones. This might indicate that socioeconomic homophily needs some time to unveil itself, in a similar fashion to what was observed in Paris 1; but it may also pertain to other processes, or even be the product of chance, given the relatively small number of adolescents that are considered. Therefore, there does not seem to be any generic trend for socioeconomic homophily over time, except perhaps the fact that it did not decrease in any of the studied contexts.

1.2. Homophily and Centralization

The results mentioned above are different to what part of the literature seemed to suggest, particularly within the Social Network Analysis tradition. In chapter 1, we saw that some authors expected socioeconomic origin to affect students' centrality in friendship networks, more so than to elicit homophily (Hjalmarsson and Mood 2015; Malacarne 2017). For the most part, the pattern observed here is the exact opposite: same-background students are more likely to befriend one another, but there is no clear association of socioeconomic background and network centrality. This is true of three out of the four schools, as well as of the summer camp; the only exception is Savoie 2, which exhibits very weak homophily and where upper-background students emit and receive on average more friendship nominations. Note that, even though the theme of socioeconomic centralization was not directly addressed in my research questions, getting a holistic picture of how friendship networks are structured along socioeconomic lines is important in order to better understand homophily specifically. As was mentioned in the introduction, homophily is one particular feature that we can abstract from the relational structure, but, at the end of the day, this structure is an interdependent whole and should be envisioned as such.

Why, then, is there a centrality effect in Savoie 2 only? In chapters 6 and 7, we found school norms to be more legitimate overall among Savoie 2's students, probably due to the academic selection that the school administration operates in its recruitment. This may explain why high-achieving students have on average more friends, which itself explains a good part of the centrality of high-background students. In the same vein, it is likely that middle- and upper-class cultural orientations are more dominant within this school: in chapter 6, I mentioned the fact that clothing styles seem less oriented toward working-class tastes in Savoie 2 compared to the other schools (less sport brands, more high-rise jeans for girls), as well as the fact that there is more cultural homogeneity overall (lower standard deviation of cultural status scores). In the short then, it seems plausible that working-class students attending this school are characterized by a form of "cultural good will" (Bourdieu [1979] 2016) that creates more agreement between students about legitimate school and cultural norms, to the benefit of upper-background ones. This is not the only possible explanation, though, and dedicated studies would be needed to understand the relationship between schools' socioeconomic composition and centrality phenomena[211].

In any case, this raises an interesting question: is there some sort of trade-off between homophily and centralization in friendship networks? Indeed, the case of Savoie 2 suggests that social and cultural homogeneity may favor the emergence of unidimensional statutory hierarchies – as all students agree on what should be done, the criteria for achieving status are also more consensual; yet this homogeneity is also the most likely explanation for the very weak socioeconomic homophily found in that school. By contrast, in highly diverse schools, it could be that no single group of students is able to endorse the norms and values of the entire student body, thus decreasing hierarchical patterns overall. This interpretation is supported by studies that found peer hierarchies to be associated with group consensus about social norms. In the context of higher education, Kane argues that elite colleges with a diverse population (particularly in terms of nationality) tend to have weaker status hierarchies compared to homogeneous colleges and to colleges with active community-building policies (Kane 2011, p.6). Coleman (1960) also considers that peer popularity strongly depends on one's ability to incarnate the values of the group – which conversely implies that a heterogeneous population with few shared values is less susceptible to hierarchical patterns.

Coming back to the difference between my results and the expectations of (part of) the literature, it may be explained by differences in the levels of socioeconomic and cultural homogeneity of the studied populations. If previous research primarily took place in

211 Another hypothesis could be that upper-background students have more local capital on average, living closer to the school, having siblings in it, or entering middle school together with many of their neighbors or primary school friends; whereas working-class students would more often experience a split between middle school and their other sociability circles. However, this is not supported by the data: in Savoie 2, students with a working-class origin do not on average live further away from their schoolmates, nor do they have less primary school peers in middle school.

socioeconomically homogeneous schools, the association of socioeconomic origin and friendship centrality may be stronger than in the diverse schools studied here.

Alternatively, these diverging results may pertain to national differences: many studies in US schools found that friendship networks were strongly centralized, and that students could clearly identify a group of “popular” peers or a “leading crowd” (cf. chapter 1, section 1.1.2). By contrast, we saw in chapter 5 that centralization is weak in our four schools, with no popularity process in the SAOMs – in fact, in Paris 1 and Savoie 1, there is even a propensity to the de-concentration of ties in the generative models. Moreover, although this is something I lacked the space to develop, the interviews did not reveal any clear group of popular students; and this is coherent with the rest of the ethnographic literature in French schools (here as well, see chapter 1, section 1.1.2). Thus, peer hierarchies as a whole might be less pronounced (or at least less apparent) in the French than in the US cultural context, which would make hierarchies based on socioeconomic origin less salient as well. As a side note, and in relation to the aforementioned idea that hierarchy stems from normative consensus, it may be the case that the differences in hierarchical patterns between the French and US contexts are linked to community-building policies, which are much more common in the US case (e.g. inter-school sport competitions, school-specific uniforms, extra-curricular activities organized by school). Thus community-building school policies could provide a joint explanation for both weaker hierarchy and higher homophily in French schools compared to US ones – although this should be seen as a working hypothesis, which goes far beyond what my data can test for (in particular, it is not clear whether socioeconomic homophily truly is weaker in US schools than in French ones, since we simply lack enough studies on socioeconomic homophily in both countries).

Finally, the lack of social integration associated with working-class backgrounds may only be apparent for the most precarious or destitute households; to the extent that many working-class families in my four schools still have relatively stable jobs and decent living conditions, their children would not suffer the same penalties in terms of friendship making at school.

A final hypothesis raised in chapter 1 concerning the relationship between peer status and socioeconomic origin was that of a conformity pattern, with extreme origins – either very high or very low – being depreciated within peer groups. However, there was no evidence for a curved pattern of association between socioeconomic background and friendship centrality in the studied schools; that is, there does not seem to be a statutory advantage of middle-class backgrounds (cf. chapter 4 as well as Appendix 4E).

2. Type and Strength of Relationships (RQ 3)

Socioeconomic mixing or, on the contrary, segregation, takes place at different levels. Students from different backgrounds can be simple acquaintances, positively disposed toward one another without being close; they can be “pals” or “comrades”, in the sense of a superficial friendship tie; they can be what may be called “local” or “situational” friends, that frequent a common sociability circle and spend time together, but always as part of a broader group or within a single context, such as school; or they can be friends in the strongest sense, i.e. having an interpersonal relationship that has become autonomous from its context of formation (Bidart 1997). Among the latter, we may again distinguish best friends, with whom trust and emotional intimacy is higher than with other friends, confidants, who share secrets with each other, or we may single out those friends who have been introduced into new sociability circles, being presented to other friends or to one’s family.

We saw in chapter 4 that homophily generally gets stronger as one considers more demanding relationships. Moreover, there also seems to be a sharp split between in-school and out-of-school relationships, net of their strength: close friendships at school are not necessarily more homophilic than weak ones outside of it. Most likely, this comes from the greater role of segregated foci outside of school, where residential propinquity, parental networks and extra-curricular activities should play a greater role in friendship formation. This is something that we observed among the friendships with schoolmates that also extend outside of school; and, though it cannot be verified with the data at hand, the chances are that this is even more pronounced for friends that are not in the same school, and/or not of the same age. Thus, diverse schools are probably less homophilic than most other sociability contexts of adolescents. In that sense, they can be seen, to an extent at least, as factors of socioeconomic heterophily within students’ personal networks.

In any case, mixing and segregation are not necessarily exclusive from one another, as they can correspond to different levels of relationships. Though perfect mixing would imply a complete absence of socioeconomic homophily down to the most demanding of these levels, one should not disregard the importance of weak ties either. That youths from different origins can chat, play and get along with each other is already a significant step compared to outright spatial segregation, even if they do not necessarily become close friends. It is an open question as to how much weak ties matter relative to strong ones in peer-to-peer socialization: part of the outcomes that are generally expected from socioeconomic mixing – in terms of reducing schooling inequalities or disrupting the reproduction of class-specific dispositions – may pass by such "light" ties. In that regard, one important result is that, even in the most segregated of the studied schools, namely Paris 1, there is no positive association of socioeconomic distance with the odds of conflict among students. While this does not mean that there is no conflict nor tension associated to socioeconomic origin whatsoever, it does suggest that these remain minor for the most part. The fear, sometimes expressed by upper-class parents to justify their school choice (van Zanten 2015), that socioeconomic mixing may result in systematic violence, exclusion, or dominance among different-background students, does not seem justified[212].

3. Relational Processes Underlying Socioeconomic Homophily (RQ4)

Socioeconomic homophily does not imply the homophilic selection of same-background peers. Unlike gender, ethnicity or academic results, socioeconomic origin is hardly part of students' perceptive frames, at least not in a direct way: parental occupation and education are simply absent from students' discourses about their friends and enemies, and conspicuous consumption is both relatively rare and misleading as per one's actual background, as students from the upper-classes mostly refrain from displaying their wealth. Clothing and musical tastes are probably what come closest to a form of socioeconomic distinction among students. This is, however, relatively imprecise, as there are not clearly bounded cultural genres in which each student could be classified, and as it does not always correspond to one's actual background: an upper-class student may well decide to dress in a "street" way and to listen to rap music – especially considering that, in most schools, working-class tastes seem to be dominant in youth peer groups in general.

With that said, quantitative models did find an effect for homophilic selection in certain schools – a strong one in Paris 1, a moderate to weak one in Paris 2, and possibly a weak one in Savoie 1 as well, although this is much less clear (it only appears between waves 4 and 6, which is the less "reliable" of the studied periods due to the larger time gap and to the covid-related disruptions). Technically, these estimates could simply capture the effect of

212 To be precise, my results show that, within diverse schools, different-background students are not more likely to dislike or bully one another than same-background ones. It may still be the case, however, that the socioeconomic composition of a school affects the overall prevalence of conflicts within it, which could only be investigated by comparing schools with different compositions.

unobserved foci; but the models from chapter 5 integrate many relevant controls, so the confidence that we can have in the existence of a "true" selection effect is reasonably high. Students, it seems, tend to select friends from similar backgrounds,-which implies that they do perceive something of their peers' socioeconomic origin. Beyond cultural genres acting as a fuzzy but correct-on-average frame of perception, this may also pass by subtler forms of matching on certain personality traits that, though not immediately apparent, impact the ease with which a relationship might form or its resilience over time, even without students being aware of it. Moreover, selection effects in quantitative models should also capture what I called "micro-foci", that is, small scale and short time chosen foci that, in practice, are almost indistinguishable from disposition-driven selection (playing football during the break, going out for lunch with classmates, sitting in the same part of the classroom, going to the school library, etc.).

Several other processes beyond socioeconomic selection contribute to the formation of socioeconomic homophily. Imposed foci play a relatively modest role overall. In Parisian schools, residential distance has little impact on friendship formation once other relevant covariates are accounted for, and the role of the former primary school is either non-existent (Paris 1) or quickly decreases in time (Paris 2). Spatial factors matter more in Savoyard schools, especially Savoie 1, which has the most spread-out residential structure and little public transportation. The inducing impact this has on socioeconomic homophily remains modest, however, as residential segregation is relatively small itself (and, in particular, smaller than in the Parisian schools). The classroom distribution has a strong impact on students' friendships but, for the most part, the studied schools have relatively balanced classrooms in socioeconomic terms (except for special tracks). Moreover, homophilic selection also operates on attributes correlated to socioeconomic background, notably ethnicity and academic results and attitudes. This entails either more or less socioeconomic homophily depending on the school but in generally has slightly more impact than imposed foci – ethnic selection in particular. Finally, special tracks notably contribute to socioeconomic homophily, particularly the international section in Paris 1 – as for the pre-vocational track in Savoie 1, it has less of an inducing impact because it is smaller relative to the size of the school, but the processes at work are essentially similar. In both cases, these tracks are very homogeneous socioeconomically, and their impact on friendship networks pertains both to propinquity through imposed foci, and to selection processes due to labeling and prejudices among students.

A word on this notion that there are some attributes “correlated to socioeconomic origin”. By construction, the statistical models used throughout the thesis are agnostic as to where these correlations come from: whether the association of, say, academic results and socioeconomic origin is due to a causal relation (e.g. transmission of cultural capital) or to some unobserved confounder (e.g. IQ of parents) makes no difference in modeling terms. Theory-wise, though, this can have very different implications. In particular, to the extent that social class as a concept has a broader meaning than that of simply occupational groups (though this is the proxy most often used), it can be debatable which properties are merely correlated to it, which are derived from it, and which are part of the very definition of what a social class is. For example, saying something like “*there is no socioeconomic selection once cultural selection is accounted for*” may feel like a strange statement, if one is to treat cultural tastes as being primarily a manifestation of class positions (which is for instance the thesis defended in Bourdieu’s*The Distinction;* Bourdieu [1979] 2016*).* To what extent does the measure of cultural status used in the empirical chapters contain information about the variability of the level of cultural (i.e. how high-brow it might be) not explained by socioeconomic origin? Likewise, to what extent does it merely act as a complementary proxy to socioeconomic origin by capturing nuances of a household’s cultural capital not properly reflected in declared occupations? Similar questions could arise for students’ places of residence – is living in a wealthy neighborhood rather than in social housing not a constitutive part of what we call social classes?

Alas, there is no easy answer to these questions. On the one hand, none of these attributes can be written off as a strict derivative of socioeconomic positions – some working-class individuals do enjoy reading novels or listening to classical music, some upper-class households do choose to live in mixed neighborhoods, and these cannot simply be written off as “false” members of their class. On the other hand, the issues raised by the decomposition of social class into a number of component properties are made apparent by a *reduction ad absurdum*: controlling for cultural tastes, place of residence and academic levels, but also for income, personality traits, clothing, ways of speaking, moral values and political positions, one could certainly reach the conclusion that there is not any socioeconomic selection anywhere – but this would not mean much, as all the elements that give socioeconomic origin its power on relationships would have been taken away. In that regard, the socioeconomic homophily induced by residential propinquity, academic selection or cultural selection should not be thought of as a “fortuitous” or “by-product” homophily: that students from different backgrounds tend to have different results, tastes and living places is itself a mark of the socioeconomic factors that shape their social lives. Altogether then, part of the covariates tested in my statistical models can be seen as operating a form of decomposition of the way in which socioeconomic selection functions, rather than capturing an entirely different process – this is most notably the case of the cultural covariates.

In any case, taken in isolation, none of these correlated factors have a strong impact on socioeconomic homophily. However, they get combined and reinforced by network endogenous processes, which were captured in the quantitative models by the effect of transitivity. In practice, this means that students often find friends through common acquaintances, as well as by socializing in groups rather than in dyadic configurations. Transitivity is a well-known phenomenon in Social Network Analysis, and is generally considered inherent to most friendship networks (Block 2018). However, in the present case, it also comes from the way in which friendship functions among adolescents specifically: belonging to the same social circles is often more important than person-to-person affinities or feelings. Bidart (2010) talks of a “communitarian” mode of friendship making, that would progressively be replaced by more elective modes as individuals grow older (in Western societies at least). As a result, students’ networks are more homogeneous than what their individual dispositions alone would imply, since group processes structurally imply an aggravation of individual homophilic inclinations.

This is not specific to socioeconomic origin but should occur for virtually any individual attribute (Asikainen et al. 2020; Foster et al. 2011). Nevertheless, what is unique to socioeconomic homophily is that this transitive aggravation appears as the strongest of all inducing processes. By contrast, ethnic and gender homophilies are primarily determined by ethnic and gender selection, even though transitivity aggravation occurs as well. This specificity of socioeconomic homophily is the joint result of direct socioeconomic selection being weak, and of a myriad of factors relevant to friendship formation being correlated to socioeconomic origin. In turn, this comes from the fact that students’ awareness of their own backgrounds and of that of their peers is weak, and yet socioeconomic origin still affects most aspects of their lives.

4. Low-order Relational Processes: Time Trends and Variations across Contexts (RQ5)

The low-order processes identified do not vary much over time. In some schools (Paris 2 and possibly Savoie 1), socioeconomic selection peaks at specific waves, but this seems to capture periodical variations rather than a consistent time trend. The importance of cultural similarity may increase as students grow older, but the evidence for this remains very slight – the transition from middle to high school will likely matter more in that regard than the effect of students' aging through middle school years does. Finally, the role of the former primary school decreases over time, which seems logical. As for the strong increase in socioeconomic homophily observed in Paris 1, it does not seem to result from a change in students' behavior; rather, the effect of homophily-inducing processes accumulates over time, such that the networks progressively shift toward more and more homophily.

If variation in time is uncertain, variation across contexts is extremely clear. Depending on the school (or camp), the extent to which adolescents tend to select same-background peers sharply differ – moderate in Paris 1, weak and inconsistent in time in Paris 2 and Savoie 1, either null or very slight in the summer camp, and most likely null in Savoie 2. This contrasts with gender or ethnic selection, both of which are much more consistent across cases. Academic selection is also different across schools, more or less following the same pattern as socioeconomic selection (high in Paris 1, moderate in Paris 2 and Savoie 1 and weak or null in Savoie 2). How can we explain such variation?

Selection processes are a manifestation of students' expressed dispositions. However, these are not necessarily a faithful reflection of their internalized, or *latent*, dispositions, defined as the psychological traits engraved in their mind, and which we never get to observe directly. In the theoretical framework exposed in chapter 2 and based on Lahire's grammar of human action (2005), individuals' dispositions can be activated, or on the contrary inhibited, by contextual incentives. Therefore, variations across contexts in expressed dispositions (i.e. in homophilic selection) can come from two (non-exclusive) sources: either the pool of latent dispositions found in the respective populations differ, or the contextual moderators impacting the expression of these latent dispositions do.

Distinguishing between these incentives and latent dispositions is important, because it informs us about the behaviors that may be expected under a change of context. As such, it is crucial to potential interventions meant to reduce homophily. However, the methodological challenge that this represents is substantial; and the research design adopted here was not meant to address it, since the studied adolescents were always considered in a single context (either school or the camp). Still, given the theoretical importance of this distinction, I wish to take a few moments to discuss these points, based on the available information. This should be seen as suggestive for the most part and calls for corroboration (or refutation) based on future research.

The pool of latent dispositions in a given population will be dependent on the selection biases that control the entry into this population. In general, one would expect individuals to be selected on their positive inclinations toward heterophilic bonding when they attend a context in which their background is underrepresented, and, conversely, to be selected on their homophilic dispositions in contexts that are homogeneous from their point of view. This implies contradictory biases depending on the socioeconomic group that is considered. On the one hand, upper-class households attending diverse public schools should be relatively open to socioeconomic diversity, since the most "segregative" ones will try to avoid this type of school. On the other hand, working-class households that opted for private schooling should be selected on their academic ambitions, desire for ascending mobility and possibly cultural "good will" regarding middle- and upper-class norms. Therefore, one would expect upper-class students to have less homophilic latent dispositions in Paris 1 and Savoie 1 than in Paris 2 and Savoie 2, but this should be the opposite for working-class students.

This makes it difficult to anticipate what type of selection biases will prove more favorable to socioeconomic mixing overall. However, in the case of Paris 1, families that choose the international section, even though sending their child to a diverse school, still choose a distinct, elitist track. Whether this is either more or less of a segregative choice compared to the other options that may have been available to them is up for debate, but at the very least, it is not obvious that these households in particular were selected on pro-mixing dispositions. Moreover, based on the analyses from chapter 7, it seems that the two private schools had a pool of latent dispositions that were more favorable to socioeconomic mixing overall: both cultural tastes and academic attitudes were more homogeneous in Savoie 2 than in the other schools, whereas socioeconomic origin was a remarkably poor predictor of academic performance in Paris 2.

With that said, there are also good reasons to think that latent dispositions are not sufficient to explain the observed differences in homophilic selection. First is the case of the summer camp. One of the main interests of studying socioeconomic homophily there was that selection biases in the population should be weak (at least those pertaining to the attitudes toward socioeconomic mixing), given the medical nature of the camp and the associated enrolment process (cf. chapter 8). If anything, this should result in adolescents being less intrinsically disposed to socioeconomic mixing than those of the school samples[213]. Yet, as we saw, there is little to no homophilic selection there. This supports the idea that the lower selection observed there has to do with contextual moderators, that either tone homophily down in the camp or raise it in certain schools (even though the fact that the camp is a single context with a relatively small number of attendees makes it difficult to generalize from there).

A second point to support the idea that contextual moderators may be crucial in explaining variation in homophilic selection pertains to the perceptive frames used by students in selecting their peers, which were discussed in chapter 6. As we saw, students' perceptions of the socioeconomic origins of their friends is indirect and imprecise at best, which likely explains why direct homophilic selection is weak in most cases (and still moderate even in Paris 1). However, this also implies that socioeconomic selection may become stronger if students were provided with the mental categories that made socioeconomic origin more apparent, and that would let them make sense of certain small traits or behaviors related to one's background. This may be one of the particular effects of the presence of the international section in Paris 1: since internationals are overwhelmingly from an upper-class background, and since they are clearly identified among students as being the "privileged" track, certain traits typical of upper-class origins may become more noticeable to students as being typical of internationals. For instance, we saw in chapter 7 how some non-international girls would describe the clothing tastes of an upper-background non-international peer as similar to those of the internationals. This also suggests that high levels of socioeconomic homophily, regardless of how they emerged, might foster direct socioeconomic selection in return: the more socioeconomic origins correspond to well-bounded groups in the friendship networks, the more likely it is that students will associate these origins with certain traits or behaviors and use them to explain the friendship structure as they observe it, eventually increasing the perceptive salience of socioeconomic attributes as a whole. Therefore, not only does selection explain homophily, but the opposite may also be

213 Given that the four schools are socioeconomically diverse, I would expect the selection biases at the summer camp to be favorable to pro-mixing households overall, relative to a random draw from the general population (even parents from the international track in Paris 1 chose a public and diverse school, whereas downtown Paris, where they live, is certainly not short of elitist private schools where they could have "escaped" to).

true: homophily might sometimes explain selection, to the extent that it acts as a contextual moderator enhancing individuals' latent homophilic dispositions.

5. Policy Implications for Socioeconomic Mixing at School

Changing the contextual moderators to which students are exposed is, together with the desegregation of imposed foci, the main way in which school administrations (or summer camp staffs) can impact the strength of socioeconomic homophily. What kind of moderators, then, could favor mixing among youths? Again, the answer provided here can only be suggestive, since the research design was not meant to investigate changes in students' practices under an external change of context. Still, some exploratory insights are worth considering.

Firstly, the role of special tracks, and of institutional labels more generally. As the cases of Paris 1 and Savoie 1 have shown, these are clear obstacles to socioeconomic mixing: they are at the same time both a socially segregated foci and a contextual moderator that fosters selection processes. Assuming that these tracks exist nevertheless – after all, in the case of Paris 1, there would simply be far less socioeconomic diversity without the international section, which would imply mostly homogeneous friendships among students – it is probably preferable, in terms of homophily reduction, to split the concerned students across different classrooms. The interviews conducted in Paris 1 also showed that students are very sensitive to unequal treatment among them, such as the internationals benefiting from more out-of-school trips. The fact that the tracks are treated differently by the school staff can only support students' perceptions of these tracks as a relevant means of distinction and of classification[214]. As for the pre-vocational track of Savoie 1, I could not investigate it as much as the international section of Paris 1, but the clear isolation of its members in the friendship networks suggests that similar processes may be at work.

Another important factor pertains to the polarization of the distribution of socioeconomic origins. Having a relatively homogeneous continuum of origins, with a fair share of middle-class backgrounds (as in Savoie 1), is probably preferable to polarized distributions with two clear poles of upper- and lower-class origins and few intermediate backgrounds (as in Paris 1). Apart from the fact that middle-class students may act as intermediaries between extreme socioeconomic positions, polarized distributions probably

214 See Chabot (2017) for an example of a Parisian school relatively similar to Paris 1, but where the administration opted for a different strategy by "diluting" the selective track (a rare language option) into the other classrooms while downplaying differences between students, with some success.

contribute to making socioeconomic factors more apparent in students' perceptive frames, since it is easier to perceive discrete categories rather than a continuous spectrum.

The case of academic results and attitudes is more complicated. Based on the interviews, we saw that these attitudes are one of the main perceptive frameworks used by students to sort out their peers. Given the correlation of socioeconomic background and academic results and attitudes, this implies that reducing academic selection should also reduce socioeconomic homophily. In order to do so, schools may try to attenuate the interpersonal competition among students, for instance by promoting group works or by using other forms of evaluations beyond numerical grades. In that regard, the weak and non-significant effect of academic similarity in Savoie 2 may come from the fact that students there do not have such numerical grades – though it may also come from the higher academic homogeneity of the school. This may also be one of the reasons why socioeconomic selection and homophily were weak during the summer camp, where academic factors are entirely absent. However, the contribution scores estimated in chapter 5 seemed to suggest that academic selection has a relatively weak impact on socioeconomic homophily, once other relevant factors are accounted for. Moreover, we saw in chapter 6 that the tensions born from students having different academic attitudes may be less pronounced among different-background peers, potentially mitigating the impact of academic hierarchies on socioeconomic homophily. Furthermore, an attenuation of academic hierarchies already seems to be at work within students' subcultures, as they resort to different strategies meant to flatten academic hierarchies, with varied success (stating that grades are a private matter, high-achieving students remaining discrete and humble, etc.).

Altogether then, although academic hierarchies probably play a role in fostering socioeconomic homophily, this might not be the most decisive factor at work. With that said, one property of socioeconomic homophily is precisely that there is no single decisive factor underlying it, so policy interventions aimed at reducing it will have to tackle several minor inducers anyway. Therefore, it seems reasonable to treat academic hierarchies as one piece of the broader puzzle of socioeconomic homophily reduction.

Regarding contextual moderators favorable to socioeconomic mixing, one case that deserves particular attention is that of the summer camp studied in chapter 8. Homophilic selection appeared null or very weak there, despite teenagers having a lot of freedom in how they organized their sociability. What types of contextual moderators may have toned down homophilic dispositions during the camp? According to McFarland et al. (2014), homophilic behaviors should be less pronounced in schools with a "positive climate", i.e. where adolescents feel safe and are not subject to stress. According to them, individuals that feel threatened in some way would be more prone to look for similar peers, as this would provide them with a greater sense of security. This is a rather generic argument, and, as I explained in chapter 2 (section 2.2.2), I am relatively skeptical about the empirical demonstration provided in the article. Nevertheless, it resonates with Intergroup Contact Theory in social psychology, which makes an essentially similar argument (Allport 1954). In a relational context marked by uncertainty and the fear of marginalization (which very much tends to be the case in middle school), students will most likely fall back on "easier" relationships, with people that they know well and whose reaction they can anticipate more easily – which, on average, should correspond to same-background peers. By contrast, if they feel at ease and secured in their interactions with others, they should be most likely to try and reach out to peers that feel more different or that they do not know well, making it more likely for heterophilic relationships to form.

In that regard, an interesting feature of the camp is the role of the animation staff, which actively tried to facilitate relationships among attendees by organizing shared games and activities, offering emotional and social support to shy or isolated adolescents, or mediating conflicts and disputes. This contrasts with the behavior of teachers and school staff in French middle schools, which typically do not tend to meddle in students' social lives. The fact that they maintain a sharp distinction between adults and youth also makes social and emotional support effectively impossible; typically, a student playing or spending a lot of time with an adult is hard to imagine in a school context and would likely result in a form of stigma among same-age peers, whereas this was common and well-accepted in the context of the summer camp.

Consequently, schools may try to reduce homophilic selection by promoting safer and less stressful relational contexts, in which students do not feel pressured into joining a protective group of friends at any cost, and where they feel confident enough to try and reach out to dissimilar peers. This may be done by regulating student relations and facilitating positive contacts among classmates through greater adult intervention, in order to "grease the wheels" of the relational structure. With that said, these recommendations essentially come from a theoretical extrapolation, and there is no direct empirical result that supports them in the thesis. Therefore, they should be considered with care. Finally, in the conclusion of chapter 8, I also mentioned other potential moderators that may have toned down homophilic selection during the summer camp (the short duration of the stay, or a feeling of sameness among attendees due to their common experience of the chronic disease that they suffer from), but these are specific to this particular camp, and do not have clear implications for school year contexts.

The last potential lever that may be used to reduce socioeconomic homophily pertains to network clustering. Given the key result that transitivity processes strongly aggravate socioeconomic homophily in the studied schools – further supported by formal simulation models (Asikainen et al. 2020; Foster et al. 2011) –, it is tempting to conclude that weakly clustered friendship networks are generally favorable to socioeconomic mixing. Looking back on our sample of schools, it is indeed the case that Paris 1 has the most transitive network, and Savoie 2 the least transitive one (Appendix 3B). As for the summer camp, even though it has very transitive friendship networks, roughly at the same level as Paris 1, we also saw that socioeconomic homophily only appeared in co-sittings during meals, not in declared friendship nominations. One of the potential explanations put forward in chapter 8 was that the forced discretization implied by the table structure – i.e. the fact that students are necessarily organized in discrete groups – may reinforce endogenous group processes compared to declarative friendship networks, thus entailing more socioeconomic homophily. This would also support the idea that more clustered relational structures are more favorable to socioeconomic homophily.

This idea, however, should be considered with care: because transitivity affects the network structure in very deep ways, reducing it may have unforeseen implications in empirical contexts. For example, Blau (1977) and McFarland et al. (2014) predict less homophily in contexts where individuals have less freedom in choosing their friends based on elective, person-to-person affinities. While they primarily had in mind the size and diversity of a population, this should also apply to network clustering: highly clustered networks impose more external constraints on individuals' friendship choices, since friends are selected based on their belonging to common social circles rather than dyadic affinities. In a similar vein, Block (2018) argues that there is a negative interaction effect between transitivity and homophily, because shared foci matter more for dissimilar individuals – again, a variation on this idea that homophily is higher the more unconstrained the pairing process tends to be (cf. the discussion in chapter 2, section 2.2.2). Interestingly, these arguments seem to rest on the assumption that dispositional selection is one of the prime drivers of homophily (McFarland et al. explicitly speak of "choice homophily"), thus the belief that the more "room" dispositions have to be expressed the higher the homophily. However, as we saw, direct selection plays a rather minor role in the case of socioeconomic homophily, so this idea that less structural constraints imply more homophily might not apply to the case of socioeconomic origin. Furthermore, the extent to which the network structure is clustered may act, not just as a structural effect (which is what simulations can test for), but also as a contextual moderator eventually impacting the salience of socioeconomic attributes in selection processes. Thus, more research is needed to understand the exact role of network and group processes in fostering homophilic network structures, including empirical comparisons of schools with different levels of transitivity, as well as stylized simulations meant to formally explore the structural processes at work.

In any case, and assuming that reducing clustering does indeed favor socioeconomic mixing, it is not clear how this could be done concretely. One solution may be to diversify the social situations in which friendship formation can occur at school; for instance, re-shuffling classrooms and learning groups more frequently, mixing students from different classrooms for certain lessons (which is already the case for certain language options), or organizing more extra-curricular activities (which tend to be relatively rare in the French context).

6. Insights into the Origins of Socioeconomic Homophily among Adults

What can we learn from these results about the ways in which socioeconomic homophily emerges in society at large? What seems clear is that teenagers – at least those in diverse schools – do not have strong and rigid internalized preferences (yet?) for same-background peers. Their homophilic dispositions, though they exist, appear rather moderate, and are susceptible to being toned down by appropriate contextual moderators. This suggests that the inter-generational transmission of homophilic dispositions through primary socialization within the family is an overall modest *explanans* of socioeconomic homophily, not only at school but, by extension, across the course of life. At the very least, it is fair to say that socioeconomic origin does not automatically entail precocious dispositions that would prevent any form of heterophilic bonding whatsoever.

From there, three broad classes of explanation can be put forward to reconcile this observation of rather weak homophilic dispositions among children and adolescents with what we know of the strength of socioeconomic homophily among adults (Bidart et al. 2011; McPherson et al. 2001).

First, it may be the case that the adolescents studied here differ from those found in other, non-diverse contexts: in segregated schools, individuals may exhibit stronger internalized dispositions toward socioeconomic similarity in friendship. This may come in part from a selection bias: the households found in diverse schools would transmit less homophilic dispositions to their children than the households that search for segregated contexts. However, it may also underline the role of spatial segregation in the socialization of children and in their acquisition of homophilic dispositions: acquiring dispositions for socioeconomic homophily in childhood would not simply depend on the occupation or education of parents, but on the broader social environment to which children are exposed (which is actually quite a trivial hypothesis to make). In that regard, the strategies of spatial closure or of "surrounding" of the child that are found in certain upper-class households (van Zanten 2015) would be decisive in the reproduction of the psychological dispositions responsible for homophily and, more generally, practices of socioeconomic distinction among adults.

Second is of course the idea that, just like the adolescents studied here, most adults do not have strong homophilic inclinations toward same-class friends. This would point to imposed foci and spatial segregation as the driving factors of socioeconomic homophily among adults, as well as contextual moderators that are unfavorable to heterophilic bonding in most adult social contexts (e.g. hierarchical relations in the workplace likely hinder the formation of friendship between managers and workers).

Finally, the internalization of homophilic dispositions may occur later during the course of life. Bidart et al. (2011), when looking at the friendship networks of young adults during the transition from studies to work, suggest that university years may be decisive in the constitution of personal networks and in the type of sociability practices to which individuals are accustomed. Similarly, Kane (2011), looking at the ego-networks of students from an elite college, found that the structure of freshmen's ego-networks quickly evolved to match their new environment, becoming both sparser and more diverse than their former high school ego-networks were. By extension, it might be the case that many dispositions related to friendship formation, such as one's expectations about one's friends or preferred sociability contexts, form during university years (at least for individuals going into tertiary education). Alternatively, socialization to sociability might be continuous over the course of life, and pertain to an individual's class of destination rather than to their socioeconomic origin[215].

215 A more subtle version of the argument could be that certain latent dispositions are indeed acquired during early family socialization, but that they only become effective in guiding the choice of friends later in life

Of course, these three explanations are not mutually exclusive. In any case, they remain very generic at this stage, and dedicated research would be needed to explore them further. What we can say is that (most) adolescents in mixed contexts have an overall weak intrinsic propensity to look for same-background friends. This implies that more processes are needed to account for the society-wise level of socioeconomic homophily, beyond the inter-generational transmission of homophilic psychological dispositions through family socialization.

7. Limitations and Directions for Research

The main limitations of the thesis have already been made apparent in the discussion so far. First comes the difficulty of distinguishing between the effects of latent dispositions and that of contextual moderators. In order to circumvent this issue, future research should try to extend the range of the sociability contexts that are considered. An ideal design would be to follow the same individuals across a variety of contexts. However, this appears particularly challenging from a methodological point of view, at least if one wishes to use complete networks: for instance, it would demand us to model friendship networks within a school, and then to follow some of the students of that school in their out-of-school activities or during the holidays and, in each of those contexts, to collect complete network data as well. A less ambitious but probably more feasible solution would be to use a survey method centered on individuals' ego-networks, rather than on complete networks (similar, for instance, to the methodology of Bidart et al. 2011). Ego-networks usually have less modeling potential than complete networks do, but, together with detailed and longitudinal interviews, they may allow for reconstituting respondents' relational inclinations across their various sociability contexts. Furthermore, it would be valuable to examine adolescent friendships in a broad diversity of situation beyond school settings, even if they do not concern the same individuals. Summer camps such as the one examined in chapter 8 appear as a promising lead, and investigating different camps with various activities, pedagogical orientations and socioeconomic compositions would greatly enrich our understanding of how youth sociability in general, and socioeconomic homophily in particular, differ across settings. Other interesting contexts to explore may be out-of-school activities, street sociability, and online sociability (social media, multiplayer games, online communities, etc.).

(for instance, certain musical tastes are acquired during childhood, but individuals talk more of more about music with their friends as they grow older).

Second, the small number of schools that were studied in this dissertation severely limited the possibilities for inter-school comparison. A larger sample would permit a more rigorous examination of school-level associations between features of the context and levels of socioeconomic homophily and would provide more solid ground for generalization. To the best of my knowledge, there is no large network database on adolescent friendship in France as of now, or at least nothing that can be compared to what exists in the United States with the Add Health project (Harris 2013) or in other European countries with the CILS4EU project (Kalter et al. 2016). Using these two data bases for testing some of the hypotheses formulated in chapter 7 on inter-school comparison thus seems promising, although the issue of the cross-national validity of results should be considered as well (for example, I mentioned above that the processes pertaining to popularity and network centralization may differ between the French and US contexts). Larger samples would also allow for an examination of the effect of schools' socioeconomic composition on student sociability. How different are the friendship networks of diverse schools compared to those of segregated ones, both toward the top and bottom of the socioeconomic ladder? And what is the relational situation of students whose socioeconomic backgrounds contrast with that of their environment – i.e., how do upper-class students fare in dominantly working-class schools, and vice-versa? Note that all these questions could (and should) also be investigated with ethnographic approaches, not only with large quantitative databases. In particular, this will be essential to get a better understanding of how local features of the school context affect friendship formation. To this end, and compared with what already exists in the literature, the use of Social Network Analysis methods would greatly help in making the results of several monographs readily comparable, thus facilitating the constitution of a cumulative body of knowledge over separate studies and opening the door to large scale meta-analyses.

A third limitation pertains to the imbalance between quantitative and qualitative methods in the dissertation, and notably the fact that I could not collect longitudinal interviews. A longitudinal examination of how students' perceptive categories for sorting out their peers evolve over time would likely prove valuable. Additionally, ethnographic work may try to examine socioeconomic homophily through direct observation, rather than self-reports from adolescents. This should include fine-grained descriptions of youths' social practices, of course, but also the collection of observational data meant to objectify sociability practices, which could then be analyzed quantitatively, similar to the meal-sharing partitions of chapter 8. One question worth exploring through these methods is whether the substantial content of social ties is the same for homophilic and heterophilic pairings. Are friends from different backgrounds doing the same kind of activities together, or talking about the same things, that they are with same-background peers? Or does heterophilic bonding rest on specific types of relationships, that tend to give a different "weight" to various components of friendship such as practical help, secret sharing, emotional support and shared activities? In this second case, it would be interesting to examine the topics, interests or modes of socializing that are favorable to either homophilic or heterophilic mating. Incidentally, this would also be a way to open the "black box" of homophilic dispositions, in order to study the specific traits or attitudes that push socioeconomically similar individuals toward one another – something that quantitative estimates like those presented in chapter 5 are typically powerless to do.

Fourth, the age group studied here – teenagers aged 10 to 14 – is a narrow one. All the questions raised so far could thus equally be raised for primary school children, high school students, and all the other stages of the life course. Apart from the similarities and differences between these periods, one key interrogation pertains to the resilience of friendship ties during the transitions between different stages. In particular, I mentioned the fact that middle school may be one of the most diverse contexts, in socioeconomic terms, that students will encounter throughout their lives (cf. chapter 3, section 2.1). Since we saw that some heterophilic friendships can indeed form in diverse schools, it stands to question whether these will survive when students move on to high school, especially in cases where friendship pairs are separated in different schools, tracks and eventually career paths. Longitudinal panels that cover the transitions between schooling levels, or retrospective interviews with high school students where they are asked about their former middle school friends, would be appropriate designs in that regard (again, Bidart et al. 2011 provides a good methodological example, even though it concerns older individuals). Given the result that socioeconomic homophily tends to be stronger for more intimate relationships, and that out-of-school relations are less favorable to mixing that in-school ones, I would expect the survival rate of heterophilic ties to be lower on average compared to homophilic ones; but this would demand empirical confirmation. Furthermore, even if they are relatively rare, the heterophilic friendships formed during middle school that manage to persist over the years could hold important hints as to the ways in which cross-background ties can be nurtured.

Finally, a few methodological innovations have been proposed throughout the thesis. While these proved important in answering my research questions, their novelty also implied that their generic properties, or their expected behaviors under different situations, remain relatively obscure at this stage. Testing and applying these approaches in different studies would therefore be necessary to explore their potential, on top of (hopefully!) proving useful to answering other research puzzles. First is the longitudinal ERPM (*Exponential Random Partition Model*) presented in chapter 8 and which is an extension of the original cross-sectional ERPM framework (Hoffman et al. 2020). ERPM applies the logic of inferential network modeling of discrete group partitions. The utility of this approach is rather straightforward: many relational phenomena are organized in group structures, rather than in networks. This includes for instance partying, group holidays, or working groups at school or in the workplace. Moreover, research in social psychology has shown that the mental representation of friendship structures by individuals is typically organized in group form: for example, when asked who is friends with who in their classroom, students will usually report several groups of peers that often hang out together. This is the basis of the method known as *cognitive maps*, which is an alternative to network analysis in investigating friendship structures: several students are asked to report friendship groups in their classroom or grade-level, and their answers are aggregated into a common partition. ERPM makes it possible to model these partitions in the same way as network data, which opens the way for easier comparison and integration of the two approaches in the study of sociability.

A second tool proposed in the dissertation – though not a methodological one strictly speaking – is the theoretical model of friendship formation presented in chapter 2, which essentially consisted of piecing together known processes in the SNA literature within the broader contextual-dispositional framework of Lahire (2005). It proved to be heuristic in explaining the emergence of socioeconomic homophily in friendship networks but could be applied to a variety of relational phenomena, such as other types of homophily or centralization patterns.

Finally, the contribution scores developed in chapter 5, which incidentally pair well with the aforementioned theoretical model. In principle, these could be used to assess the contribution of low-order processes to any aggregate network statistic captured by either ERGM or SAOM. Moreover, they could also be applied to what Snijders and Steglich (2015) call “true macro properties”, that is, network features that cannot be described as a sum of node-level configurations (for example, the average geodesic distance of the network, or core-periphery structures). From a technical point of view, the exact structural implications of setting a parameter to 0, and how this affects the expression of other model parameters, should be further explored; for instance, whether this can lead to degenerate simulations and to unrealistic network structures in certain cases. The sensitivity of the estimated contribution scores to mis-specified models or to the omission of important covariates would also be worth considering.

Therefore, many promising leads exist for further research on socioeconomic homophily among children and adolescents. Crucially, the integration of monographic approaches and of the toolbox of Social Network Analysis, as well as the bridging of the different research traditions that share an interest in youth sociability, open up new methodological and theoretical perspectives – not only regarding socioeconomic homophily, but, more broadly, for the study of human sociability as well as of socioeconomic inequalities and social reproduction.

Bibliography

Adler, Patricia, and Peter Adler. 1998. *Peer Power: Preadolescent Culture and Identity*. Rutgers University Press.

Allport, Gordon Willard. 1954. *The Nature of Prejudice*. Addison-Wesley Publishing Company.

Anderson, Brigham S., Carter Butts, and Kathleen Carley. 1999. "The Interaction of Size and Density with Graph-Level Indices." *Social Networks* 21(3):239–67. doi: 10.1016/S0378-8733(99)00011-8.

Anon. 2017. *Les panels d'élèves de la DEPP: source essentielle pour connaître et évaluer le système éducatif*. Paris: Ministère de l'Éducation Nationale.

Asikainen, Aili, Gerardo Iñiguez, Javier Ureña-Carrión, Kimmo Kaski, and Mikko Kivelä. 2020. "Cumulative Effects of Triadic Closure and Homophily in Social Networks." *Science Advances* 6(19):eaax7310. doi: 10.1126/sciadv.aax7310.

Baerveldt, Chris, Marijtje A. J. Van Duijn, Lotte Vermeij, and Dianne A. Van Hemert. 2004. "Ethnic Boundaries and Personal Choice. Assessing the Influence of Individual Inclinations to Choose Intra-Ethnic Relationships on Pupils' Networks." *Social Networks* 26(1):55–74. doi: 10.1016/j.socnet.2004.01.003.

Ball, Stephen J. 1981. *Beachside Comprehensive: A Case-Study of Secondary Schooling*. CUP Archive.

Ball, Stephen J., Carol Vincent, and Sophie Kemp. 2004. "'Un agréable mélange d'enfants... ' : prise en charge de la petite enfance, mixité sociale et classes moyennes." *Education et sociétés* no 14(2):13–31.

Battaglia, Mattea. 2016. "Comme au lycée, l'affectation des collégiens parisiens passera par un algorithme." *Le Monde.fr*, September 12.

Baumeister, Roy F., Kathleen D. Vohs, and David C. Funder. 2007. "Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior?" *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 2(4):396–403. doi: 10.1111/j.1745-6916.2007.00051.x.

Becker, Howard S. [1963] 2008. *Outsiders*. Simon and Schuster.

Bernstein, Basil. [1977] 2003. *Class, Codes and Control: Applied Studies towards a Sociology of Language*. Psychology Press.

Bidart, Claire. 1991. "L'amitié, les amis, leur histoire. Représentations et récits." *Sociétés Contemporaines* 5(1):21–42. doi: 10.3406/socco.1991.984.

Bidart, Claire. 1997. *L'amitié, un lien social*. La Découverte.

Bidart, Claire. 2010. "Les âges de l'amitié." *Transversalités* (113):65–81. doi: 10.3917/trans.113.0065.

Bidart, Claire, Alain Degenne, and Michel Grossetti. 2011. *La vie en réseau: Dynamique des relations sociales*. Paris: Presses Universitaires de France.

Blau, Peter M. 1994. *Structural Contexts of Opportunities*. University of Chicago Press.

Blau, Peter Michael. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. Free Press.

Block, Per. 2015. “Reciprocity, Transitivity, and the Mysterious Three-Cycle.” *Social Networks* 40:163–73. doi: 10.1016/j.socnet.2014.10.005.

Block, Per. 2018. “Network Evolution and Social Situations.” *Sociological Science* 5:402–31. doi: 10.15195/v5.a18.

Block, Per, and Thomas Grund. 2014. “Multidimensional Homophily in Friendship Networks*.” *Network Science* 2(2):189–212. doi: 10.1017/nws.2014.17.

Block, Per, James Hollway, Christoph Stadtfeld, Johan Koskinen, and Tom Snijders. 2019. “‘Predicting’ after Peeking into the Future: Correcting a Fundamental Flaw in the SAOM -- TERGM Comparison of Leifeld and Cranmer (2019).” *ArXiv:1911.01385 [Stat]*.

Block, Per, Johan Koskinen, James Hollway, Christian Steglich, and Christoph Stadtfeld. 2018. “Change We Can Believe in: Comparing Longitudinal Network Models on Consistency, Interpretability and Predictive Power.” *Social Networks* 52:180–91. doi: 10.1016/j.socnet.2017.08.001.

Block, Per, Christoph Stadtfeld, and Tom A. B. Snijders. 2019. “Forms of Dependence: Comparing SAOMs and ERGMs From Basic Principles.” *Sociological Methods & Research* 48(1):202–39. doi: 10.1177/0049124116672680.

Bonnéry, Stéphane. 2006. “La Question de «l’ethnicité» Dans l’École: Essai de Reconstruction Du Problème.” *Sociétés et Jeunesses En Difficulté. Revue Pluridisciplinaire de Recherche* (1).

Boudon, Raymond. [1973] 2011. *L’inégalité des chances: la mobilité sociale dans les sociétés industrielles*. Pluriel.

Bourdieu, Pierre. [1979] 2016. *La Distinction: Critique sociale du jugement*. Minuit.

Bourdieu, Pierre, and Jean Claude Passeron. [1970] 1990. *La Reproduction*. Sage in association with Theory, Culture & Society, Dept. of Administrative and Social Studies, Teesside Polytechnic.

Boyer, Régine. 2006. “Pasquier Dominique. Cultures lycéennes : la tyrannie de la majorité. Paris : Éd. Autrement, 2005. – 180 p. (Mutations).” *Revue française de pédagogie. Recherches en éducation* (155):170–72.

Brantlinger, Ellen, Massoumeh Majd-Jabbari, and Samuel L. Guskin. 1996. “Self-Interest and Liberal Educational Discourse: How Ideology Works for Middle-Class Mothers.” *American Educational Research Journal* 33(3):571–97. doi: 10.2307/1163277.

Brechwald, Whitney A., and Mitchell J. Prinstein. 2011. “Beyond Homophily: A Decade of Advances in Understanding Peer Influence Processes.” *Journal of Research on Adolescence* 21(1):166–79. doi: 10.1111/j.1532-7795.2010.00721.x.

de Bruyn, E. H., A. H. N. Cillessen, and I. B. Wissink. 2010. “Associations of Peer Acceptance and Perceived Popularity With Bullying and Victimization in Early Adolescence.” *The Journal of Early Adolescence* 30(4):543–66. doi: 10.1177/0272431609340517.

Burk, William J., Margaret Kerr, and H\a akan Stattin. 2008. “The Co-Evolution of Early Adolescent Friendship Networks, School Involvement, and Delinquent Behaviors.” *Revue Française de Sociologie* 49(3):499–522.

Cartwright, Dorwin, and Frank Harary. 1956. “Structural Balance: A Generalization of Heider’s Theory.” *Psychological Review* 63(5):277–93. doi: 10.1037/h0046049.

Cepić, Dražen, and Željka Tonković. 2020. “How Social Ties Transcend Class Boundaries? Network Variability as Tool for Exploring Occupational Homophily.” *Social Networks* 62:33–42. doi: 10.1016/j.socnet.2020.02.003.

Chabot, Timothée. 2017. “Rapports de Classe(s) : Mixité Sociale et Scolaire Dans Un Collège Parisien.” Mémoire de master, Sciences Po Paris, Paris.

Chan, Tak Wing, and John H. Goldthorpe. 2007. “Class and Status: The Conceptual Distinction and Its Empirical Relevance.” *American Sociological Review* 72(4):512–32. doi: 10.1177/000312240707200402.

Cillessen, Antonius H. N., and Lara Mayeux. 2004. “From Censure to Reinforcement: Developmental Changes in the Association Between Aggression and Social Status.” *Child Development* 75(1):147–63. doi: 10.1111/j.1467-8624.2004.00660.x.

Cohen, Albert Kircidel. 1955. *Delinquent Boys: The Culture of the Gang*. Free Press.

Cohen, Jere. 1979. “Socio-Economic Status and High-School Friendship Choice: Elmtown’s Youth Revisited.” *Social Networks* 2(1):65–74. doi: 10.1016/0378-8733(79)90011-X.

Coie, John D., Kenneth A. Dodge, and Heide Coppotelli. 1982. “Dimensions and Types of Social Status: A Cross-Age Perspective.” *Developmental Psychology* 18(4):557–70. doi: 10.1037/0012-1649.18.4.557.

Coleman, James S. 1960. “The Adolescent Subculture and Academic Achievement.” *American Journal of Sociology* 65(4):337–47. doi: 10.1086/222722.

Coleman, James S. 1988. “Social Capital in the Creation of Human Capital.” *American Journal of Sociology* 94:S95–120.

Coleman, James Samuel. 1961. *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education*. Free Press of Glencoe.

Coleman, James Samuel. [1966] 1995. *Equality of Educational Opportunity (Coleman) Study (EEOS), 1966*. Inter-university Consortium for Political and Social Research.

Coleman, James Samuel, and Thomas Hoffer. 1987. *Public and Private High Schools: The Impact of Communities*. Basic Books.

Cook, Philip J., and Jens Ludwig. 1997. “Weighing the ‘Burden of “Acting White”’: Are There Race Differences in Attitudes toward Education?” *Journal of Policy Analysis and Management* 16(2):256–78.

Coulangeon, Philippe. 2003. “La stratification sociale des goûts musicaux: Le modèle de la légitimité culturelle en question.” *Revue Française de Sociologie* 44(1):3. doi: 10.2307/3323104.

Coulangeon, Philippe. 2010. “Les métamorphoses de la légitimité: Classes sociales et goût musical en France, 1973-2008.” *Actes de la recherche en sciences sociales* 181–182(1):88. doi: 10.3917/arss.181.0088.

Coulmont, Baptiste, and Patrick Simon. 2021. “Transmission Des Prénoms Dans Les Familles Immigrées : Réponse à Jean- François Mignot.”

Cousin, Olivier, and Georges Felouzis. 2002. *Devenir collégien: l’entrée en classe de sixième*. Paris: ESF.

Da Costa, Sylvie, and Agnès van Zanten. 2011. “L’enseignement Privé Entre Mission, Management et Marché.” Pp. 291–307 in *L’État et l’enseignement privé. L’application de la loi Debré (1959)*. Rennes: Presses Universitaires de Rennes.

Delalande, Julie. 2003. “Culture enfantine et règles de vie. Jeux et enjeux de la cour de récréation.” *Terrain. Anthropologie & sciences humaines* (40):99–114. doi: 10.4000/terrain.1555.

Delalande, Julie. 2005. “Comment le groupe s’impose aux enfants.” *Empan* no48(4):27–31.

Delvaux, Bernard, and Agnès Van Zanten. 2006. “Les établissements scolaires et leur espace local d’interdépendance.” *Revue française de pédagogie* (156):5–8. doi: 10.4000/rfp.260.

Dijkstra, Jan Kornelis, Tina Kretschmer, Kim Pattiselanno, Aart Franken, Zeena Harakeh, Wilma Vollebergh, and René Veenstra. 2015. “Explaining Adolescents’ Delinquency and Substance Use: A Test of the Maturity Gap: The SNARE Study.” *Journal of Research in Crime and Delinquency* 52(5):747–67. doi: 10.1177/0022427815582249.

Dijkstra, Jan Kornelis, Siegwart Lindenberg, Frank C. Verhulst, Johan Ormel, and René Veenstra. 2009. “The Relation between Popularity and Aggressive, Destructive, and Norm-Breaking Behaviors: Moderating Effects of Athletic Abilities, Physical Attractiveness, and Prosociality.” *Journal of Research on Adolescence* 19(3):401–13.

Doyle, Jamie Mihoko, and Grace Kao. 2007. “Friendship Choices of Multiracial Adolescents: Racial Homophily, Blending, or Amalgamation?” *Social Science Research* 36(2):633–53. doi: 10.1016/j.ssresearch.2006.12.001.

Dubet, François, and Danilo Martuccelli. 1996. “Théories de la socialisation et définitions sociologiques de l’école.” *Revue Française de Sociologie* 37(4):511. doi: 10.2307/3322131.

Dubet, François, and Danilo Martuccelli. 2014. *A l’école . Sociologie de l’expérience scolaire*. Le Seuil.

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach*. Cambridge: Cambridge University Press.

Duru-Bellat, Marie. 2015. *Les inégalités sociales à l'école: Genèse et mythes*. Presses Universitaires de France.

Duru-Bellat, Marie, and Alain Mingat. 1997. "La Constitution de Classes de Niveau Dans Les Collèges; Les Effets Pervers d'une Pratique à Visée Égalisatrice." *Revue Française de Sociologie* 38(4):759. doi: 10.2307/3322627.

Eder, Donna. 1985. "The Cycle of Popularity: Interpersonal Relations Among Female Adolescents." *Sociology of Education* 58(3):154. doi: 10.2307/2112416.

Efron, David. 1941. *Gesture and Environment: A Tentative Study of Some of the Spatio-Temporal and "Linguistic" Aspects of the Gestural Behavior of Eastern Jews and Southern Italians in New York City, Living Under Similar as Well as Different Environmental Conditions*. King's crown Press.

Erikson, Robert. 1984. *Social Class of Men, Women and Families*. Swedish Institute for Social Research.

Felouzis, Georges. 2003. "La ségrégation ethnique au collège et ses conséquences." *Revue française de sociologie* Vol. 44(3):413–47.

Flache, Andreas, Michael Mäs, Thomas Feliciani, Edmund Chattoe-Brown, Guillaume Deffuant, Sylvie Huet, and Jan Lorenz. 2017. "Models of Social Influence: Towards the Next Frontiers." *Journal of Artificial Societies and Social Simulation* 20(4):2.

Flashman, Jennifer. 2012. "Academic Achievement and Its Impact on Friend Dynamics." *Sociology of Education* 85(1):61–80.

Fordham, Signithia, and John U. Ogbu. 1986. "Black Students' School Success: Coping with the 'burden of "Acting White."'" *The Urban Review* 18(3):176–206. doi: 10.1007/BF01112192.

Fortuin, Janna, Mitch van Geel, and Paul Vedder. 2016. "Peers and Academic Achievement: A Longitudinal Study on Selection and Socialization Effects of in-Class Friends." *The Journal of Educational Research* 109(1):1–6. doi: 10.1080/00220671.2014.917257.

Foster, David V., Jacob G. Foster, Peter Grassberger, and Maya Paczuski. 2011. "Clustering Drives Assortativity and Community Structure in Ensembles of Networks." *Physical Review E* 84(6):066117. doi: 10.1103/PhysRevE.84.066117.

Freeman, Linton C. 1978. "Centrality in Social Networks Conceptual Clarification." *Social Networks* 1(3):215–39.

Furman, Wyndol. 1996. "The Measurement of Friendship Perceptions: Conceptual and Methodological Issues."

Galitz, Tara, and Dominique Robert. 2014. "Governing Bullying through the New Public Health Model: A Foucaultian Analysis of a School Anti-Bullying Programme." *Critical Public Health* 24(2):182–95. doi: 10.1080/09581596.2013.784394.

Galland, Olivier. 2011. “Chapitre 8 - Sociabilité et loisirs.” *U* 217–42.

Ganzeboom, Harry. 2019. “Tools for Occupation Status Measures for ISCO-68.” Retrieved August 27, 2021 (http://www.harryganzeboom.nl/isco08/index.htm).

Ganzeboom, Harry B. G., Paul M. De Graaf, and Donald J. Treiman. 1992. “A Standard International Socio-Economic Index of Occupational Status.” *Social Science Research* 21(1):1–56. doi: 10.1016/0049-089X(92)90017-B.

Ganzeboom, Harry B. G., and Donald J. Treiman. 1996. “Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations.” *Social Science Research* 25(3):201–39. doi: 10.1006/ssre.1996.0010.

Garnier, Pascale. 2015. “L’‘agency’ des enfants. Projet scientifique et politique des ‘childhood studies.’” *Education et sociétés* n° 36(2):159–73.

Gayet, Daniel, Martine Greffier, Thierry Vasse, and Louisette Guibert. 2003. *L’élève, côté cour, côté classe*. Paris: INRP.

Geven, Sara, Jan O. Jonsson, and Frank van Tubergen. 2017. “Gender Differences in Resistance to Schooling: The Role of Dynamic Peer-Influence and Selection Processes.” *Journal of Youth and Adolescence* 46(12):2421–45. doi: 10.1007/s10964-017-0696-2.

Glevarec, Hervé, and Michel Pinet. 2009. “La « tablature » des goûts musicaux : un modèle de structuration des préférences et des jugements.” *Revue française de sociologie* 50(3):599. doi: 10.3917/rfs.503.0599.

Goodreau, Steven M., James A. Kitts, and Martina Morris. 2009. “Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks.” *Demography* 46(1):103–25.

Hallinan, Maureen T., and Ruy A. Teixeira. 1987. “Students’ Interracial Friendships: Individual Characteristics, Structural Effects, and Racial Differences.” *American Journal of Education* 95(4):563–83.

Hallinan, Maureen T., and Richard A. Williams. 1987. “The Stability of Students’ Interracial Friendships.” *American Sociological Review* 52(5):653. doi: 10.2307/2095601.

Hallinan, Maureen T., and Richard A. Williams. 1989. “Interracial Friendship Choices in Secondary Schools.” *American Sociological Review* 54(1):67. doi: 10.2307/2095662.

Hamm, Jill V. 2000. “Do Birds of a Feather Flock Together? The Variable Bases for African American, Asian American, and European American Adolescents’ Selection of Similar Friends.” *Developmental Psychology* 36(2):209–19. doi: 10.1037/0012-1649.36.2.209.

Hargreaves, Dr David H., and David Hargreaves. 2006. *Social Relations in a Secondary School*. Routledge.

Harpalani, Vinay. 2002. *What Does ‘Acting White’ Really Mean? Racial Identity Formation and Academic Achievement among Black Youth*. *SSRN Scholarly Paper*. ID 3153173. Rochester, NY: Social Science Research Network.

Harris, Kathleen Mullan. 2013. “The Add Health Study: Design and Accomplishments.” doi: 10.17615/C6TW87.

Herpin, Nicolas. 1996. “Les Amis de Classe: Du Collège Au Lycée.” *Économie et Statistique* 293(1):125–36.

Hjalmarsson, Simon, and Carina Mood. 2015. “Do Poorer Youth Have Fewer Friends? The Role of Household and Child Economic Resources in Adolescent School-Class Friendships.” *Children and Youth Services Review* 57:201–11. doi: 10.1016/j.childyouth.2015.08.013.

Hoffman, Marion, Per Block, and Tom A. B. Snijders. 2020. “Modeling Partitions of Individuals.” *ArXiv:2009.13974 [Stat]*.

Hollingshead, August de Belmont. 1949. *Elmtown’s Youth: The Impact of Social Classes on Adolescents*. J. Wiley.

Hunter, David R. 2007. “Curved Exponential Family Models for Social Networks.” *Social Networks* 29(2):216–30. doi: 10.1016/j.socnet.2006.08.005.

Ichou, Mathieu. 2018. *Les enfants d’immigrés à l’école: Inégalités scolaires du primaire à l’enseignement supérieur*. Paris: Presses Universitaires de France.

Indlekofer, Natalie, and Ulrik Brandes. 2013. “Relative Importance of Effects in Stochastic Actor-Oriented Models.” *Network Science* 1(3):278–304. doi: 10.1017/nws.2013.21.

International Labor Office. 2012. *International Standard Classification of Occupations: Structure, Group Definitions and Correspondence Tables*. International Labour Office.

James, Allison, and Alan Prout. 1997. *Constructing and Reconstructing Childhood: Contemporary Issues in the Sociological Study of Childhood*. Psychology Press.

Kalter, Frank, Antony F. Heath, Miles Hewstone, Jan O. Jonsson, Matthijs Kalmijn, Irena Kogan, and Frank Tubergen. 2016. “Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) – Full Version. Data File for On-site Use.”

Kane, Danielle. 2011. “The Gendered Transition to College: The Role of Culture in Ego-Network Evolution.” *Poetics* 39(4):266–89. doi: 10.1016/j.poetic.2011.05.003.

Kossinets, Gueorgi, and Duncan J. Watts. 2009. “Origins of Homophily in an Evolving Social Network.” *American Journal of Sociology* 115(2):405–50. doi: 10.1086/599247.

Krackhardt, David. 1988. “Predicting with Networks: Nonparametric Multiple Regression Analysis of Dyadic Data.” *Social Networks* 10(4):359–81. doi: 10.1016/0378-8733(88)90004-4.

Krivitsky, Pavel N., and Mark S. Handcock. 2014. “A Separable Model for Dynamic Networks.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 76(1):29–46. doi: 10.1111/rssb.12014.

Kruse, Hanno, and Clemens Kroneberg. 2019. “More Than a Sorting Machine: Ethnic Boundary Making in a Stratified School System.” *American Journal of Sociology* 125(2):431–84. doi: 10.1086/705561.

Kruse, Hanno, Sanne Smith, Frank van Tubergen, and Ineke Maas. 2016. “From Neighbors to School Friends? How Adolescents’ Place of Residence Relates to Same-Ethnic School Friendships.” *Social Networks* 44:130–42. doi: 10.1016/j.socnet.2015.07.004.

Kuha, Jouni, and Colin Mills. 2018. “On Group Comparisons With Logistic Regression Models.” *Sociological Methods & Research* 004912411774730. doi: 10.1177/0049124117747306.

Labov, William. 1993. *Le parler ordinaire: la langue dans les ghettos noirs des États-Unis*. Les Ed. de Minuit.

Lacey, Colin. 1971. *Hightown Grammar: The School as a Social System*. Manchester University Press.

Lahire, Bernard. 2005. *L’homme pluriel: Les ressorts de l’action*. Armand Colin.

Lahire, Bernard. 2017. *Tableaux de familles. Heurs et malheurs scolaires en milieux populaires*. Média Diffusion.

Lakon, Cynthia M., John R. Hipp, Cheng Wang, Carter T. Butts, and Rupa Jose. 2015. “Simulating Dynamic Network Models and Adolescent Smoking: The Impact of Varying Peer Influence and Peer Selection.” *American Journal of Public Health* 105(12):2438–48. doi: 10.2105/AJPH.2015.302789.

Lareau, Annette. 2003. *Unequal Childhoods: Class, Race, and Family Life*. University of California Press.

Lazarseld, Paul, and Robert King Merton. 1954. “Friendship as a Social Process: A Substantive and Methodological Analysis.” in *Berger, M., Abel, T. and Charles, H., Eds., Freedom and Control in Modern Society*. New York.

Lazega, Emmanuel. 1999. “Le Phénomène Collégial: Une Théorie Structurale de l’action Collective Entre Pairs.” *Revue Française de Sociologie* 639–70.

Lazega, Emmanuel. 2014. *Réseaux sociaux et structures relationnelles: « Que sais-je ? » n° 3399*. Presses Universitaires de France.

Lazega, Emmanuel, and Tom A. B. Snijders. 2015. *Multilevel Network Analysis for the Social Sciences: Theory, Methods and Applications*. New York, NY: Springer Berlin Heidelberg.

Lehman-Frisch, Sonia, Jean-Yves Authier, and Frédéric Dufaux. 2012. “Les enfants et la mixité sociale dans les quartiers gentrifiés à Paris, Londres et San Francisco.” *Rapport de recherche CNAF* 185.

Leifeld, Philippe, and Skyler J. Cranmer. 2016. “A Theoretical and Empirical Comparison of the Temporal Exponential Random Graph Model and the Stochastic Actor-Oriented Model.” doi: arXiv:1506.06696v3.

Lepoutre, David. 1997. *Cœur de banlieue: Codes, rites, et langages*. Paris: Odile Jacob.

Lewis, Kevin, Marco Gonzalez, and Jason Kaufman. 2012. “Social Selection and Peer Influence in an Online Social Network.” *Proceedings of the National Academy of Sciences of the United States of America* 109(1):68–72.

Lignier, Wilfried, and Julie Pagis. 2012. “Quand les enfants parlent l’ordre social, Children Articulating Social Order.” *Politix* n° 99(3):23–49.

Lignier, Wilfried, and Julie Pagis. 2014. “Inimitiés enfantines, Childhood animosity.” *Genèses* n° 96(3):35–61.

Lomi, Alessandro, Tom A. B. Snijders, Christian E. G. Steglich, and Vanina Jasmine Torló. 2011. “Why Are Some More Peer than Others? Evidence from a Longitudinal Study of Social Networks and Individual Academic Performance.” *Social Science Research* 40(6):1506–20. doi: 10.1016/j.ssresearch.2011.06.010.

Lospinoso, Joshua A., Michael Schweinberger, Tom A. B. Snijders, and Ruth M. Ripley. 2011. “Assessing and Accounting for Time Heterogeneity in Stochastic Actor Oriented Models.” *Advances in Data Analysis and Classification* 5(2):147–76. doi: 10.1007/s11634-010-0076-1.

Lusher, Dean, Johan Koskinen, and Garry Robins. 2013. *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press.

Malacarne, Timothy. 2017. “Rich Friends, Poor Friends: Inter–Socioeconomic Status Friendships in Secondary School.” *Socius* 3:2378023117736994. doi: 10.1177/2378023117736994.

Marsden, Peter V. 1990. “Network Data and Measurement.” *Annual Review of Sociology* 1:435–63.

McFarland, Daniel A., James Moody, David Diehl, Jeffrey A. Smith, and Reuben J. Thomas. 2014. “Network Ecology and Adolescent Social Structure.” *American Sociological Review* 0003122414554001. doi: 10.1177/0003122414554001.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. “Birds of a Feather: Homophily in Social Networks.” *Annual Review of Sociology* 27(1):415–44. doi: 10.1146/annurev.soc.27.1.415.

Mercken, Liesbeth, Tom A. B. Snijders, Christian Steglich, and Hein de Vries. 2009. “Dynamics of Adolescent Friendship Networks and Smoking Behavior: Social Network Analyses in Six European Countries.” *Social Science & Medicine (1982)* 69(10):1506–14. doi: 10.1016/j.socscimed.2009.08.003.

Mercklé, Pierre. 2016. *La sociologie des réseaux sociaux*. La Découverte.

Merle, Pierre. 2012. “Carte scolaire et ségrégation sociale des établissements. Une analyse monographique des collèges rennais, Abstract, Resumen.” *Espaces et sociétés* (151):103–21.

Merle, Pierre. 2015. *L’élève humilié: L’école, un espace de non-droit ?* Presses Universitaires de France.

Merten, Don E. 1997. “The Meaning of Meanness: Popularity, Competition, and Conflict among Junior High School Girls.” *Sociology of Education* 70(3):175. doi: 10.2307/2673207.

Mignot, Jean-François. 2021. “Prénoms des descendants d’immigrés en France : Essai de reproduction d’un article scientifique.” 357.

Monchaux, Philippe. 2011. “Éditorial.” *Carrefours de l’education* n° 31(1):3–3.

Mood, Carina. 2010. “Logistic Regression: Why We Cannot Do What We Think We Can Do, and What We Can Do About It.” *European Sociological Review* 26(1):67–82. doi: 10.1093/esr/jcp006.

Moody, James. 2001. “Race, School Integration, and Friendship Segregation in America.” *American Journal of Sociology* 107(3):679–716. doi: 10.1086/338954.

Moody, James, Wendy D. Brynildsen, D. Wayne Osgood, Mark E. Feinberg, and Scott Gest. 2011. “Popularity Trajectories and Substance Use in Early Adolescence.” *Social Networks* 33(2):101–12. doi: 10.1016/j.socnet.2010.10.001.

Mouw, Ted, and Barbara Entwisle. 2006. “Residential Segregation and Interracial Friendship in Schools.” *American Journal of Sociology* 112(2):394–441. doi: 10.1086/506415.

Müller, Christoph Michael, Verena Hofmann, Janine Fleischli, and Felix Studer. 2016. “Classroom Peer Influence From the Entire Class, Dominant Students, and Friends.” *Journal of Cognitive Education and Psychology* 15(1):122–45. doi: 10.1891/1945-8959.15.1.122.

Oberti, Marco, and Edmond Préteceille. 2011. *Cadres supérieurs et professions intermédiaires dans l’espace urbain, entre séparatisme et mixité sous contrôle*. Armand Colin.

Oberti, Marco, Edmond Préteceille, and Clément Rivière. 2012. *Les Effets de l’assouplissement de La Carte Scolaire Dans La Banlieue Parisienne*. HALDE - défenseur des droits et DEPP - ministère de l’éducation nationale.

Octobre, Sylvie, Christine Détrez, Pierre Mercklé, and Nathalie Berthomier. 2010. *L’enfance des loisirs. Trajectoires communes et parcours individuels de la fin de l’enfance à la grande adolescence*. Ministère de la Culture - DEPS.

Oles, Andrzej. 2021. “Query Openrouteservice from R.” *Github*. Retrieved February 12, 2021 (https://giscience.github.io/openrouteservice-r/articles/openrouteservice.html).

Olsson, E. 2007. “The Economic Side of Social Relations: Household Poverty, Adolescents’ Own Resources and Peer Relations.” *European Sociological Review* 23(4):471–85. doi: 10.1093/esr/jcm016.

Palheta, Ugo. 2012. “Le collège divise. Appartenance de classe, trajectoires scolaires et enseignement professionnel.” *Sociologie* 2(4):363–86.

Palheta, Ugo. 2015. *La domination scolaire: Sociologie de l’enseignement professionnel et de son public*. Paris: Presses Universitaires de France.

Papapolydorou, Maria. 2014. "'When You See a Normal Person …': Social Class and Friendship Networks among Teenage Students." *British Journal of Sociology of Education* 35(4):559–77. doi: 10.1080/01425692.2013.782811.

Pasquier, Dominique. 2005. *Cultures lycéennes: la tyrannie de la majorité*. Autrement.

Passerieux, V. 2010. "L'intello Ou Le Bouffon." *Diversité* (162):91–95.

Passeron, Jean-Claude. [1991] 2013. *Le Raisonnement sociologique: Un espace non poppérien de l'argumentation*. Albin Michel.

Pattison, Philippa, and Garry Robins. 2002. "Neighborhood-Based Models for Social Networks." *Sociological Methodology* 32:301–37.

Pellegrini, Anthony D., and Jeffrey D. Long. 2002. "A Longitudinal Study of Bullying, Dominance, and Victimization during the Transition from Primary School through Secondary School." *British Journal of Developmental Psychology* 20(2):259–80.

Penissat, Étienne, and Jay Rowell. 2012. "Note de recherche sur la fabrique de la nomenclature socio-économique européenne ESeC." *Actes de la recherche en sciences sociales* n° 191-192(1):126–35.

Plenty, Stephanie, and Jan O. Jonsson. 2017. "Social Exclusion among Peers: The Role of Immigrant Status and Classroom Immigrant Density." *Journal of Youth and Adolescence* 46(6):1275–88. doi: 10.1007/s10964-016-0564-5.

Plenty, Stephanie, and Carina Mood. 2016. "Money, Peers and Parents: Social and Economic Aspects of Inequality in Youth Wellbeing." *Journal of Youth and Adolescence* 45(7):1294–1308. doi: 10.1007/s10964-016-0430-5.

Raveaud, Maroussia. 2003. "Minorités, ethnicité et citoyenneté : les modèles français et anglais sur les bancs de révole." *Revue française de pédagogie* 144(1):19–28. doi: 10.3406/rfp.2003.2966.

Rayou, Patrick. 1998. *La cité des lycéens*. Harmattan.

Rayou, Patrick, and Agnès van Zanten. 2004. *Enquête sur les nouveaux enseignants: changeront-ils l'école ?* Bayard.

Ripley, Ruth, and Tom A. B. Snijders. 2020. "Manual for SIENA Version 4.0." 65.

Robins, Garry, Philippa Pattison, and Jodie Woolcock. 2005. "Small and Other Worlds: Global Network Structures from Local Processes." *American Journal of Sociology* 110(4):894–936. doi: 10.1086/427322.

Robins, Garry, Pip Pattison, Yuval Kalish, and Dean Lusher. 2007. "An Introduction to Exponential Random Graph (P*) Models for Social Networks." *Social Networks* 29(2):173–91. doi: 10.1016/j.socnet.2006.08.002.

Rocher, Thierry. 2016. "Construction d'un Indice de Position Sociale Des Élèves." *Éducation & Formations* (90):5–27. doi: 10.48464/hal-01350095.

Rodkin, Philip C., Thomas W. Farmer, Ruth Pearl, and Richard Van Acker. 2006. “They’re Cool: Social Status and Peer Group Supports for Aggressive Boys and Girls.” *Social Development* 15(2):175–204. doi: 10.1046/j.1467-9507.2006.00336.x.

Rose, Amanda J., Lance P. Swenson, and Wendy Carlson. 2004. “Friendships of Aggressive Youth: Considering the Influences of Being Disliked and of Being Perceived as Popular.” *Journal of Experimental Child Psychology* 88(1):25–45. doi: 10.1016/j.jecp.2004.02.005.

Rouff-Fiorenzi, Katia. 2019. “Colonies de vacances.” *Lien Social* N° 1245(4):22–29.

Safi, Mirna. 2013. *Les inégalités ethno-raciales*. Editions La Découverte.

Salmivalli, Christina. 2010. “Bullying and the Peer Group: A Review.” *Aggression and Violent Behavior* 15(2):112–20. doi: 10.1016/j.avb.2009.08.007.

Schelling, Thomas C. 1969. “Models of Segregation.” *The American Economic Review* 59(2):488–93.

Shrum, Wesley, Neil H. Cheek, and Saundra MacD. Hunter. 1988. “Friendship in School: Gender and Racial Homophily.” *Sociology of Education* 61(4):227–39. doi: 10.2307/2112441.

Simmel, Georg. 1950. *The Sociology of Georg Simmel*. Simon and Schuster.

Simmel, Georg, and Everett C. Hughes. 1949. “The Sociology of Sociability.” *American Journal of Sociology* 55(3):254–61.

Sirota, Régine. 1988. *L’École primaire au quotidien*. FeniXX.

Smirnov, Ivan, and Stefan Thurner. 2017. “Formation of Homophily in Academic Performance: Students Change Their Friends Rather than Performance.” *PLOS ONE* 12(8):e0183473. doi: 10.1371/journal.pone.0183473.

Smith, Sanne, Ineke Maas, and Frank van Tubergen. 2014. “Ethnic Ingroup Friendships in Schools: Testing the by-Product Hypothesis in England, Germany, the Netherlands and Sweden.” *Social Networks* 39:33–45. doi: 10.1016/j.socnet.2014.04.003.

Snijders, T. A. B., and Chris Baerveldt. 2003. “A Multilevel Network Study of the Effects of Delinquent Behavior on Friendship Evolution.” *The Journal of Mathematical Sociology* 27(2–3):123–51. doi: 10.1080/00222500305892.

Snijders, Tom A. B. 2013. “Network Dynamics.” in *The Handbook of Rational Choice Social Research*. Standford University Press.

Snijders, Tom A. B. 2017. “Stochastic Actor-Oriented Models for Network Dynamics.” *Annual Review of Statistics and Its Application* 4(1):343–63. doi: 10.1146/annurev-statistics-060116-054035.

Snijders, Tom A. B., Gerhard G. van de Bunt, and Christian E. G. Steglich. 2010. “Introduction to Stochastic Actor-Based Models for Network Dynamics.” *Social Networks* 32(1):44–60. doi: 10.1016/j.socnet.2009.02.004.

Snijders, Tom A. B., Philippa E. Pattison, Garry L. Robins, and Mark S. Handcock. 2006. "New Specifications for Exponential Random Graph Models." *Sociological Methodology* 36(1):99–153. doi: 10.1111/j.1467-9531.2006.00176.x.

Snijders, Tom A. B., and Christian E. G. Steglich. 2015. "Representing Micro–Macro Linkages by Actor-Based Dynamic Network Models." *Sociological Methods & Research* 44(2):222–71. doi: 10.1177/0049124113494573.

Stadtfeld, Christoph. 2018. "The Micro–Macro Link in Social Networks." Pp. 1–15 in *Emerging Trends in the Social and Behavioral Sciences*. American Cancer Society.

Stark, Tobias H., and Andreas Flache. 2012. "The Double Edge of Common Interest: Ethnic Segregation as an Unintended Byproduct of Opinion Homophily." *Sociology of Education* 85(2):179–99. doi: 10.1177/0038040711427314.

Testanière, Jacques. 1967. "Chahut Traditionnel et Chahut Anomique Dans l'enseignement Du Second Degré." *Revue Française de Sociologie* 8:17. doi: 10.2307/3319085.

Thaning, Max, and Martin Hällsten. 2020. "The End of Dominance? Evaluating Measures of Socio-Economic Background in Stratification Research." *European Sociological Review* 36:533–47. doi: 10.1093/esr/jcaa009.

Troop-Gordon, Wendy, Kari J. Visconti, and Kayla J. Kuntz. 2011. "Perceived Popularity During Early Adolescence: Links to Declining School Adjustment Among Aggressive Youth." *The Journal of Early Adolescence* 31(1):125–51. doi: 10.1177/0272431610384488.

Tuma, Nancy Brandon, and Maureen T. Hallinan. 1979. "The Effects of Sex, Race, and Achievement on Schoolchildren's Friendships." *Social Forces* 57(4):1265–85. doi: 10.2307/2577270.

Turner, John C. 1987. *Rediscovering the Social Group: A Self-Categorization Theory*. Basil Blackwell.

Verbrugge, Lois. 1977. "The Structure of Adult Friendship Choices." *Social Forces* 56:576. doi: 10.2307/2577741.

Wasserman, Stanley, Katherine Faust, and Stanley (University of Illinois Wasserman Urbana-Champaign. 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Whyte, William Foote. [1943] 1993. *Street Corner Society: The Social Structure of an Italian Slum*. University of Chicago Press.

Willis, Paul E. 1977. *Learning to Labor: How Working Class Kids Get Working Class Jobs*. Columbia University Press.

Willmott, Peter. 1966. *Adolescent Boys of East London*. Routledge & K. Paul.

Wittek, Rafael, Tom A. B. Snijders, and Victor Nee. 2013. *The Handbook of Rational Choice Social Research*. Stanford University Press.

Young, Michael F. D. 1972. “Knowledge and Control: New Directions for the Sociology of Education.” *British Journal of Educational Studies* 20(2):247.

van Zanten, Agnès. 2000. “Le quartier ou l’école ? Déviance et sociabilité adolescente dans un collège de banlieue.” *Déviance et société* 24(4):377–401. doi: 10.3406/ds.2000.1737.

van Zanten, Agnes. 2003. “Middle-Class Parents and Social Mix in French Urban Schools: Reproduction and Transformation of Class Relations in Education.” *International Studies in Sociology of Education* 13(2):107–24. doi: 10.1080/09620210300200106.

van Zanten, Agnès. 2012. *L’école de la périphérie*. Presses Universitaires de France.

van Zanten, Agnès Van. 2015. *Choisir son école: Stratégies familiales et médiations locales*. Presses Universitaires de France.

Zárate, Román Andrés(Zárate Vasquez), Joshua David Angrist, and Parag A. Pathak. 2019. “Essays in the Economics of Education.” Thesis, Massachusetts Institute of Technology.

Zarca, Bernard. 1999. “Le sens social des enfants.” *Sociétés contemporaines* 36(1):67–101. doi: 10.3406/socco.1999.1731.

Appendices

List of Appendices

Appendices for chapter 3

Appendix 3A: Exploratory Analyses about the Impact of the Covid Crisis on Student Sociability

Friendship Networks between Waves 4 and 6

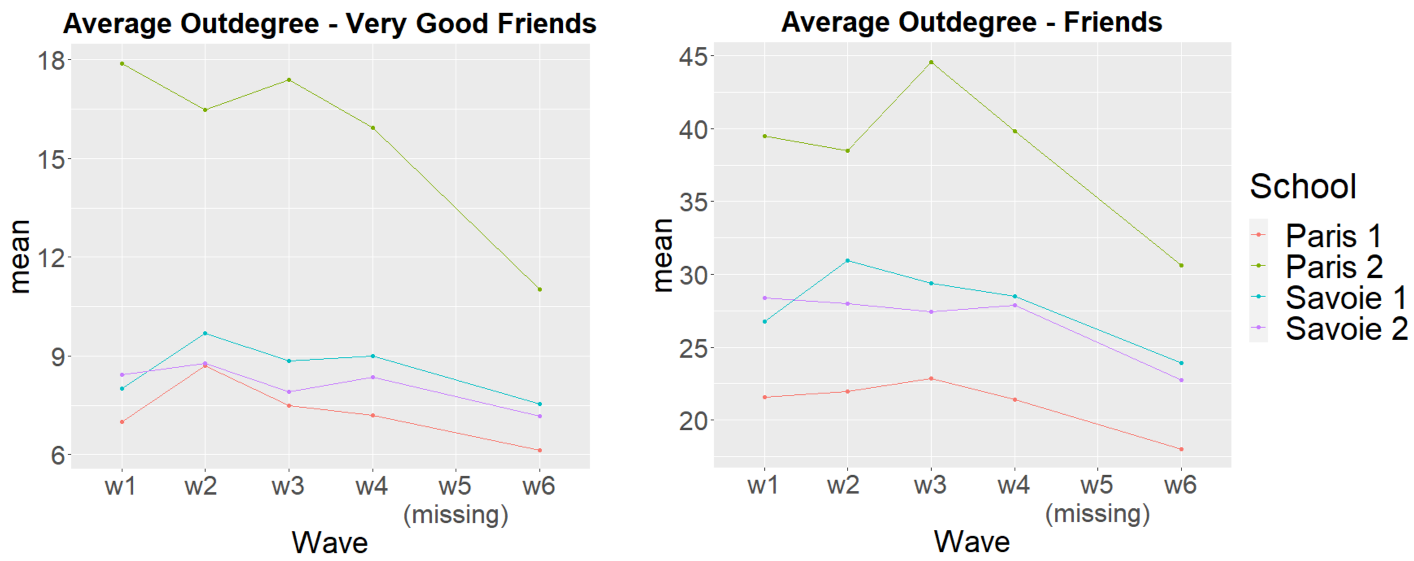
The first way to look at the impact of the lock-down period of the Spring 2020 on student sociability is to look at changes in the friendship networks between waves 4 and 6. Since we do not have a control group not exposed to the treatment (i.e. the lock-down), it is difficult to say if the evolution between these waves is due to its impact, or to the normal aging process of teenagers. However, we can see if there are abrupt changes in trends compared to the period from wave 1 to 4, which would hint at a potential effect of the COVID crisis (aging processes might be expected to be relatively homogeneous over middle school years, although this is not guaranteed either).

Figure 1 shows the evolution of the average outdegree (number of emitted nominations per student) in the friendship networks (very good friends and friends) over time. In all four schools and for both networks, there is a clear drop in the average number of friends declared by students. In some cases, the decrease seems to have started as early as wave 3 (Paris 2, perhaps Paris 1), but for other it only starts in wave 4 (Savoie 1 and 2). Moreover, even for Paris 2, the magnitude of the decrease between wave 4 and 6 is much larger than that of previous waves. Therefore, it seems plausible that the COVID crisis and related lock-down induced a reduction in the number of friends that students could keep. Possible mechanisms may be the difficulty of keeping a large number of social contacts while being stuck at home; a reduction in the number of foci and opportunities for social contact, even after the lock-down was over (e.g. many sport clubs stayed shut for several months in Fall 2020 and Winter 2020-21 as well); an increase in stress for some students, making them less inclined to positive contacts with peers; or perhaps an effect of wearing surgical masks on school ground, which might make interactions in large groups more difficult (muffled voices and lack of facial expression could increasingly hinder communication as group size increases). With that said, the reduction of the number of declared friends as students transition from childhood to adolescence is a relatively expected process (children tend to have a particularly extensive conception of what a “friend” is; Bidart 2010), so it does not seem unreasonable to think that aging played a role as well (moreover, the fact that students in wave 6 entered their last year of middle school might contribute to a form of threshold pattern in aging processes, as they suddenly become the “big ones” in the school). The hypothesis that COVID played a role is plausible, but uncertain at this stage – comparison with similar cohorts of students in recent years would be needed to confirm it.

Note that apart from density, there is no clear structural changes in friendship networks in wave 6. Transitivity remains fairly stable. There is a slight drop in reciprocity, as well as the proportion of ties carried over from wave 4 to 6 is lower compared to previous transitions, but these are both logical correlates of the drop in density (for density to drop, more ties need to disappear between waves 4 and 6; as for reciprocity, the ties disappearing between the two waves are primarily weak ties, meaning that they were more frequently unreciprocated). See Appendix 3B below for further descriptives about the friendship networks in time. Finally, to the extent that the drop in density between waves 4 and 6 should largely concern weak ties, this might induce a slight increase in socioeconomic homophily (weaker ties tend to be less homophilic, cf. chapter 4). This might contribute to the slight increase in socioeconomic homophily observed in certain schools in wave 6, but the effect remains very slight at the aggregate level – granted that it may be counter-balanced by other social processes linked to time and aging (cf. chapter 4, sections 2 and 3)[1]. As things stand, it is impossible to conclude that the lock-down episode had a clear impact on homophily patterns among students.

1 There is a very strong increase in socioeconomic homophily in Paris 1, but it started before wave 4. In fact, the one school where the transition from wave 4 to 6 seems more socioeconomically homophilic than others is Savoie 1: multivariate models suggest an effect of socioeconomic homophilic selection in the transition from wave 4 to 6 there, which was completely absent in previous waves (see chapter 5, section 1.3.6). However, given that this is a single case, and taking into account the uncertainty associated with these measures – multivariate network models are particularly complex (again, see chapter 5) –, the evidence remains too slight for a definite conclusion.

Figure 1: Average Outdegree in the Friendship Networks per Wave and School



Networks of Distant Contact during Lock-Down

In order to further investigate the impact of the lock-down on student sociability, dedicated items were included in the questionnaires from wave 6 (Fall 2020). Students were asked four questions:

- "During the lock-down, did you meet other students from the school in person (for instance during the daily hour allowed outside[2])?"
- "During the lock-down, did you talk with other students from the school through remote audio or video conversations (phone, zoom, skype, whatsapp, facetime, chat rooms of video games, etc)?"
- "During the lock-down, did you talk with other students from the school through written messages (text, email, messenger, instagram or snapchat groups, messaging apps, etc)?"
- "During the lock-down, did you work with other students? For instance, doing your homework together, helping each other out for lessons, or sharing your notes on lessons?"[3]

2 During the lock-down of Spring 2020, French citizens were allowed one hour per day of going outside for walks or sport (they had to fill administrative authorization files giving the precise time of the outing, and carry it with them in case of a police control).

3 Schools tried to provide online courses and homework during the lock-down (though the drop-out rates were high, and many students hardly worked during the period).

For each of these four questions, students could tick one out of four boxes indicating the frequency with which they had the considered type of contact: “never”, “rarely (two or three times during the entire lock-down”, “sometimes (about once a week or every two weeks)” and “often (several times a week)”.

Then, for all the answers except “never”, they were asked to nominate up to 10 schoolmates with whom they had talked (again, for the four types of contact separately). This results in four networks of lock-down contact, which I will call “face-to-face”, “vocal”, “written” and “work” networks. By construction, these networks only concern students that were here both in wave 4 and wave 6: those that left between the waves were not interrogated in wave 6, and those that arrived at the start of the fourth year of middle school were not yet in the school when the lock-down was implemented at the end of the third year (i.e. what should have been wave 5).

Table 1 provides descriptives about these lock-down networks for each school. Unsurprisingly, the network of face-to-face interactions is very sparse, although a few students nevertheless saw one another during the lock-down (mostly students living extremely close from one another). The networks of written and vocal contacts, on the other hand, are relatively dense – slightly less than the 5-friend networks (cf. Appendix 3B), which means that students declared between 3 and 4 contacts on average. There are also few isolates (around ten in each school), meaning that most students had some distant contact at least during the lock-down. We can also note that these networks are very weakly centralized, which means that there are not many inequalities in the distribution of distant contacts (much less so than for in-school friendship networks).

Figure 2 gives the number of nominations in the lock-down networks per type of relationship in the friendship networks of wave 4. The majority of the peers with whom students were in contact during the lock-down were already very good friends in wave 4 (more or less two thirds) or at least friends (more or less one quarter). This shows that these networks are mostly a continuation of in-school friendship networks, and particularly of strong friendships[4]. In other words, during the lock-down, students seem to have mostly

4 One might be concerned that students under-declared the number of contacts that they had during the lock-down due to the format of the question (directly writing up to ten names on the questionnaire sheet, instead of ticking boxes in an exhaustive list like for the items pertaining to in-school friendships). However, with 3 to 4 nominations per student on average, lock-down nominations are very far from the upward limit of 10 nominations. Therefore, it does not seem that the format of the question limited the number of emitted nominations. With that said, students may still have forgotten some of the contacts they had during lock-down (especially since these are recollection questions, taking place several months after the lock-down), so

maintained contact with their closest friends. This may also contribute to explaining the drop in friendship density between waves 4 and 6: “weak friends” did not keep contact during the lock-down, thus the dissolution of certain ties during this period.

Are there socioeconomic differences in how students kept contact during the lock-down? **Table 2** shows the results of multinomial logit models whose dependent variables are the frequencies of distant contact declared by students (four-post ordinal categories: “never”, “rarely”, “sometimes” and “often”; “never” is taken as the reference category). Independent variables are students’ school, gender, and socioeconomic background (Socio-Academic Scores). Note that the model is only fitted to the frequencies of vocal, written and work contact – face-to-face contact is simply too rare, such that none of the estimated parameters would be statistically significant. For vocal contact, there seems to be no statistically significant difference depending on socioeconomic background. However, for written contact, upper-background students are a bit more likely to have answered “often” rather than “never” (log-odd coefficient of 0.24, significant at the 90% threshold). More importantly, they are much more likely to have declared higher frequencies for work contact (i.e. having worked with peers during the lock-down). This seems to reflect academic and cultural inequalities more than relational ones: upper-background students did not necessarily have more contact with friends during the lock-down, but they resorted to written communication mediums (slightly) more often, and they organized co-working among friends more often. Most likely, this is due to the fact that they followed lessons and maintained some academic continuity during the lock-down more easily than their lower-background peers.

Finally, **Table 3** shows dyad-independent ERGMs applied to the four networks of distant contact, in the four schools separately. Similar to the basic models of chapter 4, only four parameters are included: density, SAS homophily (absolute difference of SA scores), SAS emission and SAS reception. There is socioeconomic homophily in the four types of networks, more or less of the same magnitude as in the friendship networks of the corresponding schools. Effect sizes are probably a bit larger than in the networks of very good friends (cf. chapter 4), at least for Paris 1 and Savoie 1. This probably reflects the fact that the friends with whom students kept contact during the lock-down are closer friends on average – and we know that socioeconomic homophily tends to be stronger for more demanding ties (cf. chapter 4 as well). The only school with no socioeconomic homophily is Savoie 2, which is also coherent with what is observed in the in-school friendship networks.

these estimates should be taken with caution indeed.

Regarding the emission and reception effects, it is interesting to note that upper-class students tend to receive and emit more nominations in three out of the four schools for vocal, written and working contact (not every single coefficient is statistically significant, but the overall pattern over schools and types of nominations is very clear). This is the case in Paris 1, Savoie 1 and Savoie 2, but not in Paris 2. Interestingly, we saw above (multinomial model of Table 2) that declared frequencies of contact do not differ much across socioeconomic backgrounds for vocal and written contacts. What this suggests is that upper-class students in these three schools had contacts with more friends during the lock-down (higher in- and outdegrees in the networks), but that these contacts were not more frequent altogether (i.e. working-class students spent more time talking with the same persons).

Table 1: Descriptives for the Lock-Down Networks

	Face-to-face	Vocal	Written	Work
Paris 1				
Nb of Nodes	73	73	73	73
Density	0.006	0.039	0.044	0.028
Reciprocity (% of reciprocated ties)	0.28	0.65	0.60	0.52
Transitivity	0.56	0.53	0.45	0.41
Indegree Centralization (Freeman Index)	0.04	0.07	0.07	0.06
Outdegree Centralization (Freeman Index)	0.08	0.06	0.10	0.08
Nb of Isolates (no incoming nor outgoing tie)	49	4	4	12
Paris 2				
Nb of Nodes	215	215	215	215
Density	0.002	0.014	0.016	0.008
Reciprocity (% of reciprocated ties)	0.31	0.55	0.54	0.56
Transitivity	0.42	0.35	0.29	0.42
Indegree Centralization (Freeman Index)	0.02	0.06	0.03	0.03
Outdegree Centralization (Freeman Index)	0.03	0.03	0.03	0.04
Nb of Isolates (no incoming nor outgoing tie)	126	19	8	46
Savoie 1				
Nb of Nodes	162	162	162	162
Density	0.003	0.014	0.020	0.006
Reciprocity (% of reciprocated ties)	0.36	0.54	0.62	0.46
Transitivity	0.47	0.38	0.33	0.45
Indegree Centralization (Freeman Index)	0.02	0.04	0.04	0.02
Outdegree Centralization (Freeman Index)	0.03	0.04	0.04	0.02
Nb of Isolates (no incoming nor outgoing tie)	87	14	5	41
Savoie 2				
Nb of Nodes	167	167	167	167
Density	0.002	0.020	0.020	0.010
Reciprocity (% of reciprocated ties)	0.37	0.62	0.59	0.55
Transitivity	0	0.43	0.36	0.49
Indegree Centralization (Freeman Index)	0.02	0.05	0.04	0.03
Outdegree Centralization (Freeman Index)	0.02	0.04	0.04	0.05
Nb of Isolates (no incoming nor outgoing tie)	102	12	12	24

Figure 2: Number of Lock-Down Nominations per Type of Friendship Tie in Wave 4

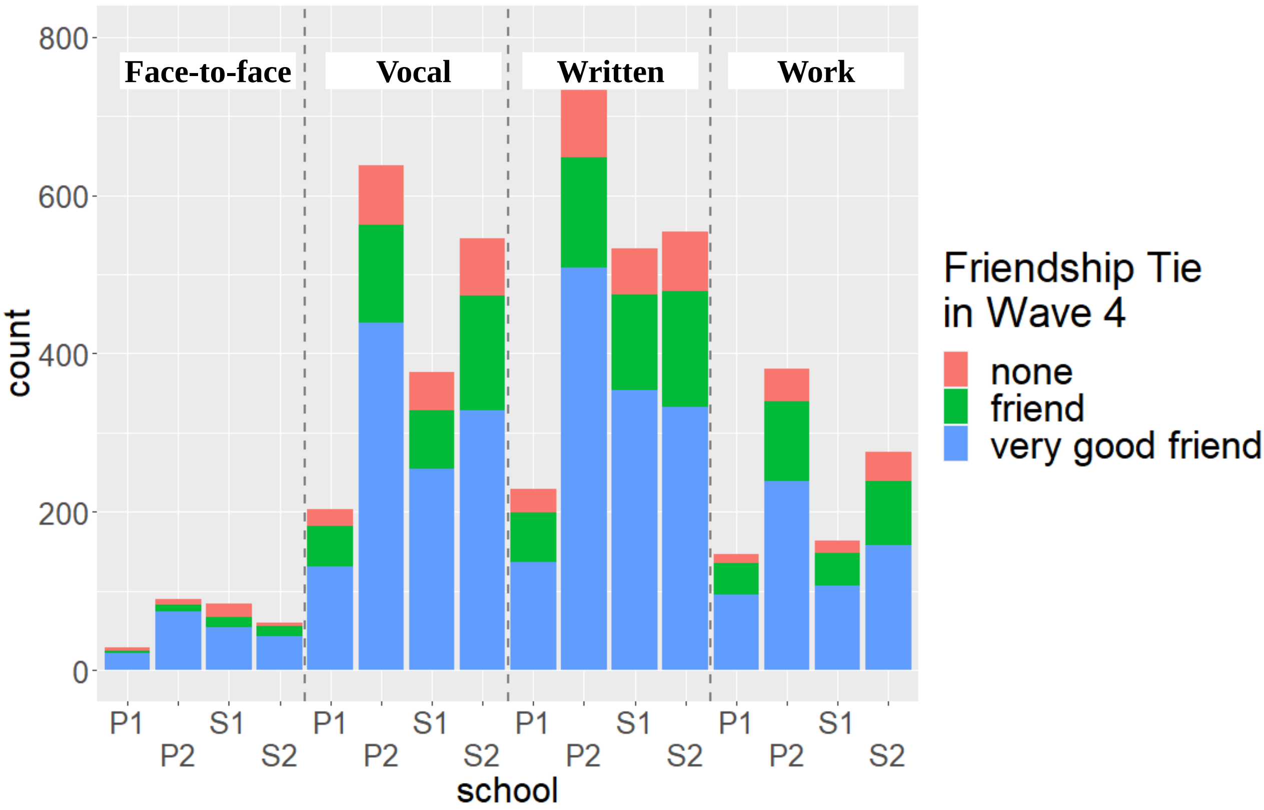


Table 2: Multinomial Regression on the Declared Frequency of Contact during the Lock-Down Period

	Vocal			Written			Work		
	Rarely	Sometimes	Often	Rarely	Sometimes	Often	Rarely	Sometimes	Often
Intercept	-0.71* (0.41)	0.15 (0.33)	1.66*** (0.26)	-1.61*** (0.57)	-0.1 (0.38)	2.11*** (0.29)	-0.55** (0.27)	-0.47* (0.28)	0.49** (0.22)
SAS	0.18 (0.17)	-0.18 (0.15)	0.08 (0.11)	0.15 (0.2)	0.01 (0.17)	0.24* (0.14)	0.27** (0.12)	0.5*** (0.12)	0.29*** (0.11)
Sex – male	-0.09 (0.33)	-0.06 (0.29)	-0.28 (0.22)	0.07 (0.42)	-0.07 (0.34)	-0.85*** (0.27)	-0.52** (0.24)	-1.01*** (0.25)	-0.97*** (0.21)
School – Paris 2	-0.02 (0.46)	-0.1 (0.37)	-0.22 (0.3)	1.49*** (0.59)	0.75* (0.41)	0.25 (0.33)	0.21 (0.31)	-0.06 (0.33)	0 (0.25)
School – Savoie 1	0.39 (0.44)	-0.53 (0.39)	-0.61** (0.3)	1.36** (0.61)	0.35 (0.45)	0.28 (0.35)	-0.1 (0.31)	-0.09 (0.32)	-1.53*** (0.32)
school – Paris 1	-0.23 (0.69)	-0.09 (0.53)	0.04 (0.41)	0.73 (0.78)	0.43 (0.54)	-0.07 (0.43)	-1* (0.59)	0.46 (0.41)	0.09 (0.34)

Note: the category of reference is “never”. The frequency of face-to-face contact is not modeled (most students answered “never”).

Table 3: Socioeconomic Homophily, Emission and Reception in the Networks of Lock-Down Contact (ERGM)

		Face-to-face	Vocal	Written	Work
Paris 1	Density	-4.898*** (0.348)	-2.212*** (0.109)	-2.511*** (0.113)	-2.946*** (0.148)
	Absdiff SAS	-0.313 (0.306)	-1.122*** (0.110)	-0.760*** (0.101)	-0.990*** (0.131)
	SAS receiver	-0.077 (0.256)	0.262*** (0.082)	0.103 (0.078)	0.324*** (0.099)
	SAS sender	-0.848*** (0.254)	0.242*** (0.082)	0.149* (0.078)	0.397*** (0.099)
Paris 2	Density	-5.866*** (0.167)	-3.994*** (0.061)	-4.123*** (0.065)	-4.799*** (0.086)
	Absdiff SAS	-0.374** (0.154)	-0.116** (0.048)	-0.126** (0.052)	0.041 (0.065)
	SAS receiver	0.137 (0.124)	-0.008 (0.040)	0.015 (0.043)	-0.054 (0.054)
	SAS sender	-0.218* (0.124)	-0.012 (0.040)	-0.039 (0.043)	-0.085 (0.054)
Savoie 1	Density	-5.325*** (0.177)	-3.711*** (0.075)	-3.949*** (0.086)	-5.005*** (0.142)
	Absdiff SAS	-0.443*** (0.163)	-0.215*** (0.059)	-0.290*** (0.072)	-0.214** (0.106)
	SAS receiver	-0.221* (0.129)	0.107** (0.047)	0.061 (0.057)	0.182** (0.085)
	SAS sender	0.199 (0.129)	0.131*** (0.047)	0.019 (0.057)	0.304*** (0.085)
Savoie 2	Density	-6.017*** (0.221)	-3.917*** (0.075)	-3.898*** (0.075)	-4.619*** (0.107)
	Absdiff SAS	-0.225 (0.179)	-0.037 (0.055)	-0.080 (0.056)	-0.115 (0.078)
	SAS receiver	0.324** (0.145)	0.063 (0.045)	0.095** (0.046)	0.128** (0.063)
	SAS sender	-0.109 (0.144)	0.106** (0.045)	0.102** (0.046)	0.214*** (0.063)

Appendix 3B: Descriptives about the Main Networks used in the Dissertation

Table 1: Descriptives – “Very good Friend” Networks

	wave 1	wave 2	wave 3	wave 4	wave 6
Paris 1					
Nb of Nodes	83	82	82	80	81
Density	0.093	0.109	0.097	0.092	0.080
Reciprocity (% of reciprocated ties)	0.58	0.64	0.59	0.59	0.61
Transitivity	0.42	0.43	0.38	0.44	0.40
Indegree Centralization (Freeman Index)	0.11	0.16	0.15	0.10	0.13
Outdegree Centralization (Freeman Index)	0.42	0.23	0.43	0.27	0.17
Proportion of ties maintained to the next wave	0.70	0.57	0.61	0.56	-
Paris 2					
Nb of Nodes	264	259	257	242	233
Density	0.069	0.065	0.069	0.067	0.049
Reciprocity (% of reciprocated ties)	0.49	0.54	0.51	0.50	0.56
Transitivity	0.33	0.32	0.34	0.33	0.31
Indegree Centralization (Freeman Index)	0.14	0.17	0.17	0.12	0.10
Outdegree Centralization (Freeman Index)	0.34	0.44	0.50	0.45	0.34
Proportion of ties maintained to the next wave	0.55	0.59	0.61	0.46	-
Savoie 1					
Nb of Nodes	169	178	178	179	180
Density	0.051	0.057	0.051	0.051	0.043
Reciprocity (% of reciprocated ties)	0.49	0.50	0.53	0.56	0.61
Transitivity	0.29	0.30	0.33	0.34	0.33
Indegree Centralization (Freeman Index)	0.15	0.11	0.12	0.10	0.06
Outdegree Centralization (Freeman Index)	0.15	0.25	0.26	0.16	0.18
Proportion of ties maintained to the next wave	0.54	0.52	0.58	0.49	-
Savoie 2					
Nb of Nodes	219	207	204	193	180
Density	0.039	0.044	0.040	0.044	0.041
Reciprocity (% of reciprocated ties)	0.50	0.54	0.58	0.57	0.63
Transitivity	0.26	0.26	0.26	0.26	0.31
Indegree Centralization (Freeman Index)	0.08	0.08	0.07	0.06	0.07
Outdegree Centralization (Freeman Index)	0.17	0.15	0.14	0.13	0.10
Proportion of ties maintained to the next wave	0.55	0.50	0.59	0.41	-

Table 2: Descriptives – “Friend” Networks (friend + very good friend nominations)

	wave 1	wave 2	wave 3	wave 4	wave 6
Paris 1					
Nb of Nodes	83	82	82	80	81
Density	0.29	0.27	0.30	0.27	0.23
Reciprocity (% of reciprocated ties)	0.71	0.68	0.70	0.70	0.70
Transitivity	0.52	0.52	0.52	0.51	0.49
Indegree Centralization (Freeman Index)	0.24	0.19	0.20	0.22	0.22
Outdegree Centralization (Freeman Index)	0.55	0.63	0.36	0.46	0.26
Proportion of ties maintained to the next wave	0.70	0.72	0.72	0.66	-
Paris 2					
Nb of Nodes	264	259	257	242	233
Density	0.15	0.15	0.198	0.17	0.13
Reciprocity (% of reciprocated ties)	0.61	0.63	0.61	0.60	0.60
Transitivity	0.39	0.37	0.41	0.40	0.38
Indegree Centralization (Freeman Index)	0.23	0.21	0.25	0.24	0.21
Outdegree Centralization (Freeman Index)	0.36	0.36	0.48	0.48	0.46
Proportion of ties maintained to the next wave	0.61	0.72	0.66	0.55	-
Savoie 1					
Nb of Nodes	169	178	178	179	180
Density	0.17	0.18	0.17	0.16	0.13
Reciprocity (% of reciprocated ties)	0.61	0.61	0.62	0.62	0.66
Transitivity	0.38	0.40	0.41	0.40	0.37
Indegree Centralization (Freeman Index)	0.24	0.27	0.18	0.19	0.14
Outdegree Centralization (Freeman Index)	0.53	0.39	0.55	0.61	0.26
Proportion of ties maintained to the next wave	0.63	0.66	0.64	0.56	-
Savoie 2					
Nb of Nodes	219	207	204	193	180
Density	0.13	0.14	0.14	0.15	0.13
Reciprocity (% of reciprocated ties)	0.59	0.60	0.62	0.63	0.65
Transitivity	0.34	0.33	0.34	0.34	0.34
Indegree Centralization (Freeman Index)	0.16	0.15	0.16	0.18	0.18
Outdegree Centralization (Freeman Index)	0.31	0.31	0.36	0.31	0.30
Proportion of ties maintained to the next wave	0.59	0.66	0.63	0.52	-

Table 3: Descriptives – 5-Friends Networks ("the five friends you spend the most time with at school")

	wave 1	wave 2	wave 3	wave 4	wave 6
Paris 1					
Nb of Nodes	83	82	82	80	81
Density	0.052	0.056	0.053	0.053	0.053
Reciprocity (% of reciprocated ties)	0.59	0.60	0.67	0.63	0.65
Transitivity	0.45	0.34	0.43	0.39	0.48
Indegree Centralization (Freeman Index)	0.08	0.08	0.09	0.06	0.09
Outdegree Centralization (Freeman Index)	0.01	0.01	0.01	0.01	0.01
Proportion of ties maintained to the next wave	0.53	0.51	0.63	0.60	-
Paris 2					
Nb of Nodes	264	259	257	242	233
Density	0.017	0.018	0.018	0.018	0.019
Reciprocity (% of reciprocated ties)	0.54	0.53	0.57	0.57	0.64
Transitivity	0.30	0.30	0.34	0.32	0.35
Indegree Centralization (Freeman Index)	0.05	0.04	0.05	0.04	0.03
Outdegree Centralization (Freeman Index)	0.00	0.00	0.00	0.00	0.00
Proportion of ties maintained to the next wave	0.47	0.48	0.50	0.49	-
Savoie 1					
Nb of Nodes	169	178	178	179	180
Density	0.024	0.023	0.023	0.023	0.024
Reciprocity (% of reciprocated ties)	0.54	0.54	0.59	0.63	0.66
Transitivity	0.36	0.34	0.45	0.444	0.45
Indegree Centralization (Freeman Index)	0.05	0.07	0.05	0.05	0.04
Outdegree Centralization (Freeman Index)	0.01	0.01	0.01	0.01	0.01
Proportion of ties maintained to the next wave	0.41	0.50	0.50	0.50	-
Savoie 2					
Nb of Nodes	219	207	204	193	180
Density	0.020	0.021	0.021	0.022	0.025
Reciprocity (% of reciprocated ties)	0.57	0.58	0.60	0.60	0.69
Transitivity	0.37	0.30	0.32	0.34	0.43
Indegree Centralization (Freeman Index)	0.054	0.03	0.03	0.04	0.03
Outdegree Centralization (Freeman Index)	0.00	0.00	0.00	0.00	0.00
Proportion of ties maintained to the next wave	0.45	0.59	0.47	0.39	-

Table 4: Descriptives – “Dislike” Networks

	wave 1	wave 2	wave 3	wave 4	wave 6
Paris 1					
Nb of Nodes	83	82	82	80	81
Density	0.077	0.057	0.050	0.041	0.044
Reciprocity (% of reciprocated ties)	0.23	0.16	0.15	0.16	0.25
Transitivity	0.21	0.24	0.18	0.17	0.14
Indegree Centralization (Freeman Index)	0.18	0.37	0.23	0.31	0.14
Outdegree Centralization (Freeman Index)	0.27	0.33	0.34	0.52	0.57
Proportion of ties maintained to the next wave	0.25	0.31	0.21	0.26	-
Paris 2					
Nb of Nodes	264	259	257	242	233
Density	0.033	0.024	0.026	0.022	0.019
Reciprocity (% of reciprocated ties)	0.21	0.16	0.19	0.15	0.18
Transitivity	0.18	0.12	0.13	0.11	0.10
Indegree Centralization (Freeman Index)	0.09	0.10	0.16	0.14	0.09
Outdegree Centralization (Freeman Index)	0.24	0.17	0.21	0.18	0.15
Proportion of ties maintained to the next wave	0.22	0.33	0.27	0.19	-
Savoie 1					
Nb of Nodes	169	178	178	179	180
Density	0.051	0.041	0.042	0.047	0.035
Reciprocity (% of reciprocated ties)	0.18	0.19	0.18	0.16	0.07
Transitivity	0.18	0.16	0.14	0.21	0.19
Indegree Centralization (Freeman Index)	0.17	0.23	0.22	0.25	0.16
Outdegree Centralization (Freeman Index)	0.28	0.22	0.30	0.30	0.26
Proportion of ties maintained to the next wave	0.28	0.30	0.36	0.29	-
Nb of Nodes	219	207	204	193	180
Density	0.050	0.048	0.045	0.038	0.029
Reciprocity (% of reciprocated ties)	0.18	0.17	0.18	0.18	0.15
Transitivity	0.18	0.17	0.17	0.13	0.12
Indegree Centralization (Freeman Index)	0.14	0.16	0.12	0.09	0.09
Outdegree Centralization (Freeman Index)	0.25	0.22	0.21	0.34	0.23
Proportion of ties maintained to the next wave	0.26	0.36	0.29	0.21	-

Appendix 3C: A Quick Presentation of Popularity Measures

Four questions pertained to popularity in the questionnaires. They were part of the permanent module, but only starting in wave 2; in wave 1, only the first question ('cool' item) was asked. The questions were:

1. *"Some students are sometimes considered to be "cool" or "popular" by their classmates. Do you think this applies to some students in your grade-level? If yes, which ones? (you can name up to 5 persons)*
2. *Among your grade-level, are there students that almost everyone in the school knows about? (you can name up to 5 persons)*
3. *Among you grade-level, are there students that many people fear? (you can name up to 5 persons)*
4. *Among your grade-level, are there students that many people like? (you can name up to 5 persons)*

NB: the questions were asked in this order, all on the same page.

Table 1: Questionnaire Items Related to Popularity

Format of the question	Waves	Item	Average outdeg.	Median outdeg.	St. dev. for outdeg.
Name up to 5 persons	All	Cool	1.12	0	1.61
	All except wave 1	Known	0.78	0	1.30
	All except wave 1	Feared	0.18	0	0.57
	All except wave 1	Liked	0.67	0	1.23

Table 2: Connection between Popularity Items: Proportion of the Nominations in the Row Item that also Correspond to a Nomination in the Column Item (averaged over all waves and schools)

	Cool	Known	Fear	Liked
Cool	1	0.38	0.03	0.26
Known	0.50	1	0.04	0.26
Fear	0.15	0.18	1	0.06
Liked	0.40	0.29	0.02	1

Table 3: Socioeconomic Homophily, Emission and Reception in Popularity Networks (ERGM) – Wave 2

		Cool	**Known**	**Fear**	**Liked**
Paris 1	Density	-3.679*** (0.154)	-4.619*** (0.232)	-5.307*** (0.374)	-4.323*** (0.226)
	Absdiff SAS	-0.360*** (0.116)	-0.215 (0.160)	-0.306 (0.361)	-0.651*** (0.198)
	SAS receiver	0.070 (0.092)	0.169 (0.128)	-1.049*** (0.305)	0.532*** (0.159)
	SAS sender	0.093 (0.092)	0.045 (0.128)	0.038 (0.308)	-0.175 (0.158)
Paris 2	Density	-4.959*** (0.091)	-5.215*** (0.104)	-7.344*** (0.285)	-5.403*** (0.110)
	Absdiff SAS	-0.429*** (0.085)	-0.463*** (0.096)	-0.112 (0.247)	-0.346*** (0.095)
	SAS receiver	0.087 (0.068)	0.109 (0.076)	-0.651*** (0.210)	0.084 (0.076)
	SAS sender	-0.022 (0.068)	0.096 (0.076)	-0.194 (0.211)	0.168** (0.076)
Savoie 1	Density	-4.890*** (0.127)	-5.161*** (0.139)	-7.429*** (0.454)	-5.308*** (0.157)
	Absdiff SAS	-0.185** (0.093)	-0.068 (0.107)	-0.473 (0.536)	-0.203* (0.117)
	SAS receiver	0.035 (0.075)	-0.247*** (0.088)	-0.994** (0.462)	0.132 (0.094)
	SAS sender	0.408*** (0.075)	0.185** (0.088)	0.642 (0.464)	0.233** (0.094)
Savoie 2	Density	-5.262*** (0.118)	-5.505*** (0.134)	-7.320*** (0.331)	-5.363*** (0.128)
	Absdiff SAS	0.009 (0.086)	0.011 (0.103)	0.179 (0.264)	-0.080 (0.104)
	SAS receiver	0.065 (0.071)	-0.143* (0.086)	-0.753*** (0.228)	0.125 (0.086)
	SAS sender	0.117* (0.071)	0.078 (0.086)	-0.194 (0.229)	-0.204** (0.086)

Table 4: Socioeconomic Homophily, Emission and Reception in Popularity Networks (ERGM) – Wave 6

		Cool	Known	Fear	Liked
Paris 1	Density	-7.063*** (1.147)	-6.744*** (0.912)	-5.656*** (0.411)	-4.576*** (0.287)
	Absdiff SAS	-699.859 (12,497.960)	-743.128 (13,008.960)	-0.397 (0.368)	-0.738*** (0.218)
	SAS receiver	-698.842 (12,497.960)	-742.608 (13,008.950)	-0.519* (0.314)	0.234 (0.168)
	SAS sender	700.362 (12,497.950)	743.714 (13,008.950)	0.795** (0.315)	0.503*** (0.168)
Paris 2	Density	-5.402*** (0.126)	-5.264*** (0.117)	-7.024*** (0.311)	-6.087*** (0.163)
	Absdiff SAS	-0.377*** (0.112)	-0.276*** (0.103)	-0.866** (0.422)	-0.103 (0.130)
	SAS receiver	0.015 (0.089)	-0.090 (0.084)	-0.862** (0.357)	0.081 (0.107)
	SAS sender	0.056 (0.089)	-0.155* (0.084)	0.645* (0.357)	-0.145 (0.107)
Savoie 1	Density	-4.742*** (0.115)	-4.606*** (0.111)	-6.591*** (0.376)	-5.044*** (0.144)
	Absdiff SAS	-0.175** (0.088)	-0.290*** (0.091)	-0.874** (0.425)	-0.426*** (0.125)
	SAS receiver	0.233*** (0.071)	0.086 (0.073)	-1.042*** (0.344)	0.134 (0.098)
	SAS sender	0.047 (0.071)	0.059 (0.073)	-0.071 (0.344)	0.057 (0.098)
Savoie 2	Density	-5.230*** (0.141)	-4.762*** (0.119)	-10.012*** (1.001)	-5.619*** (0.165)
	Absdiff SAS	-0.140 (0.115)	-0.220** (0.108)	0.778 (0.571)	-0.007 (0.129)
	SAS receiver	0.034 (0.093)	-0.286*** (0.088)	-0.363 (0.516)	-0.101 (0.106)
	SAS sender	-0.066 (0.093)	-0.191** (0.088)	0.536 (0.518)	-0.030 (0.106)

***: p<0.01; **: p<0.05; *: p<0.1

Appendix 3D: Additional Details on Socio-Academic Scores

Regression Model used for the Construction of SA Scores: Performance at the *Brevet des Collèges* predicted by the PCS Category of the Two Parents (DEPP nationally representative sample)

Table 1: Performance at the *Brevet* Predicted by Parental Occupation (Linear Regression)

parameter	coefficient	parameter	coefficient
pcsmother10	reference	pcsfather10	reference
pcsmother21	-0.048 (0.231)	pcsfather21	-0.948*** (0.134)
pcsmother22	-0.033 (0.208)	pcsfather22	-0.649*** (0.136)
pcsmother23	-0.156 (0.390)	pcsfather23	-0.127 (0.162)
pcsmother31	1.257*** (0.228)	pcsfather31	0.473*** (0.157)
pcsmother32	1.639*** (0.217)	pcsfather32	1.082*** (0.175)
pcsmother33	1.139*** (0.251)	pcsfather33	0.608*** (0.167)
pcsmother34	1.641*** (0.237)	pcsfather34	0.585*** (0.207)
pcsmother35	0.684*** (0.260)	pcsfather35	0.120 (0.206)
pcsmother37	0.760*** (0.203)	pcsfather37	0.328** (0.130)
pcsmother38	1.076*** (0.240)	pcsfather38	0.526*** (0.132)
pcsmother42	1.281*** (0.201)	pcsfather42	0.148 (0.171)
pcsmother43	0.752*** (0.192)	pcsfather43	-0.290* (0.164)
pcsmother44	-2.571 (1.820)	pcsfather44	1.340 (0.907)
pcsmother45	0.508** (0.213)	pcsfather45	0.102 (0.190)
pcsmother46	0.359* (0.193)	pcsfather46	-0.365*** (0.137)
pcsmother47	0.745*** (0.240)	pcsfather47	-0.367*** (0.135)
pcsmother48	-0.124 (0.305)	pcsfather48	-0.645*** (0.136)
pcsmother52	-0.265 (0.188)	pcsfather52	-0.834*** (0.142)
pcsmother53	-0.152 (0.299)	pcsfather53	-0.544*** (0.144)
pcsmother54	0.058 (0.188)	pcsfather54	-0.448*** (0.171)
pcsmother55	-0.405** (0.194)	pcsfather55	-0.994*** (0.166)
pcsmother56	-0.636*** (0.188)	pcsfather56	-1.534*** (0.188)
pcsmother61	-0.713*** (0.202)	pcsfather61	-1.259*** (0.118)
pcsmother66	-0.953*** (0.196)	pcsfather66	-1.383*** (0.127)
pcsmother69	-0.966*** (0.271)	pcsfather69	-1.683*** (0.197)
pcsmother81	-1.353** (0.549)	pcsfather81	-1.791** (0.782)
pcsmother82	-0.957*** (0.193)	pcsfather82	-2.291*** (0.192)
pcsmother99	-0.671*** (0.192)	pcsfather99	-1.744*** (0.124)
Constant	0.779*** (0.163)	-	-
Observations	30,753		
R^2	0.175		
Adjusted R^2	0.174		

***: p<0.01; **: p<0.05; *: p<0.1

Note: the dependent variable is the coordinate of individuals on the first axis of a Principle Component Analysis (PCA) performed on the 14 different grades obtained at the exam (end-of-the-year exam as well as continuous evaluation). All treatments are done using sampling weights (cf. the codebook of the DEPP panel, Anon 2017).

Management of Ambiguous Declarations for Parental Occupation

It was sometimes the case that a declared occupation could fall in two distinct PCS categories. Typically, I would be unable to tell a craftsman apart from a skilled worker (e.g. the student said the parent was a “carpenter”); or to differentiate between different hierarchical levels for employees (e.g. many parents are sales representatives, which could be either an intermediary or manager occupation depending on one’s exact position in the company’s hierarchy). In such cases, the parent was attributed two categories instead of one; and the SA score of the student was determined by averaging the two possible scores that would result from these two alternatives. This was a way to mitigate the uncertainty associated with these declarations, while still making use of the information available, since the two possible categories were generally closer to one another than they would be to any random occupation (note that this would only be done for a single parent, such that the final score was also impacted by the second parent’s accurate information; if both parents had two possible occupational categories, then the student’s score was simply considered as missing).

Comparison of SA Scores and ISEI Predictive Power on Friendship Relations

Students’ declaration about their parental occupations were coded into the ISCO categories, then turned into ISEI scores (using the ‘ISCO08ConverRsions’ R package). Students’ score is defined as the average of the ISEI of the parents in the household (so for separated parents, only the one living with the student is considered). Students’ ISEI are then standardized across the four schools. In Table 2, both SA scores and ISEI are used as covariates in an ERGM model in the networks of “very good friend” nominations (wave 1). We can see that the homophily coefficient is statistically significant for SA scores, but not for ISEI, suggesting that patterns of socioeconomic homophily are better captured by the SA scores than by the ISEI measure.

Table 2: Socioeconomic Homophily, Activity and Popularity in Very Good Friend Networks for SAS and ISEI (ERGM) – Wave 1

	Paris 1	**Paris 2**	**Savoie 1**	**Savoie 2**
Density	-1.740*** (0.086)	-2.486*** (0.029)	-2.646*** (0.061)	-3.168*** (0.047)
Absdiff SAS	-0.380*** (0.092)	-0.190*** (0.029)	-0.184*** (0.053)	-0.076* (0.045)
SAS receiver	-0.072 (0.107)	0.135*** (0.032)	0.073 (0.060)	0.102** (0.051)
SAS sender	-0.421*** (0.109)	0.167*** (0.032)	-0.011 (0.60)	0.214*** (0.051)
Absdiff ISEI	-0.012 (0.082)	0.007 (0.028)	-0.132** (0.058)	-0.071 (0.044)
ISEI receiver	0.007 (0.098)	-0.159*** (0.031)	-0.129* (0.066)	0.064 (0.049)
ISEI sender	0.372*** (0.099)	-0.236*** (0.031)	-0.132** (0.066)	0.016 (0.050)

Appendix 3E: Coding of Ethnic Categories

Three information were used for ethnic coding: students' last name, first name, and the foreign languages that they declared to speak (questionnaires wave 3; note that they could also indicate their level of fluency in that language, such that simply understanding simple conversations or knowing a few words was enough for a language to be declared). For each of these 3 variables *separately*, I made a qualitative decision as to whether it could be associated to a geographic region of the world. These regions are typically large cultural areas with common languages and/or similar names (e.g. Arabic names define a group of countries in North Africa and the Middle-East). In some cases, one information could be coded as indicating two or three regions at once (notably Muslim or Christian first names, which can refer to many different countries). Subregions were also defined when information were more precise (e.g. Berber dialects in North Africa). When in doubt, no region was associated to the information. Natives (i.e. French students) were not a category, but were simply defined as having no migratory information.

The 12 regions in which each of the three variables could be classified are: (1) Eastern and South-Eastern Asia (2) Indian subcontinent (3) Eastern Europe (4) Romani (5) Italy (6) Southern Europe (except Italy) (7) Northern Europe, Northern America, Australia and New Zealand (8) Latin America (9) Subsaharian Africa (10) Middle-East and North Africa (11) Turkey and (12) French oversea territories.

Note that categories 4 and 5 (Romani and Italy) are more specific than other regions, and they could arguably be subsumed under broader categories. Italy has been kept as a stand-alone category due to its particular importance in certain schools (Savoyard schools tend to have a fair share of Italian backgrounds, probably due to the geographical and historical proximity of Savoie and Italy), whereas Romani backgrounds are not straightforward to associate to a single European region due to their nomadic history (this only concerns a few students anyway).

Also note that some regions could sometimes be difficult to distinguish from one another, in which case a priority rule was applied. Thus, when in doubt, Middle-East and North Africa prevailed over Turkey and Subsaharian Africa (e.g. Muslim first and last name, no language spoken). Similarly, Southern Europe (Spain) was preferred over Latin America. These choices were made based on the relative frequency of the corresponding groups in my sample (students are altogether more likely to be from a MENA or South European background). In any case, a trace of this arbitrary decision was kept in the coding process, so different classification choice would be possible with the same data base.

Finally, in some cases, sub-regions could be determined when the information available was more precise than these 12 large groups. This is generally the case for the languages spoken, many of which are specific to a country and/or ethnic group (e.g. Berber, Portuguese, Polish, Farsi, etc). The information about these sub-regions was kept in the database, but it is not used in any of the treatments presented in the PhD.

Then, information about the three variables were aggregated following these rules (NB: all the examples given are of course fictional to protect students' anonymity):

1. When student last name is associated to a single region, and the other pieces of information are either concordant or missing, the region is attributed.

   E.g. 1: family name is COMPAGNO (Italy), first name is MARIO (Italy) and Italian is declared as a spoken language. The student is coded as Southern Europe, sub-region Italy.

   E.g. 2: family name is BOUKRAA (MENA region), first name is ADAM (could be from different regions, but coherent with MENA) and there are no foreign language declared. The student is coded as Middle-East and North Africa (MENA), no sub-region specified.

2. When student last name is not associated to a single region, but both the first name and languages spoken are, it is attributed.

E.g.: family name is MOUS (unclear), first name is ALI (muslim), language spoken is Arabic. The student is coded as Middle-East and North-Africa (MENA).

3. When only language is associated to a region, a qualitative decision is made: if the language is not taught in the French schooling system, that the student declared a relatively high level of mastery, and that the first and last name could be from this region (though not clearly indicative of it either), I could decide to attribute the region. Otherwise, it was not attributed.

   E.g. 1: family name is FEBRE (France), first name is INES (unclear), language spoken is Arabic with a high level of mastery (fluent). The student is attributed MENA.

   E.g. 2: family name is PIERRE (France), first name is Clara (unclear), language spoken is Italian with a low level of mastery (can understand a basic conversation). The student is not attributed any non-native background (Italian can be taught in French schools).

4. First names alone are not considered sufficient to attribute a region.

An important implication of these rules is that, in many cases, **students from a mixed ascendance – i.e. immigrant and native – are treated as having a migratory background**. For instance, in the example 1 of point 3 I gave above (INES FEBRE, speaks Arabic), this fictional student likely has a French father or grand-father (patronym is native French). However, since natives are not a category that can be attributed, but are simply defined by lack of any other category, this student has been classified as MENA based on her first name and mastery of Arabic. This coding choice is motivated by the fact that **I preferred to have conservative estimates of ethnic homophily in students' friendships**. Indeed, we can reasonably hypothesize that ethnic homophily will generally be stronger among student with a homogeneous background compared to those with a mixed background[5]. Given the uncertainty associated with the coding of ethnic origin through proxy variables, and given the relative political sensitivity of immigration-related topics in France (both within schools and

5 On a distinct but related topic, there has been a recent controversy in France between Jean-Baptiste Mignot on the one hand, and Baptiste Coulmont and Patrick Simon on the other, about the prevalence of French first names among descendants of immigrants (Coulmont and Simon 2021; Mignot 2021). One clear result that came out of this controversy is that descendants of immigrants are much more likely to have a French first name when they come from a mixed rather than homogeneous ascendance (e.g. their mother is a native and father is an immigrant). This seems to suggest that the cultural specificity of descendants of immigrants compared to a national population (whatever that specificity may be and however it may be explained) is generally weaker among those with a mixed background.

at the national level), I deemed it preferable to have conservative rather than inflated estimates.

In any case, it is worth remembering that ethnic homophily is very much not the focus of the present dissertation, and that the variable of ethnic origin is primarily meant to act as a control for the estimation of socioeconomic homophily.

Finally, the 13 origins (including natives) were further aggregated into 5 large groups:

1. Natives: students with no identifiable ethnic background + oversea territories
2. Middle-East and North Africa (MENA) + Turkey
3. Subsaharian Africa
4. Europe outside of France
5. Other origins

Appendix 3F: Cultural Items in the Questionnaires

Items used for the Construction of the Cultural Matching Covariates

Questionnaire Extract (Wave 2): List of Musicians Presented to Students

	J'aime	Je n'aime pas	Je n'ai pas d'avis	Je ne connais pas
Niska				
Dadju				
Marwa Loud				
Kaaris				
Maitre Gims				
PNL				
JUL				
Ninho				
Vald				
Soprano				
BigFlo et Oli				
Orelsan				
Daft Punk				
The Beatles				
ACDC				
Coldplay				
Imagine Dragons				
Mika				
Jean-Jacques Goldman				
Louane				
Ed Sheeran				
Katy Perry				
Ariana Grande				
Drake				
Eminem				

Questionnaire Extract (Wave 4): List of Musicians Presented to Students

	J'aime	Je n'aime pas	Je n'ai pas d'avis	Je ne connais pas
RK				
PNL				
Niska				
Kaaris				
JUL				
Kalash Criminel				
Damso				
Soprano				
BigFlo et Oli				

Orelsan				
Lomepal				
Maitre Gims				
Marwa Loud				
Aya Nakamura				
Eva Queen				
Angèle				
Louane				
Jean-Jacques Goldman				
Johnny Hallyday				
ACDC				
Coldplay				
The Beatles				
Imagine Dragons				
Billie Eilish				
Marshmello				
Martin Garrix				
Daft Punk				
Katy Perry				
Ariana Grande				
Cardi B				
Drake				
Eminem				
XXXTentacion				

Questionnaire Extract (Wave 2): List of Video Makers Presented to Students

	J'aime	Je n'aime pas	Je n'ai pas d'avis	Je ne connais pas
Norman				
Natoo				
Squeezie				
Enjoy Phoenix				
Les Parodie Bro				
Fahd El				
Tibo in Shape				
Ibra TV				
Gamemixtreize				
Le Rire Jaune				
Golden Moustache				
Gotaga				
L'atelier de Roxane				
Troom Troom Fr				
Néo et Swan				
Dr Nozman				
Lolywood				
Ninja				

Questionnaire Extract (Wave 4): List of Video Makers Presented to Students

	J'aime	Je n'aime pas	Je n'ai pas d'avis	Je ne connais pas
Norman				
Squeezie				
Natoo				
Amixem				
Golden Moustache				
Lolywood				
McFly et Carlito				
Fahd El				
Les Parodie Bros				
Le monde à l'envers				
Tibo InShape				
IbraTV				
Lama Faché				
Dr Nozman				
Le grand JD				
Parole de chat				
Néo et Swan				
5 minutes craft				
L'atelier de Roxane				
Troom Troom Fr				
Lena Situations				
Yanissaxoxo				
Michou				
Laink et Terracid				
GameMixTreize				
Gotaga				

Items Used for the Construction of the Cultural Status Covariates

1. You listen to music:

- □ never or almost never
- □ 1, 2 or 3 times a month
- □ 1, 2 or 3 times a week
- □ Every day or almost every day

Name three singers or groups that you like:

2. You read books (apart from those that you have to read for school, and without counting comics or manga):

- □ never or almost never
- □ 1, 2 or 3 times a month
- □ 1, 2 or 3 times a week
- □ Every day or almost every day

What are the last three that you have red (apart from those for school)?

3. When you are not in school, do you do drawing, music, dance or any other artistic activity?

- □ never or almost never
- □ 1, 2 or 3 times a month
- □ 1, 2 or 3 times a week
- □ Every day or almost every day

Which artistic activity(ies)?

You do this activity:

- □ in a club or an association
- □ at home

- ◻ at some friends' place

Construction of the Cultural Status Covariate

The four posts of book-reading frequency where associated a score from 1 (never reading) to 5 (almost every day). There is no value for 2 ("1, 2 or 3 times a month" is worth a 3 and "1,2 and 3 times a week is worth a 4")[6].

The three declared musicians and the declared artistic activities were coded into several genres or categories (30 for musicians and 17 for artistic activities). Then categories were given a score from 1 to 5 as well, following the ordering shown in the two next pages[7]. Since students could declare up to three musicians, and as many artistic activities as they wished, they were always attributed the score of their highest category.

Therefore, each student had three distinct scores ranging from 1 to 5, one for book-reading, one for musical tastes and one for artistic activities (the Pearson correlation between two of these scores is typically of 0.3). The scores of cultural status are defined as the average of these three scores.

The scores were computed twice, in waves 1 and 3. Scores from wave 1 are also used in wave 2 treatments, and scores from wave 3 in all treatments up to wave 6. When a student's answer were missing on either wave 1 or 3, she was attributed her score at the other wave.

6 The "space" in-between the categories (in this case, the fact that reading a few times a month is considered to be closer from reading a few times a week than from never reading) was determined inductively, by tabulating the different cultural items together. In other words, students who never read had musical tastes and artistic practices more distinct from their peers than those from the other 3 categories of book-reading, thus the "skip" from 1 to 3 of the score.

7 Note that there is no value "4" for musicians, with the score jumping fro 3 to 5. The reason is the same as that mentioned for book-reading in the previous note.

Music

28. Soul, Blues, Funk (alicia keys, gregory porter, Bruno Mars)

29. jazz

30. Classical Music

26. Rock/Electro-pop US (imagine dragons, twenty-one pilot, Lindsey stirling, Sam Smith, the chainsmokers, coldplay, Allan Walker, marshmallow, Bastille)

27. Metal and Hard Rock

Score = 5

24. French Rock or Pop-Rock (Camille, M, LEJ)

25. old rock (queen, stones, ac/dc, beatles) + reggae (bob, alpha blondy)

21. Slam in French (grand corps malade)

22. Musicals

23. Old French Songs (Brel, Gainsbourg, brassens)

---

19. Electro / remixs

20. French Engaged Songs (tryo, 3 cafes gourmands, cowboys fringants, arcadian, debout sur le zinc)

17. French Variety or Pop (Angele, Frero Delavega, Vianney, Calogero, Lenni-Kim, Louanne)

18. Pop US (demi lovato, adele, ariana grande, katy perry, ava max, one direction, justin bieber, shawn mendes, ed sheeran)

Score = 3

15. French 'Non-Gangsta' Rap (soprano, biglo et oli, orelsan)

16. French 'Melodic' Rap (Columbine, Lomepal, Roméo Elvis, stupeflip)

---

12. Traditional Music from Home Country

13. French Hits from the 1980s' (demons de minuit, indochine, jj goldman, celine dion, johnny, cloclo, balavoine)

14. US hits from the 1980s' (mickael jacson, toto, abba, a-ha)

9. French humorous songs and parody (Max Boubil, Cyprien)

10. Musics from Disney Movies

11. Anime Openings / Video Games Musics

Score = 2

6. French 'Underground' Rap (nekfeu, damso, fianso)

7. Old French Rap (NTM, IAM)

8. Kpop

3. Rap US (xxxtentacion, lil wayne, cardi B, lucidious, NF)

4. Pop/RnB US (rihanna, lady gaga, anne marie, camila cabello, beyonce, dj snake)

5. Pop Latino (despacito & cie, anitta)

---

1. French 'Gangsta' Rap (Kaaris, Koba laD, Ninho, Zola)

2. French 'Pop/Joyful' Rap (aya nakamura, black m, maitre gims, dadju)

Score = 1

Artistic Activities

16. Writing - Poetry, Fictions		17. Dance - Conservatoire	**Score = 5**
14. Music - Unknown Place of Learning - Rare instrument (other than 11.)		15. Music - Conservatoire (any instrument)	
11. Music - Unknown Place of Learning - Common Instrument (piano, guitare)	12. Coding / Computer Languages	13. Writing - Thoughts, Diary	**Score = 4**
8. Music - Self-Taught, Home Teacher or Musical School	9. Drawing and Plastical Arts - in a Club or Class	10. Theatre	
5. Plastical Arts - at Home - Painting, Hand-made Creations, Grafitti	6. Photographs / Movies	7. Danse - Gymnatics - School or Club	**Score = 3**
2. Drawing - at Home - Manga, Pixel Arts, Pencil Drawing	3. Music - Singing at Home	4. Cooking	**Score = 2**
1. No artistic Activity			**Score = 1**

Appendices for Chapter 4

Appendix 4A: Complete Permutation Tests for Socioeconomic Homophily

Table 1: Average SAS Absolute difference in Friendship Ties – 5-Friend Networks

	Wave 1			**Wave 2**		
School	Observed	Mean in Permutations	P-Value	Observed	Mean in Permutations	P-Value
Paris 1	0.99	1.24	0.00	1.01	1.26	0.00
Paris 2	1.00	1.09	0.00	1.02	1.08	0.01
Savoie 1	1.02	1.13	0.00	1.00	1.12	0.00
Savoie 2	1.07	1.08	0.40	1.04	1.08	0.09
	Wave 3			**Wave 4**		
Paris 1	0.90	1.27	0.00	0.94	1.28	0.00
Paris 2	1.02	1.08	0.01	1.02	1.12	0.00
Savoie 1	0.95	1.10	0.00	1.01	1.11	0.01
Savoie 2	1.06	1.08	0.21	1.04	1.08	0.11
	Wave 5 (missing)			**Wave 6**		
Paris 1	-	-	-	0.88	1.29	0.00
Paris 2	-	-	-	1.05	1.12	0.02
Savoie 1	-	-	-	0.97	1.11	0.00
Savoie 2	-	-	-	1.11	1.09	0.66

Table 2: Average SAS Absolute difference in Friendship Ties – Very Good Friend Networks

	Wave 1			Wave 2		
School	Observed	Mean in Permutations	P-Value	Observed	Mean in Permutations	P-Value
Paris 1	0.99	1.24	0.00	0.97	1.27	0.00
Paris 2	0.98	1.09	0.00	0.97	1.08	0.00
Savoie 1	0.98	1.13	0.00	1.02	1.12	0.00
Savoie 2	1.05	1.08	0.22	1.13	1.08	0.92
	Wave 3			**Wave 4**		
Paris 1	0.93	1.27	0.00	0.93	1.28	0.00
Paris 2	0.99	1.08	0.00	0.99	1.12	0.00
Savoie 1	0.97	1.11	0.00	1.02	1.11	0.01
Savoie 2	1.12	1.09	0.87	1.06	1.08	0.19
	Wave 5 (missing)			**Wave 6**		
Paris 1	-	-	-	0.86	1.29	0.00
Paris 2	-	-	-	0.99	1.12	0.00
Savoie 1	-	-	-	0.98	1.12	0.00
Savoie 2	-	-	-	1.06	1.09	0.16

Table 3: Average SAS Absolute difference in Friendship Ties – Friend Networks

	Wave 1			Wave 2		
School	Observed	Mean in Permutations	P-Value	Observed	Mean in Permutations	P-Value
Paris 1	1.10	1.24	0.00	1.05	1.26	0.00
Paris 2	1.02	1.09	0.00	1.01	1.08	0.00
Savoie 1	1.05	1.13	0.00	1.05	1.12	0.00
Savoie 2	1.12	1.08	0.98	1.15	1.08	1.00
	Wave 3			**Wave 4**		
Paris 1	1.05	1.27	0.00	1.07	1.28	0.00
Paris 2	1.01	1.08	0.00	1.02	1.12	0.00
Savoie 1	1.04	1.11	0.00	1.04	1.11	0.00
Savoie 2	1.15	1.08	1.00	1.10	1.08	0.70
	Wave 5 (missing)			**Wave 6**		

Paris 1	-	-	-	0.94	1.29	0.00
Paris 2	-	-	-	1.04	1.12	0.00
Savoie 1	-	-	-	1.01	1.12	0.00
Savoie 2	-	-	-	1.07	1.09	0.17

Table 4: Average SAS Absolute difference in Friendship Ties – Out-of-School Friend Networks

		Wave 1			**Wave 3**	
School	Observed	Mean in Permutations	P-Value	Observed	Mean in Permutations	P-Value
Paris 1	1.02	1.24	0.00	0.87	1.27	0.00
Paris 2	0.92	1.09	0.00	1.02	1.08	0.07
Savoie 1	0.89	1.12	0.00	0.90	1.10	0.00
Savoie 2	1.04	1.08	0.20	1.10	1.08	0.63

Table 5: Average SAS Absolute difference in Friendship Ties – Parental Friend Networks

		Wave 1			**Wave 3**	
School	Observed	Mean in Permutations	P-Value	Observed	Mean in Permutations	P-Value
Paris 1	0.95	1.24	0.00	0.84	1.27	0.00
Paris 2	0.89	1.09	0.00	0.92	1.08	0.00
Savoie 1	0.88	1.12	0.00	0.91	1.11	0.00
Savoie 2	1.05	1.08	0.26	1.06	1.08	0.33

Appendix 4B: Exposure to Socioeconomically Distant Others, Per School (Waves 1, 4 and 6)

Table 1: Paris 1 (wave 1 / wave 4 / wave 6)

		Nb of Emitted Nominations toward Students from the Opposite Quartile			
		0 friend	1 friend	2 friends	3+ friends
Friends	students from lowest quartile (%)	0 / 5 / 5	0 / 15 / 32	11 / 25 / 21	89 / 55 / 42
	students from highest quartile (%)	0 / 0 / 30	16 / 15 / 15	32 / 20 / 20	52 / 65 / 35
Very Good Friends	students from lowest quartile (%)	16 / 40 / 37	47 / 40 / 42	21 / 15 / 16	16 / 15 / 5
	students from highest quartile (%)	58 / 70 / 65	16 / 5 / 5	0 / 5 / 20	26 / 20 / 10

Table 2: Paris 2 (wave 1 / wave 4 / wave 6)

		Nb of Emitted Nominations toward Students from the Opposite Quartile			
		0 friend	1 friend	2 friends	3+ friends
Friends	students from lowest quartile (%)	0 / 0 / 0	0 / 0 / 0	2 / 0 / 2	98 / 100 / 98
	students from highest quartile (%)	0 / 0 / 0	0 / 0 / 0	0 / 3 / 2	100 / 97 / 98
Very Good Friends	students from lowest quartile (%)	5 / 7 / 12	8 / 15 / 17	14 / 12 / 15	73 / 66 / 56
	students from highest quartile (%)	2 / 0 / 5	17 / 3 / 16	21 / 15 / 24	60 / 82 / 55

Table 3: Savoie 1 (wave 1 / wave 4 / wave 6)

		Nb of Emitted Nominations toward Students from the

		Opposite Quartile			
		0 friend	1 friend	2 friends	3+ friends
Friends	students from lowest quartile (%)	8 / 13 / 13	0 / 7 / 5	3 / 4 / 3	89 / 76 / 79
	students from highest quartile (%)	2 / 0 / 7	5 / 7 / 2	7 / 4 / 7	86 / 89 / 84
Very Good Friends	students from lowest quartile (%)	21 / 24 / 33	26 / 20 / 13	21 / 20 / 23	32 / 36 / 31
	students from highest quartile (%)	32 / 15 / 20	15 / 26 / 30	12 / 15 / 27	41 / 44 / 43

Table 4: Savoie 2 (wave 1 / wave 4 / wave 6)

		Nb of Emitted Nominations toward Students from the Opposite Quartile			
		0 friend	1 friend	2 friends	3+ friends
Friends	students from lowest quartile (%)	2 / 0 / 5	4 / 2 / 2	2 / 2 / 5	92 / 96 / 88
	students from highest quartile (%)	0 / 0 / 0	0 / 2 / 11	0 / 0 / 7	0 / 98 / 82
Very Good Friends	students from lowest quartile (%)	12 / 12 / 14	18 / 15 / 36	14 / 23 / 23	56 / 50 / 27
	students from highest quartile (%)	5 / 6 / 20	26 / 17 / 18	19 / 23 / 23	50 / 54 / 39

Appendix 4C: Complete ERGM Outputs for Chapter 4

Table 1: ERGMs Used for the Construction of Figure 4-3 – Very Good Friend Networks

		Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 6**
Paris 1	Density	-1.803*** (0.073)	-1.561*** (0.063)	-1.598*** (0.069)	-1.665*** (0.072)	-1.703*** (0.075)
	Absdiff SAS	-0.389*** (0.059)	-0.463*** (0.051)	-0.543*** (0.058)	-0.586*** (0.058)	-0.666*** (0.067)
	SAS receiver	-0.071 (0.047)	-0.051 (0.040)	0.009 (0.045)	0.088* (0.045)	0.008 (0.052)
	SAS sender	-0.068 (0.047)	-0.020 (0.040)	-0.115** (0.045)	0.010 (0.045)	-0.170*** (0.052)
Paris 2	Density	-2.396*** (0.025)	-2.445*** (0.026)	-2.442*** (0.026)	-2.382*** (0.027)	-2.723*** (0.033)
	Absdiff SAS	-0.196*** (0.021)	-0.206*** (0.022)	-0.149*** (0.021)	-0.200*** (0.023)	-0.201*** (0.028)
	SAS receiver	-0.006 (0.017)	0.035* (0.018)	-0.002 (0.018)	-0.007 (0.019)	0.008 (0.023)
	SAS sender	-0.044*** (0.017)	-0.124*** (0.018)	-0.154*** (0.018)	-0.174*** (0.019)	-0.119*** (0.023)
Savoie 1	Density	-2.613*** (0.048)	-2.610*** (0.043)	-2.646*** (0.044)	-2.738*** (0.043)	-2.798*** (0.047)
	Absdiff SAS	-0.272*** (0.040)	-0.192*** (0.034)	-0.266*** (0.036)	-0.182*** (0.034)	-0.275*** (0.039)
	SAS receiver	-0.036 (0.032)	0.081*** (0.027)	0.079*** (0.029)	0.084*** (0.028)	0.079** (0.031)
	SAS sender	-0.122*** (0.032)	-0.033 (0.027)	-0.069** (0.029)	-0.107*** (0.028)	-0.077** (0.031)
Savoie 2	Density	-3.207*** (0.041)	-3.163*** (0.043)	-3.234*** (0.044)	-3.014*** (0.043)	-3.069*** (0.049)
	Absdiff SAS	-0.125*** (0.031)	0.015 (0.031)	0.014 (0.033)	-0.089*** (0.034)	-0.090** (0.038)
	SAS receiver	0.157*** (0.025)	0.104*** (0.026)	0.079*** (0.027)	0.114*** (0.028)	0.056* (0.031)
	SAS sender	0.224*** (0.025)	0.115*** (0.026)	0.037 (0.027)	-0.045 (0.028)	-0.060* (0.031)

Table 2: ERGMs Used for the Construction of Figure 4-3 – Friend Networks

		Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 6**
Paris 1	Density	-0.559*** (0.049)	-0.473*** (0.046)	-0.334*** (0.048)	-0.473*** (0.049)	-0.426*** (0.051)
	Absdiff SAS	-0.242*** (0.035)	-0.387*** (0.034)	-0.429*** (0.035)	-0.405*** (0.035)	-0.622*** (0.040)
	SAS receiver	-0.036 (0.029)	-0.019 (0.027)	0.040 (0.027)	0.029 (0.028)	-0.039 (0.031)
	SAS sender	-0.161*** (0.029)	-0.096*** (0.027)	-0.090*** (0.027)	-0.021 (0.027)	-0.076** (0.031)
Paris 2	Density	-1.579*** (0.018)	-1.560*** (0.018)	-1.384*** (0.017)	-1.380*** (0.019)	-1.673*** (0.021)
	Absdiff SAS	-0.142*** (0.014)	-0.144*** (0.015)	-0.153*** (0.014)	-0.177*** (0.015)	-0.148*** (0.017)
	SAS receiver	0.026** (0.012)	0.010 (0.012)	0.024** (0.011)	-0.020 (0.012)	0.005 (0.014)
	SAS sender	-0.006 (0.012)	-0.046*** (0.012)	-0.062*** (0.011)	-0.098*** (0.012)	-0.036*** (0.014)
Savoie 1	Density	-1.367*** (0.029)	-1.345*** (0.026)	-1.427*** (0.026)	-1.515*** (0.026)	-1.585*** (0.028)
	Absdiff SAS	-0.167*** (0.022)	-0.156*** (0.020)	-0.157*** (0.020)	-0.155*** (0.020)	-0.229*** (0.022)
	SAS receiver	-0.048*** (0.018)	0.064*** (0.016)	0.060*** (0.016)	0.097*** (0.016)	0.059*** (0.018)
	SAS sender	-0.085*** (0.018)	0.007 (0.016)	-0.005 (0.016)	0.036** (0.016)	-0.035* (0.018)
Savoie 2	Density	-2.000*** (0.024)	-1.948*** (0.025)	-1.946*** (0.025)	-1.776*** (0.025)	-1.850*** (0.029)
	Absdiff SAS	0.011 (0.017)	0.078*** (0.018)	0.073*** (0.018)	-0.028 (0.019)	-0.084*** (0.022)
	SAS receiver	0.133*** (0.014)	0.081*** (0.015)	0.070*** (0.015)	0.064*** (0.016)	0.044** (0.018)
	SAS sender	0.163*** (0.014)	0.087*** (0.015)	0.041*** (0.015)	0.074*** (0.016)	0.061*** (0.018)

Table 3: ERGMs Used for the Construction of Figure 4-5 – Dislike Networks

		Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
Paris 1	Density	-2.170*** (0.081)	-2.743*** (0.091)	-3.003*** (0.103)	-3.038*** (0.113)	-3.274*** (0.116)
	Absdiff SAS	-0.257*** (0.061)	0.022 (0.064)	0.097 (0.067)	-0.125* (0.074)	0.131* (0.068)
	SAS receiver	0.008 (0.049)	-0.283*** (0.053)	-0.131** (0.055)	-0.071 (0.060)	0.048 (0.056)
	SAS sender	-0.056 (0.049)	-0.077 (0.053)	-0.035 (0.055)	0.191*** (0.060)	0.053 (0.056)
Paris 2	Density	-3.318*** (0.036)	-3.668*** (0.044)	-3.654*** (0.042)	-3.771*** (0.048)	-3.820*** (0.054)
	Absdiff SAS	-0.082*** (0.028)	-0.130*** (0.033)	-0.026 (0.031)	-0.031 (0.035)	-0.126*** (0.042)
	SAS receiver	0.065*** (0.023)	0.163*** (0.027)	0.044* (0.026)	0.119*** (0.029)	0.023 (0.035)
	SAS sender	-0.015 (0.023)	0.192*** (0.027)	0.188*** (0.026)	0.015 (0.029)	-0.004 (0.035)
Savoie 1	Density	-2.921*** (0.051)	-3.075*** (0.051)	-3.037*** (0.049)	-2.996*** (0.046)	-3.239*** (0.054)
	Absdiff SAS	-0.020 (0.037)	-0.034 (0.038)	-0.086** (0.038)	-0.019 (0.035)	-0.063 (0.041)
	SAS receiver	-0.119*** (0.031)	-0.150*** (0.031)	-0.087*** (0.031)	-0.215*** (0.029)	-0.101*** (0.033)
	SAS sender	0.072** (0.031)	0.012 (0.031)	0.047 (0.031)	0.131*** (0.029)	0.047 (0.033)
Savoie 2	Density	-2.972*** (0.036)	-3.023*** (0.040)	-3.140*** (0.041)	-3.110*** (0.045)	-3.477*** (0.058)
	Absdiff SAS	0.043 (0.028)	0.071** (0.030)	0.117*** (0.031)	-0.071* (0.038)	0.040 (0.045)
	SAS receiver	-0.127*** (0.023)	-0.112*** (0.025)	-0.139*** (0.026)	-0.153*** (0.031)	-0.124*** (0.037)
	SAS sender	0.040* (0.023)	-0.038 (0.025)	-0.066** (0.026)	-0.103*** (0.031)	-0.113*** (0.037)

Table 4: ERGMs Used for the Construction of Figure 4-5 – Bullying Networks

		Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
Paris 1	Density	-4.780*** (0.256)	-4.876*** (0.279)	-4.703*** (0.227)	-5.330*** (0.353)	-6.064*** (0.449)
	Absdiff SAS	-0.099 (0.180)	-0.644* (0.339)	0.043 (0.148)	-0.260 (0.241)	0.119 (0.208)
	SAS receiver	-0.172 (0.147)	-0.863*** (0.294)	-0.282** (0.124)	-0.108 (0.195)	0.277 (0.175)
	SAS sender	0.272* (0.147)	0.767*** (0.295)	0.430*** (0.124)	0.486** (0.195)	0.635*** (0.174)
Paris 2	Density	-5.577*** (0.113)	-6.387*** (0.176)	-6.164*** (0.148)	-7.740*** (0.316)	-7.140*** (0.266)
	Absdiff SAS	-0.208** (0.097)	-0.324** (0.151)	-0.076 (0.118)	0.188 (0.208)	-0.032 (0.199)
	SAS receiver	-0.132* (0.079)	-0.013 (0.121)	-0.075 (0.097)	0.039 (0.176)	-0.059 (0.164)
	SAS sender	-0.002 (0.079)	0.345*** (0.121)	0.096 (0.097)	0.089 (0.176)	0.158 (0.165)
Savoie 1	Density	-5.644*** (0.209)	-6.207*** (0.245)	-5.849*** (0.218)	-6.411*** (0.260)	-6.430*** (0.303)
	Absdiff SAS	-0.208 (0.177)	-0.113 (0.202)	-0.490** (0.209)	-0.221 (0.224)	-0.616** (0.309)
	SAS receiver	-0.379*** (0.143)	-0.508*** (0.166)	-0.321* (0.166)	-0.391** (0.182)	0.156 (0.244)
	SAS sender	-0.265* (0.143)	-0.119 (0.166)	0.190 (0.167)	0.059 (0.183)	-0.382 (0.244)
Savoie 2	Density	-5.973*** (0.158)	-6.211*** (0.182)	-5.992*** (0.186)	-6.757*** (0.291)	-7.429*** (0.366)
	Absdiff SAS	0.035 (0.127)	0.176 (0.135)	-0.228 (0.168)	-0.127 (0.261)	0.396 (0.259)
	SAS receiver	-0.290*** (0.106)	-0.032 (0.115)	-0.005 (0.137)	-0.289 (0.217)	-0.240 (0.225)
	SAS sender	-0.128 (0.106)	-0.265** (0.115)	-0.306** (0.137)	-0.430** (0.216)	-0.319 (0.225)

Table 5a: ERGMs Used for the Construction of Figure 4-6 – Very Good Friend Networks, Paris 1

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-3.523*** (0.193)	-3.016*** (0.176)	-3.016*** (0.193)	-3.476*** (0.198)	-3.352*** (0.221)
absdiff.sas	-0.313*** (0.063)	-0.362*** (0.055)	-0.468*** (0.061)	-0.532*** (0.062)	-0.572*** (0.073)
nodeicov.sas	0.165*** (0.061)	0.078 (0.053)	0.154*** (0.057)	0.138** (0.058)	0.116 (0.074)
nodeocov.sas	0.275*** (0.061)	0.253*** (0.053)	0.093 (0.057)	0.194*** (0.059)	-0.045 (0.075)
nodematch.ethn	0.711*** (0.110)	0.630*** (0.098)	0.709*** (0.106)	0.770*** (0.110)	0.744*** (0.119)
nodeifactor.ethn5.europe	0.276 (0.204)	0.328* (0.174)	0.370** (0.176)	0.406** (0.171)	0.330* (0.192)
nodeifactor.ethn5.mena	0.406*** (0.132)	0.387*** (0.119)	0.409*** (0.131)	0.154 (0.134)	0.340** (0.160)
nodeifactor.ethn5.other	-0.187 (0.352)	-0.137 (0.293)	0.280 (0.281)	-0.142 (0.316)	-0.042 (0.300)
nodeifactor.ethn5.subsah	0.633*** (0.154)	0.797*** (0.136)	0.688*** (0.153)	0.635*** (0.150)	0.805*** (0.168)
nodeofactor.ethn5.europe	-0.606** (0.266)	-0.211 (0.193)	-0.122 (0.193)	0.637*** (0.175)	0.463** (0.192)
nodeofactor.ethn5.mena	0.602*** (0.129)	0.223* (0.117)	0.045 (0.129)	0.247* (0.139)	-0.015 (0.162)
nodeofactor.ethn5.other	-0.589 (0.436)	0.201 (0.267)	-0.231 (0.314)	0.814*** (0.266)	0.173 (0.276)
nodeofactor.ethn5.subsah	0.620*** (0.154)	0.764*** (0.132)	0.548*** (0.145)	1.034*** (0.152)	0.789*** (0.163)
absdiff.grade	-0.401*** (0.067)	-0.391*** (0.062)	-0.480*** (0.068)	-0.397*** (0.068)	-0.385*** (0.089)
nodeicov.grade	-0.335*** (0.061)	-0.015 (0.058)	-0.061 (0.060)	0.042 (0.061)	0.002 (0.082)
nodeocov.grade	-0.484*** (0.061)	-0.296*** (0.058)	-0.330*** (0.060)	-0.205*** (0.062)	-0.214*** (0.081)
nodematch.sex	1.545*** (0.115)	1.136*** (0.092)	1.352*** (0.104)	1.465*** (0.110)	1.438*** (0.121)
nodeifactor.sex.M	0.199* (0.116)	0.044 (0.099)	-0.042 (0.109)	-0.039 (0.112)	-0.015 (0.124)
nodeofactor.sex.M	-0.365*** (0.116)	0.184* (0.100)	0.322*** (0.109)	0.515*** (0.112)	0.169 (0.124)
Akaike Inf. Crit.	*3,010.868*	*3,908.533*	*3,369.205*	*3,270.974*	*2,710.884*
Bayesian Inf. Crit.	*3,137.184*	*4,036.811*	*3,496.515*	*3,398.284*	*2,837.200*

Table 5b: ERGMs Used for the Construction of Figure 4-6 – Very Good Friend Networks, Paris 2

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-3.934*** (0.075)	-4.541*** (0.089)	-4.662*** (0.088)	-4.531*** (0.093)	-4.769*** (0.110)
absdiff.sas	-0.159*** (0.021)	-0.170*** (0.023)	-0.136*** (0.022)	-0.182*** (0.023)	-0.193*** (0.028)
nodeicov.sas	-0.029 (0.018)	0.031 (0.019)	-0.009 (0.019)	-0.010 (0.020)	0.003 (0.024)
nodeocov.sas	-0.013 (0.018)	-0.106*** (0.019)	-0.109*** (0.019)	-0.134*** (0.019)	-0.102*** (0.023)
nodematch.ethn	0.339*** (0.040)	0.320*** (0.045)	0.398*** (0.044)	0.410*** (0.047)	0.525*** (0.056)
nodeifactor.ethn5.europe	-0.147* (0.083)	0.093 (0.085)	0.191** (0.083)	0.190** (0.086)	0.109 (0.111)
nodeifactor.ethn5.mena	-0.043 (0.047)	0.124** (0.057)	0.186*** (0.057)	0.071 (0.059)	0.026 (0.072)
nodeifactor.ethn5.other	-0.393*** (0.081)	-0.020 (0.082)	-0.063 (0.084)	-0.184** (0.087)	0.002 (0.100)
nodeifactor.ethn5.subsah	0.359*** (0.054)	0.560*** (0.062)	0.608*** (0.062)	0.495*** (0.066)	0.471*** (0.080)
nodeofactor.ethn5.europe	-0.410*** (0.099)	0.550*** (0.094)	0.639*** (0.091)	0.696*** (0.099)	0.603*** (0.116)
nodeofactor.ethn5.mena	0.167*** (0.049)	0.667*** (0.066)	0.636*** (0.065)	0.768*** (0.071)	0.539*** (0.081)
nodeofactor.ethn5.other	0.160** (0.077)	0.677*** (0.085)	1.061*** (0.079)	0.889*** (0.088)	0.663*** (0.103)
nodeofactor.ethn5.subsah	0.747*** (0.055)	1.174*** (0.071)	1.125*** (0.070)	1.193*** (0.078)	0.850*** (0.090)
absdiff.grade	-0.031 (0.020)	-	-	-	-
nodeicov.grade	-0.016 (0.017)	-	-	-	-
nodeocov.grade	-0.147*** (0.017)	-	-	-	-
nodematch.sex	1.619*** (0.040)	1.610*** (0.041)	1.481*** (0.040)	1.502*** (0.042)	1.485*** (0.052)
nodeifactor.sex.M	-0.098** (0.042)	-0.127*** (0.042)	-0.101** (0.040)	-0.096** (0.042)	-0.020 (0.052)
nodeofactor.sex.M	0.289*** (0.042)	0.184*** (0.042)	0.441*** (0.040)	0.322*** (0.042)	0.440*** (0.052)
Akaike Inf. Crit.	*31,220.870*	*28,032.470*	*29,270.850*	*25,553.720*	*18,438.020*
Bayesian Inf. Crit.	*31,394.400*	*28,177.470*	*29,415.990*	*25,696.760*	*18,579.270*

Table 5c: ERGMs Used for the Construction of Figure 4-6 – Very Good Friend Networks, Savoie 1

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-4.734*** (0.134)	-4.100*** (0.112)	-4.116*** (0.115)	-4.176*** (0.112)	-3.805*** (0.124)
absdiff.sas	-0.150*** (0.043)	-0.099*** (0.036)	-0.155*** (0.038)	-0.078** (0.037)	-0.198*** (0.041)
nodeicov.sas	-0.047 (0.038)	0.032 (0.032)	0.057* (0.034)	0.071** (0.033)	0.078** (0.036)
nodeocov.sas	-0.038 (0.039)	0.013 (0.033)	-0.050 (0.034)	-0.085** (0.033)	-0.039 (0.036)
nodematch.ethn	0.697*** (0.088)	0.556*** (0.079)	0.731*** (0.079)	0.662*** (0.078)	0.530*** (0.088)
nodeifactor.ethn5.europe	0.466*** (0.090)	0.331*** (0.080)	0.577*** (0.082)	0.517*** (0.079)	0.390*** (0.091)
nodeifactor.ethn5.mena	0.437*** (0.108)	0.362*** (0.096)	0.506*** (0.097)	0.675*** (0.091)	0.558*** (0.103)
nodeifactor.ethn5.other	0.440* (0.236)	0.221 (0.185)	0.442** (0.193)	0.169 (0.209)	0.489** (0.199)
nodeifactor.ethn5.subsah	0.921*** (0.185)	0.737*** (0.176)	1.013*** (0.177)	1.007*** (0.173)	0.588*** (0.188)
nodeofactor.ethn5.europe	0.563*** (0.090)	0.348*** (0.080)	0.583*** (0.081)	0.696*** (0.078)	0.549*** (0.089)
nodeofactor.ethn5.mena	0.637*** (0.105)	0.284*** (0.097)	0.346*** (0.098)	0.408*** (0.095)	0.354*** (0.107)
nodeofactor.ethn5.other	0.039 (0.307)	0.594*** (0.179)	0.769*** (0.185)	0.866*** (0.182)	0.545*** (0.210)

nodeofactor.ethn5.subsah	0.588*** (0.210)	-0.229 (0.251)	-0.300 (0.281)	0.028 (0.245)	-0.090 (0.231)
absdiff.grade	-0.219*** (0.043)	-0.081** (0.033)	-0.262*** (0.037)	-0.335*** (0.036)	-0.399*** (0.042)
nodeicov.grade	0.048 (0.040)	0.082*** (0.030)	0.045 (0.034)	0.010 (0.033)	0.018 (0.037)
nodeocov.grade	-0.130*** (0.040)	-0.235*** (0.030)	-0.137*** (0.033)	-0.192*** (0.033)	-0.208*** (0.036)
nodematch.sex	1.853*** (0.086)	1.469*** (0.064)	1.413*** (0.066)	1.575*** (0.067)	1.333*** (0.070)
nodeifactor.sex.M	0.107 (0.087)	0.132** (0.065)	-0.0005 (0.068)	-0.033 (0.068)	0.051 (0.072)
nodeofactor.sex.M	0.320*** (0.087)	0.033 (0.066)	0.121* (0.068)	0.035 (0.068)	-0.088 (0.072)
Akaike Inf. Crit.	*8,516.306*	*11,216.430*	*10,501.080*	*10,893.830*	*9,174.586*
Bayesian Inf. Crit.	*8,669.088*	*11,371.410*	*10,656.520*	*11,049.970*	*9,329.324*

Table 5d: ERGMs Used for the Construction of Figure 4-6 – Very Good Friend Networks, Savoie 2

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-4.117*** (0.129)	-4.307*** (0.119)	-4.432*** (0.126)	-4.095*** (0.118)	-4.247*** (0.132)
absdiff.sas	-0.103*** (0.035)	0.054 (0.033)	0.034 (0.034)	-0.060* (0.035)	-0.056 (0.040)
nodeicov.sas	0.156*** (0.030)	0.034 (0.029)	0.027 (0.031)	0.067** (0.030)	0.039 (0.035)
nodeocov.sas	0.245*** (0.031)	0.086*** (0.029)	0.021 (0.031)	-0.073** (0.031)	-0.057 (0.035)
nodematch.ethn	0.143 (0.107)	0.176* (0.094)	0.166* (0.100)	0.289*** (0.094)	0.408*** (0.101)
nodeifactor.ethn5.europe	0.272*** (0.104)	0.342*** (0.090)	0.211** (0.097)	0.360*** (0.089)	0.433*** (0.097)
nodeifactor.ethn5.mena	0.020 (0.127)	0.011 (0.117)	0.048 (0.123)	0.126 (0.120)	0.246* (0.127)
nodeifactor.ethn5.other	-0.287 (0.230)	0.360* (0.186)	0.253 (0.172)	0.390** (0.167)	-0.330 (0.245)
nodeifactor.ethn5.subsah	0.535*** (0.189)	0.482*** (0.176)	0.541*** (0.184)	0.388* (0.211)	-0.076 (0.284)
nodeofactor.ethn5.europe	0.099 (0.103)	0.447*** (0.090)	0.204** (0.097)	0.297*** (0.090)	0.530*** (0.097)
nodeofactor.ethn5.mena	-0.194 (0.129)	-0.013 (0.118)	-0.065 (0.125)	0.115 (0.117)	0.114 (0.133)
nodeofactor.ethn5.other	-0.430* (0.233)	0.464*** (0.180)	0.300* (0.171)	0.538*** (0.159)	0.377** (0.191)
nodeofactor.ethn5.subsah	0.180 (0.207)	0.401** (0.183)	0.593*** (0.181)	-0.247 (0.283)	0.330 (0.263)
absdiff.grade	-0.177*** (0.035)	-0.198*** (0.034)	-0.214*** (0.035)	-0.186*** (0.035)	-0.146*** (0.039)
nodeicov.grade	0.042 (0.030)	0.160*** (0.029)	0.101*** (0.031)	0.103*** (0.030)	0.047 (0.035)
nodeocov.grade	-0.007 (0.030)	0.020 (0.029)	-0.022 (0.031)	-0.011 (0.030)	-0.080** (0.034)
nodematch.sex	1.368*** (0.063)	1.421*** (0.061)	1.598*** (0.067)	1.334*** (0.062)	1.216*** (0.069)
nodeifactor.sex.M	0.146** (0.064)	0.179*** (0.063)	0.232*** (0.070)	0.056 (0.063)	-0.062 (0.072)
nodeofactor.sex.M	-0.033 (0.064)	0.098 (0.063)	0.087 (0.070)	0.137** (0.063)	0.278*** (0.072)
Akaike Inf. Crit.	12,377.680	13,015.310	12,342.630	12,324.010	9,851.410
Bayesian Inf. Crit.	12,540.920	13,177.780	12,505.490	12,485.300	10,009.340

Table 6a: ERGMs Used for the Construction of Figure 4-6 – Friend Networks, Paris 1

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-1.453*** (0.119)	-1.195*** (0.120)	-1.127*** (0.121)	-1.764*** (0.123)	-1.765*** (0.136)
absdiff.sas	-0.193*** (0.038)	-0.327*** (0.037)	-0.394*** (0.037)	-0.369*** (0.037)	-0.564*** (0.043)

nodeicov.sas	0.144*** (0.039)	0.046 (0.037)	0.132*** (0.036)	0.102*** (0.037)	0.150*** (0.046)
nodeocov.sas	0.043 (0.038)	0.087** (0.037)	0.077** (0.036)	0.110*** (0.037)	0.100** (0.046)
nodematch.ethn	0.468*** (0.075)	0.543*** (0.074)	0.568*** (0.075)	0.615*** (0.076)	0.585*** (0.082)
nodeifactor.ethn5.europe	-0.060 (0.125)	0.112 (0.118)	0.262** (0.108)	0.306*** (0.115)	0.179 (0.122)
nodeifactor.ethn5.mena	0.241*** (0.081)	0.213*** (0.081)	0.207** (0.083)	0.379*** (0.086)	0.523*** (0.099)
nodeifactor.ethn5.other	-0.149 (0.188)	-0.171 (0.183)	0.064 (0.176)	0.005 (0.186)	0.140 (0.173)
nodeifactor.ethn5.subsah	0.495*** (0.099)	0.530*** (0.097)	0.461*** (0.099)	0.699*** (0.100)	0.796*** (0.108)
nodeofactor.ethn5.europe	-0.093 (0.131)	-0.018 (0.123)	-0.048 (0.114)	0.664*** (0.112)	0.369*** (0.119)
nodeofactor.ethn5.mena	0.360*** (0.081)	-0.116 (0.081)	-0.0002 (0.082)	0.194** (0.086)	0.217** (0.099)
nodeofactor.ethn5.other	-0.030 (0.184)	-0.182 (0.181)	0.187 (0.169)	0.721*** (0.168)	0.389** (0.163)
nodeofactor.ethn5.subsah	0.696*** (0.098)	0.635*** (0.094)	0.484*** (0.097)	0.631*** (0.099)	0.755*** (0.107)
absdiff.grade	-0.280*** (0.042)	-0.292*** (0.041)	-0.215*** (0.039)	-0.219*** (0.041)	-0.199*** (0.053)
nodeicov.grade	-0.264*** (0.040)	0.002 (0.040)	-0.088** (0.038)	0.048 (0.039)	-0.088* (0.052)
nodeocov.grade	-0.278*** (0.040)	-0.280*** (0.041)	-0.254*** (0.038)	-0.206*** (0.039)	-0.205*** (0.052)
nodematch.sex	1.009*** (0.065)	0.960*** (0.061)	0.906*** (0.061)	0.960*** (0.062)	0.817*** (0.068)
nodeifactor.sex.M	-0.035 (0.065)	0.014 (0.066)	-0.074 (0.065)	0.104 (0.064)	0.060 (0.070)
nodeofactor.sex.M	-0.237*** (0.066)	-0.180*** (0.067)	-0.068 (0.065)	0.122* (0.064)	0.079 (0.070)
Akaike Inf. Crit.	*6,220.646*	*6,757.566*	*6,718.369*	*6,520.447*	*5,592.048*
Bayesian Inf. Crit.	*6,346.963*	*6,885.844*	*6,845.679*	*6,647.757*	*5,718.364*

Table 6b: ERGMs Used for the Construction of Figure 4-6 – Friend Networks, Paris 2

	Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 6**
edges	-2.464*** (0.051)	-2.773*** (0.054)	-2.858*** (0.053)	-2.752*** (0.056)	-3.107*** (0.064)
absdiff.sas	-0.114*** (0.015)	-0.116*** (0.015)	-0.137*** (0.014)	-0.152*** (0.015)	-0.134*** (0.017)
nodeicov.sas	0.013 (0.013)	0.008 (0.013)	0.022* (0.012)	-0.026** (0.013)	-0.001 (0.015)
nodeocov.sas	0.008 (0.013)	-0.036*** (0.013)	-0.047*** (0.012)	-0.086*** (0.013)	-0.034** (0.015)
nodematch.ethn	0.255*** (0.029)	0.276*** (0.031)	0.301*** (0.030)	0.339*** (0.032)	0.372*** (0.036)
nodeifactor.ethn5.europe	-0.078 (0.055)	-0.010 (0.056)	0.136** (0.053)	0.161*** (0.057)	0.210*** (0.067)
nodeifactor.ethn5.mena	-0.079** (0.033)	-0.021 (0.037)	0.108*** (0.036)	0.049 (0.039)	0.037 (0.045)
nodeifactor.ethn5.other	-0.238*** (0.053)	-0.020 (0.053)	-0.005 (0.052)	-0.196*** (0.056)	-0.036 (0.062)
nodeifactor.ethn5.subsah	0.265*** (0.039)	0.368*** (0.042)	0.486*** (0.040)	0.472*** (0.044)	0.450*** (0.050)
nodeofactor.ethn5.europe	-0.143** (0.060)	0.287*** (0.060)	0.307*** (0.057)	0.389*** (0.061)	0.301*** (0.071)
nodeofactor.ethn5.mena	0.183*** (0.034)	0.429*** (0.040)	0.511*** (0.038)	0.493*** (0.041)	0.424*** (0.046)
nodeofactor.ethn5.other	-0.018 (0.054)	0.274*** (0.056)	0.398*** (0.053)	0.163*** (0.058)	0.201*** (0.064)
nodeofactor.ethn5.subsah	0.516*** (0.040)	0.747*** (0.045)	0.727*** (0.043)	0.814*** (0.047)	0.589*** (0.054)
absdiff.grade	-0.023 (0.014)	-	-	-	-
nodeicov.grade	-0.011 (0.012)	-	-	-	-
nodeocov.grade	-0.080*** (0.012)	-	-	-	-
nodematch.sex	1.098*** (0.024)	1.157*** (0.024)	1.136*** (0.023)	1.086*** (0.025)	1.078*** (0.029)
nodeifactor.sex.M	-0.019 (0.025)	-0.067*** (0.025)	-0.044* (0.023)	-0.025 (0.025)	-0.012 (0.029)

nodeofactor.sex.M	-0.014 (0.025)	-0.041* (0.025)	0.179*** (0.023)	0.117*** (0.025)	0.345*** (0.029)
Akaike Inf. Crit.	*55,120.200*	*51,220.670*	*55,988.270*	*47,968.540*	*38,084.120*
Bayesian Inf. Crit.	*55,293.720*	*51,365.680*	*56,133.400*	*48,111.580*	*38,225.360*

Table 6c: ERGMs Used for the Construction of Figure 4-6 – Friend Networks, Savoie 1

	Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 6**
edges	-2.582*** (0.075)	-2.364*** (0.068)	-2.441*** (0.069)	-2.430*** (0.067)	-2.218*** (0.075)
absdiff.sas	-0.082*** (0.024)	-0.080*** (0.021)	-0.048** (0.022)	-0.070*** (0.022)	-0.140*** (0.024)
nodeicov.sas	-0.022 (0.022)	0.016 (0.019)	0.024 (0.020)	0.087*** (0.020)	0.058*** (0.022)
nodeocov.sas	-0.020 (0.022)	-0.013 (0.019)	0.007 (0.020)	0.088*** (0.020)	-0.004 (0.022)
nodematch.ethn	0.355*** (0.056)	0.487*** (0.051)	0.494*** (0.052)	0.397*** (0.051)	0.400*** (0.057)
nodeifactor.ethn5.europe	0.275*** (0.056)	0.307*** (0.050)	0.314*** (0.052)	0.339*** (0.051)	0.201*** (0.059)
nodeifactor.ethn5.mena	0.438*** (0.064)	0.441*** (0.059)	0.424*** (0.059)	0.501*** (0.058)	0.411*** (0.066)
nodeifactor.ethn5.other	0.232* (0.140)	0.212* (0.113)	0.262** (0.114)	0.258** (0.114)	0.272** (0.125)
nodeifactor.ethn5.subsah	0.524*** (0.124)	0.396*** (0.122)	0.515*** (0.121)	0.651*** (0.118)	0.532*** (0.115)
nodeofactor.ethn5.europe	0.465*** (0.056)	0.520*** (0.050)	0.530*** (0.052)	0.605*** (0.050)	0.292*** (0.057)
nodeofactor.ethn5.mena	0.566*** (0.063)	0.441*** (0.059)	0.224*** (0.060)	0.481*** (0.058)	0.192*** (0.067)
nodeofactor.ethn5.other	0.265* (0.145)	0.365*** (0.113)	0.486*** (0.114)	0.379*** (0.117)	0.114 (0.136)
nodeofactor.ethn5.subsah	-0.453*** (0.165)	-0.337** (0.151)	-0.543*** (0.166)	-0.486*** (0.170)	-0.237* (0.136)
absdiff.grade	-0.165*** (0.024)	-0.115*** (0.020)	-0.211*** (0.021)	-0.190*** (0.021)	-0.310*** (0.024)
nodeicov.grade	-0.040* (0.023)	0.042** (0.019)	0.026 (0.020)	-0.027 (0.020)	-0.033 (0.021)
nodeocov.grade	-0.130*** (0.023)	-0.089*** (0.019)	-0.149*** (0.020)	-0.199*** (0.019)	-0.247*** (0.021)
nodematch.sex	1.180*** (0.039)	1.188*** (0.035)	1.178*** (0.035)	1.090*** (0.035)	1.020*** (0.038)
nodeifactor.sex.M	0.257*** (0.040)	0.174*** (0.035)	0.119*** (0.036)	0.036 (0.035)	0.065* (0.039)
nodeofactor.sex.M	0.044 (0.040)	-0.221*** (0.035)	-0.052 (0.036)	-0.075** (0.035)	-0.070* (0.039)
Akaike Inf. Crit.	*19,996.470*	*24,164.770*	*23,625.660*	*24,160.720*	*20,517.270*
Bayesian Inf. Crit.	*20,149.250*	*24,319.750*	*23,781.100*	*24,316.860*	*20,672.010*

Table 6d: ERGMs Used for the Construction of Figure 4-6 – Friend Networks, Savoie 2

	Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 6**
edges	-2.632*** (0.071)	-2.574*** (0.070)	-2.752*** (0.070)	-2.628*** (0.068)	-2.644*** (0.080)
absdiff.sas	0.022 (0.020)	0.098*** (0.020)	0.075*** (0.019)	-0.023 (0.020)	-0.064*** (0.023)
nodeicov.sas	0.147*** (0.018)	0.033* (0.017)	0.054*** (0.017)	0.040** (0.017)	0.047** (0.020)
nodeocov.sas	0.170*** (0.018)	0.071*** (0.017)	0.021 (0.017)	0.063*** (0.017)	0.052** (0.020)
nodematch.ethn	0.167*** (0.059)	0.057 (0.058)	0.178*** (0.058)	0.171*** (0.056)	0.190*** (0.065)
nodeifactor.ethn5.europe	0.294*** (0.058)	0.218*** (0.056)	0.315*** (0.055)	0.387*** (0.053)	0.382*** (0.062)
nodeifactor.ethn5.mena	0.117* (0.068)	0.116* (0.067)	0.190*** (0.068)	0.343*** (0.066)	0.355*** (0.075)
nodeifactor.ethn5.other	-0.380*** (0.133)	-0.113 (0.125)	0.034 (0.106)	0.185* (0.104)	-0.440*** (0.140)

nodeifactor.ethn5.subsah	0.381*** (0.115)	0.350*** (0.108)	0.414*** (0.110)	0.206 (0.129)	-0.100 (0.160)
nodeofactor.ethn5.europe	0.204*** (0.058)	0.266*** (0.057)	0.240*** (0.056)	0.411*** (0.053)	0.272*** (0.062)
nodeofactor.ethn5.mena	-0.039 (0.070)	0.099 (0.068)	0.085 (0.069)	0.158** (0.068)	0.010 (0.079)
nodeofactor.ethn5.other	-0.078 (0.120)	0.243** (0.113)	0.290*** (0.100)	0.410*** (0.097)	-0.026 (0.120)
nodeofactor.ethn5.subsah	0.390*** (0.110)	0.446*** (0.105)	0.729*** (0.101)	0.293** (0.129)	-0.275 (0.172)
absdiff.grade	-0.152*** (0.019)	-0.154*** (0.019)	-0.147*** (0.019)	-0.119*** (0.020)	-0.067*** (0.022)
nodeicov.grade	0.019 (0.017)	0.117*** (0.017)	0.074*** (0.018)	0.116*** (0.017)	0.038* (0.020)
nodeocov.grade	-0.004 (0.017)	0.031* (0.017)	0.047*** (0.017)	0.013 (0.017)	-0.011 (0.020)
nodematch.sex	0.969*** (0.032)	0.960*** (0.032)	1.099*** (0.032)	0.952*** (0.032)	0.819*** (0.037)
nodeifactor.sex.M	0.156*** (0.033)	0.133*** (0.033)	0.118*** (0.034)	0.095*** (0.033)	0.100*** (0.039)
nodeofactor.sex.M	-0.114*** (0.033)	0.015 (0.033)	0.053 (0.034)	0.169*** (0.033)	0.199*** (0.038)
Akaike Inf. Crit.	*29,504.280*	*29,750.850*	*29,969.710*	*28,690.440*	*22,651.920*
Bayesian Inf. Crit.	*29,667.520*	*29,913.320*	*30,132.570*	*28,851.730*	*22,809.860*

Appendix 4D: Heterophobia on other Attributes (Gender, Ethnicity and Academic Results)

Table 1: ERGM of Disliking Nominations Predicted by Gender, Ethnicity and Academic Results (homophily, emission and reception) – Paris 1

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-2.547^{***} (0.721)	-4.967^{***} (1.034)	-3.833^{***} (0.955)	-2.560^{***} (0.870)	-18.515 (449.234)
nodematch.sex	-0.360^{***} (0.104)	-0.054 (0.111)	0.070 (0.122)	0.049 (0.135)	0.315^{**} (0.144)
nodematch.ethn	0.133 (0.119)	0.233^{*} (0.137)	0.312^{**} (0.149)	0.067 (0.171)	-0.265 (0.187)
absdiff.grade	-0.045 (0.064)	0.369^{***} (0.068)	0.263^{***} (0.072)	-0.087 (0.084)	-0.095 (0.096)
nodeifactor.sex.M	0.151 (0.109)	0.184 (0.126)	0.082 (0.134)	0.184 (0.144)	-0.195 (0.157)
nodeofactor.sex.M	-0.247^{**} (0.108)	0.146 (0.125)	-0.289^{**} (0.135)	-0.535^{***} (0.149)	-0.024 (0.153)
nodeifactor.ethn.auto	0.688 (0.603)	0.436 (0.731)	0.053 (0.610)	0.235 (0.740)	1.011 (1.024)
nodeifactor.ethn.creole	0.247 (0.785)	0.721 (0.768)	-0.051 (0.665)	0.994 (0.767)	1.144 (1.096)
nodeifactor.ethn.east_eu	0.906 (0.677)	-1.250 (1.235)	-0.163 (0.710)	-0.080 (0.854)	-1.177 (1.427)
nodeifactor.ethn.india	-0.808 (1.173)	0.309 (1.022)	-0.816 (1.177)	$-Inf.000^{***}$ (0.000)	$-Inf.000^{***}$ (0.000)
nodeifactor.ethn.latino	0.207 (0.842)	0.668 (0.937)	$-Inf.000^{***}$ (0.000)	-0.504 (1.245)	-0.931 (1.433)
nodeifactor.ethn.mena	0.900 (0.601)	0.939 (0.727)	0.218 (0.607)	0.765 (0.731)	1.097 (1.023)
nodeifactor.ethn.north	1.285^{**} (0.621)	1.048 (0.756)	0.513 (0.647)	1.180 (0.762)	1.795^{*} (1.035)
nodeifactor.ethn.south_eu	0.638 (0.733)	1.602^{**} (0.787)	0.645 (0.698)	0.985 (0.838)	2.260^{**} (1.072)
nodeifactor.ethn.subsah	0.677 (0.609)	0.864 (0.736)	0.091 (0.618)	0.269 (0.747)	0.572 (1.035)
nodeofactor.ethn.auto	-0.526 (0.394)	0.435 (0.731)	0.211 (0.735)	-0.837^{*} (0.456)	14.608 (449.232)
nodeofactor.ethn.creole	$-Inf.000^{***}$ (0.000)	1.376^{*} (0.756)	1.240 (0.757)	0.162 (0.487)	14.661 (449.233)
nodeofactor.ethn.east_eu	-0.772 (0.510)	$-Inf.000^{***}$ (0.000)	-1.347 (1.015)	-1.729^{***} (0.633)	12.892 (449.233)
nodeofactor.ethn.india	-1.339^{*} (0.709)	0.969 (0.847)	-0.421 (1.027)	-1.256^{*} (0.745)	$-Inf.000^{***}$ (0.000)
nodeofactor.ethn.latino	-1.291^{*} (0.708)	1.582^{*} (0.821)	1.189 (0.823)	-0.731 (0.649)	14.039 (449.233)
nodeofactor.ethn.mena	-0.315 (0.392)	0.853 (0.729)	0.557 (0.732)	-1.060^{**} (0.461)	14.341 (449.232)
nodeofactor.ethn.north	-0.971^{**} (0.465)	-0.049 (0.772)	0.559 (0.760)	-1.161^{**} (0.555)	15.034 (449.233)
nodeofactor.ethn.south_eu	$-Inf.000^{***}$ (0.000)	1.069 (0.858)	-0.831 (1.241)	-1.128 (0.839)	12.869 (449.234)
nodeofactor.ethn.subsah	-0.580 (0.404)	0.981 (0.737)	-0.024 (0.749)	-1.418^{***} (0.488)	12.739 (449.233)
nodeicov.grade	-0.091 (0.059)	-0.199^{***} (0.069)	-0.297^{***} (0.074)	-0.231^{***} (0.080)	-0.236^{**} (0.096)
nodeocov.grade	0.055 (0.058)	0.194^{***} (0.070)	-0.026 (0.074)	0.014 (0.078)	-0.416^{***} (0.090)

Table 2: ERGM of Disliking Nominations Predicted by Gender, Ethnicity and Academic Results (homophily, emission and reception) – Paris 2

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-2.547^{***} (0.721)	-4.818^{***} (1.031)	-3.867^{***} (0.952)	-2.692^{***} (0.866)	-18.827 (445.242)
nodematch.sex	-0.360^{***} (0.104)	-0.101 (0.110)	0.048 (0.121)	0.051 (0.135)	0.313^{**} (0.144)
nodematch.ethn	0.133 (0.119)	0.186 (0.136)	0.275^{*} (0.148)	0.077 (0.171)	-0.252 (0.186)
absdiff.grade	-0.045 (0.064)	-	-	-	-
nodeifactor.sex.M	0.151 (0.109)	0.372^{***} (0.115)	0.276^{**} (0.126)	0.258^{*} (0.141)	-0.117 (0.154)
nodeofactor.sex.M	-0.247^{**} (0.108)	0.016 (0.115)	-0.268^{**} (0.126)	-0.541^{***} (0.145)	0.115 (0.151)
nodeifactor.ethn.auto	0.688 (0.603)	0.607 (0.729)	0.142 (0.608)	0.190 (0.739)	0.990 (1.023)
nodeifactor.ethn.creole	0.247 (0.785)	1.150 (0.763)	0.558 (0.654)	1.115 (0.765)	1.103 (1.095)
nodeifactor.ethn.east_eu	0.906 (0.677)	-1.241 (1.234)	-0.098 (0.709)	-0.052 (0.854)	-1.172 (1.426)
nodeifactor.ethn.india	-0.808 (1.173)	0.389 (1.020)	-0.863 (1.175)	$-\text{Inf}.000^{***}$ (0.000)	$-\text{Inf}.000^{***}$ (0.000)
nodeifactor.ethn.latino	0.207 (0.842)	0.815 (0.935)	$-\text{Inf}.000^{***}$ (0.000)	-0.470 (1.244)	-0.785 (1.431)
nodeifactor.ethn.mena	0.900 (0.601)	1.220^{*} (0.725)	0.471 (0.604)	0.866 (0.729)	1.237 (1.020)
nodeifactor.ethn.north	1.285^{**} (0.621)	1.048 (0.748)	0.338 (0.638)	0.947 (0.758)	1.547 (1.029)
nodeifactor.ethn.south_eu	0.638 (0.733)	1.951^{**} (0.783)	1.193^{*} (0.691)	1.147 (0.836)	2.469^{**} (1.066)
nodeifactor.ethn.subsah	0.677 (0.609)	1.166 (0.732)	0.489 (0.613)	0.456 (0.743)	0.737 (1.031)
nodeofactor.ethn.auto	-0.526 (0.394)	0.707 (0.728)	0.345 (0.732)	-0.867^{*} (0.455)	14.618 (445.240)
nodeofactor.ethn.creole	$-\text{Inf}.000^{***}$ (0.000)	1.436^{*} (0.752)	1.426^{*} (0.750)	0.128 (0.484)	14.583 (445.241)
nodeofactor.ethn.east_eu	-0.772 (0.510)	$-\text{Inf}.000^{***}$ (0.000)	-1.310 (1.014)	-1.737^{***} (0.633)	12.881 (445.241)
nodeofactor.ethn.india	-1.339^{*} (0.709)	1.163 (0.843)	-0.272 (1.022)	-1.284^{*} (0.743)	$-\text{Inf}.000^{***}$ (0.000)
nodeofactor.ethn.latino	-1.291^{*} (0.708)	1.487^{*} (0.818)	1.200 (0.819)	-0.737 (0.648)	14.276 (445.241)
nodeofactor.ethn.mena	-0.315 (0.392)	0.891 (0.728)	0.639 (0.730)	-1.090^{**} (0.458)	14.610 (445.240)
nodeofactor.ethn.north	-0.971^{**} (0.465)	0.486 (0.763)	0.745 (0.752)	-1.194^{**} (0.549)	14.622 (445.240)
nodeofactor.ethn.south_eu	$-\text{Inf}.000^{***}$ (0.000)	0.992 (0.854)	-0.683 (1.237)	-1.145 (0.836)	13.267 (445.241)
nodeofactor.ethn.subsah	-0.580 (0.404)	0.905 (0.734)	0.069 (0.745)	-1.452^{***} (0.481)	13.055 (445.240)
nodeicov.grade	-0.091 (0.059)	-	-	-	-
nodeocov.grade	0.055 (0.058)	-	-	-	-

Table 3: ERGM of Disliking Nominations Predicted by Gender, Ethnicity and Academic Results (homophily, emission and reception) – Savoie 1

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-15.869(116.965)	-3.808***(0.864)	-16.627(186.668)	-4.585***(0.889)	-17.290(186.559)
nodematch.sex	0.007(0.060)	0.027(0.067)	0.196***(0.063)	0.378***(0.059)	0.478***(0.074)
nodematch.ethn	-0.120(0.103)	-0.157(0.101)	-0.143(0.099)	-0.263***(0.102)	-0.178(0.128)
absdiff.grade	0.117***(0.034)	0.003(0.039)	0.110***(0.037)	0.129***(0.034)	0.056(0.042)
nodeifactor.sex.M	0.046(0.066)	0.210***(0.073)	-0.128*(0.068)	0.099(0.064)	0.127(0.078)
nodeofactor.sex.M	-0.588***(0.066)	-0.930***(0.072)	-0.589***(0.069)	-0.570***(0.064)	-0.730***(0.078)
nodeifactor.ethn.auto	0.893(0.594)	0.105(0.468)	0.607(0.592)	0.523(0.517)	1.498(1.012)
nodeifactor.ethn.creole	0.149(0.691)	-0.300(0.559)	0.012(0.673)	-0.336(0.626)	0.515(1.107)
nodeifactor.ethn.east_eu	0.651(0.632)	-0.078(0.517)	-0.276(0.665)	0.110(0.564)	1.103(1.050)
nodeifactor.ethn.roma	1.073*(0.624)	0.892*(0.498)	0.684(0.623)	0.477(0.550)	1.773*(1.044)
nodeifactor.ethn.india	1.368**(0.663)	0.297(0.619)	0.559(0.741)	0.481(0.657)	1.082(1.162)
nodeifactor.ethn.italie	0.811(0.599)	0.105(0.474)	0.465(0.598)	0.562(0.521)	1.618(1.016)
nodeifactor.ethn.latino	0.225(0.780)	-0.702(0.624)	0.206(0.673)	0.224(0.594)	1.365(1.059)
nodeifactor.ethn.mena	0.944(0.598)	0.352(0.473)	0.439(0.599)	0.311(0.523)	1.416(1.017)
nodeifactor.ethn.north	1.212*(0.671)	0.687(0.563)	0.503(0.741)	-0.328(0.773)	-0.016(1.420)
nodeifactor.ethn.south_eu	0.529(0.605)	-0.364(0.484)	0.072(0.608)	-0.088(0.532)	0.544(1.039)
nodeifactor.ethn.subsah	0.579(0.721)	-0.723(0.743)	-1.202(1.162)	-0.070(0.717)	0.784(1.080)
nodeifactor.ethn.turk	1.588***(0.608)	0.523(0.493)	0.563(0.614)	0.773(0.533)	1.596(1.027)
nodeofactor.ethn.auto	12.351(116.963)	0.915(0.720)	13.031(186.667)	1.157(0.718)	12.566(186.556)
nodeofactor.ethn.creole	12.039(116.964)	0.249(0.805)	12.474(186.668)	0.344(0.793)	11.648(186.556)
nodeofactor.ethn.east_eu	11.416(116.964)	-0.735(0.877)	13.562(186.667)	0.645(0.758)	13.257(186.556)
nodeofactor.ethn.roma	11.716(116.964)	1.403*(0.739)	12.890(186.667)	1.211(0.738)	12.538(186.556)
nodeofactor.ethn.india	13.276(116.964)	2.020***(0.766)	14.275(186.667)	2.349***(0.755)	13.245(186.556)
nodeofactor.ethn.italie	12.041(116.963)	0.986(0.725)	13.751(186.667)	0.844(0.723)	11.746(186.556)
nodeofactor.ethn.latino	11.839(116.964)	-0.289(0.815)	13.252(186.667)	-0.407(0.847)	10.275(186.557)
nodeofactor.ethn.mena	12.063(116.963)	0.998(0.724)	12.778(186.667)	0.725(0.724)	11.557(186.556)
nodeofactor.ethn.north	10.499(116.968)	0.013(1.007)	-Inf.000***(0.000)	-Inf.000***(0.000)	-Inf.000***(0.000)
nodeofactor.ethn.south_eu	12.096(116.963)	0.382(0.736)	13.550(186.667)	1.208*(0.723)	13.182(186.556)
nodeofactor.ethn.subsah	12.057(116.964)	1.467*(0.801)	13.301(186.668)	-Inf.000***(0.000)	-Inf.000***(0.000)
nodeofactor.ethn.turk	12.231(116.964)	0.725(0.743)	12.622(186.667)	0.487(0.741)	12.648(186.556)
nodeicov.grade	-0.163***(0.033)	-0.247***(0.036)	-0.240***(0.035)	-0.268***(0.031)	-0.217***(0.039)
nodeocov.grade	-0.154***(0.032)	0.0430.037)	-0.076**(0.034)	-0.181***(0.031)	-0.278***(0.039)

Table 4: ERGM of Disliking Nominations Predicted by Gender, Ethnicity and Academic Results (homophily, emission and reception) – Savoie 2

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
edges	-2.926*** (0.275)	-2.874*** (0.279)	-2.984*** (0.273)	-3.205*** (0.306)	-3.125*** (0.341)
nodematch.sex	0.001 (0.047)	0.015 (0.050)	0.093* (0.050)	0.310*** (0.056)	0.728*** (0.073)
nodematch.ethn	0.079 (0.087)	0.086 (0.092)	0.017 (0.092)	0.036 (0.105)	-0.165 (0.128)
absdiff.grade	0.016 (0.029)	0.047 (0.030)	-0.019 (0.030)	-0.018 (0.033)	-0.066 (0.041)
nodeifactor.sex.M	-0.132*** (0.050)	-0.154*** (0.053)	-0.209*** (0.055)	-0.210*** (0.060)	-0.344*** (0.078)
nodeofactor.sex.M	-0.324*** (0.050)	-0.337*** (0.053)	-0.617*** (0.056)	-0.144** (0.060)	-0.480*** (0.078)
nodeifactor.ethn.auto	0.211 (0.212)	0.538** (0.235)	0.165 (0.204)	0.651** (0.267)	0.589* (0.302)
nodeifactor.ethn.creole	-0.994 (0.616)	0.354 (0.410)	0.638** (0.315)	0.293 (0.463)	0.848* (0.435)
nodeifactor.ethn.east_eu	0.308 (0.251)	0.930*** (0.274)	0.103 (0.254)	0.481 (0.320)	0.914** (0.366)
nodeifactor.ethn.italie	0.081 (0.228)	0.585** (0.250)	0.051 (0.223)	0.607** (0.287)	0.451 (0.321)
nodeifactor.ethn.latino	0.022 (0.370)	-0.020 (0.382)	0.119 (0.302)	0.579 (0.354)	0.092 (0.401)
nodeifactor.ethn.mena	0.407* (0.218)	0.924*** (0.243)	0.382* (0.217)	0.926*** (0.275)	0.819*** (0.312)
nodeifactor.ethn.north	0.787*** (0.229)	0.862*** (0.258)	0.555** (0.227)	0.741** (0.303)	0.775** (0.361)
nodeifactor.ethn.south_eu	-0.045 (0.238)	0.460* (0.256)	-0.105 (0.232)	0.380 (0.284)	0.761** (0.310)
nodeifactor.ethn.subsah	0.499* (0.268)	0.956*** (0.291)	0.896*** (0.253)	1.194*** (0.375)	1.119*** (0.400)
nodeifactor.ethn.turk	0.327 (0.229)	0.227 (0.263)	-0.080 (0.239)	0.144 (0.305)	0.275 (0.339)
nodeofactor.ethn.auto	-0.235 (0.190)	-0.741*** (0.161)	-0.065 (0.197)	-0.772*** (0.173)	-0.878*** (0.187)
nodeofactor.ethn.creole	0.268 (0.317)	-1.252*** (0.440)	-0.614 (0.427)	-Inf.000*** (0.000)	-1.352*** (0.481)
nodeofactor.ethn.east_eu	0.034 (0.230)	-0.713*** (0.238)	0.057 (0.249)	-0.691*** (0.232)	-1.065*** (0.326)
nodeofactor.ethn.italie	0.111 (0.199)	-0.302* (0.171)	0.350* (0.205)	-0.559*** (0.193)	-0.660*** (0.202)
nodeofactor.ethn.latino	0.418 (0.270)	-0.231 (0.262)	1.076*** (0.244)	0.334 (0.216)	-1.526*** (0.344)
nodeofactor.ethn.mena	0.072 (0.198)	-0.780*** (0.183)	0.373* (0.206)	-0.207 (0.178)	-1.097*** (0.221)
nodeofactor.ethn.north	-1.269*** (0.303)	-1.359*** (0.254)	0.338 (0.228)	-1.187*** (0.262)	-1.075*** (0.311)
nodeofactor.ethn.south_eu	0.349* (0.198)	0.130 (0.166)	0.354* (0.206)	-0.850*** (0.191)	-0.688*** (0.197)
nodeofactor.ethn.subsah	-0.037 (0.248)	-0.022 (0.208)	0.828*** (0.232)	-1.047*** (0.395)	-2.168*** (0.608)
nodeofactor.ethn.turk	-0.850*** (0.249)	-2.268*** (0.314)	-1.067*** (0.279)	-2.283*** (0.338)	-1.879*** (0.303)
nodeicov.grade	-0.281*** (0.025)	-0.348*** (0.026)	-0.333*** (0.026)	-0.353*** (0.029)	-0.253*** (0.037)
nodeocov.grade	0.023 (0.025)	-0.116*** (0.026)	-0.207*** (0.026)	-0.195*** (0.029)	-0.173*** (0.037)

Appendix 4E: Probability Curves of Ties Depending on Socioeconomic Background

Figure 1: Probability of Receiving a nomination as “Very Good Friend” Depending on the Socioeconomic Background of the Sender and Receiver (ERGM) – Wave 6

Note: predictions derived from the ERGM of Table 1, Appendix 4C (four parameters included: density, SAS homophily, SAS sender and SAS receiver).

The x-axis represents the Socio-Academic Score of the sender of a nomination, and the 3 lines correspond to 3 fictive receivers with different SA Scores (-1.8 and 2.4 are the sample range of SAS, and 0.3 the middle of that range).

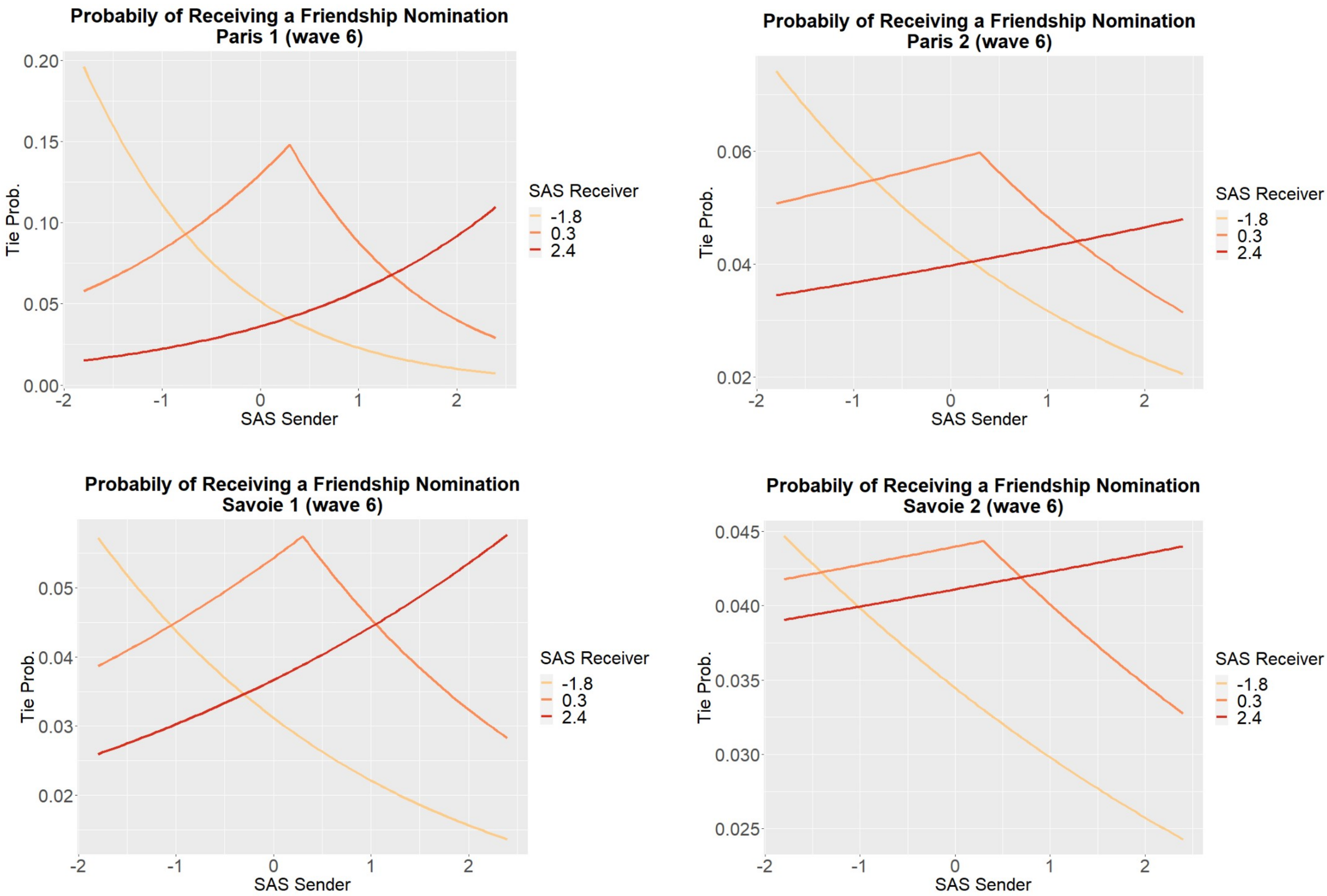


Figure 2: Probability of Receiving a nomination as "Very Good Friend" or "Friend" Depending on the Socioeconomic Background of the Sender and Receiver (ERGM) – Wave 1

Note: predictions derived from the ERGM of Table 2, Appendix 4C (four parameters included: density, SAS homophily, SAS sender and SAS receiver).

The x-axis represents the Socio-Academic Score of the sender of a nomination, and the 3 lines correspond to 3 fictive receivers with different SA Scores (-1.8 and 2.4 are the sample range of SAS, and 0.3 the middle of that range).

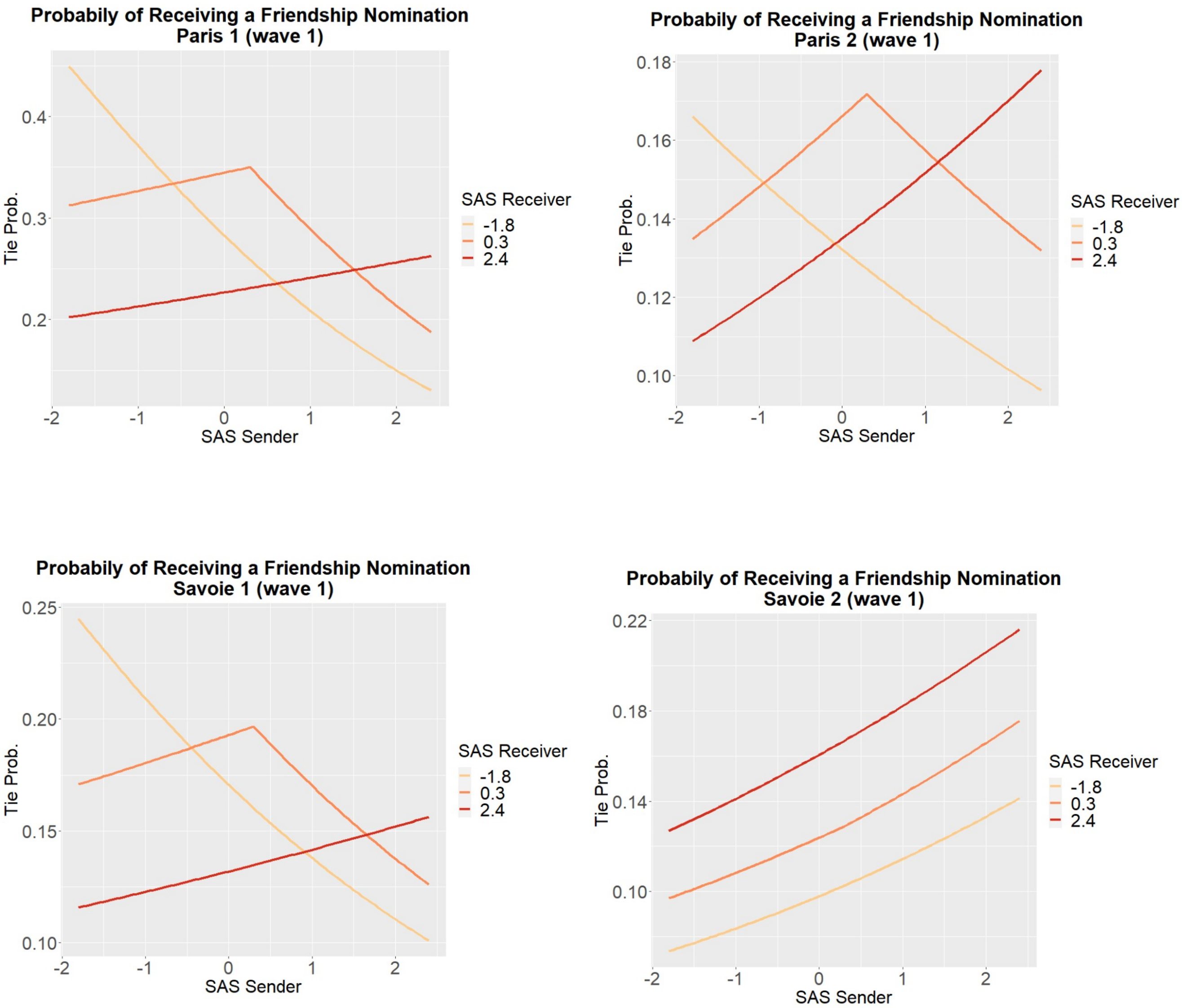
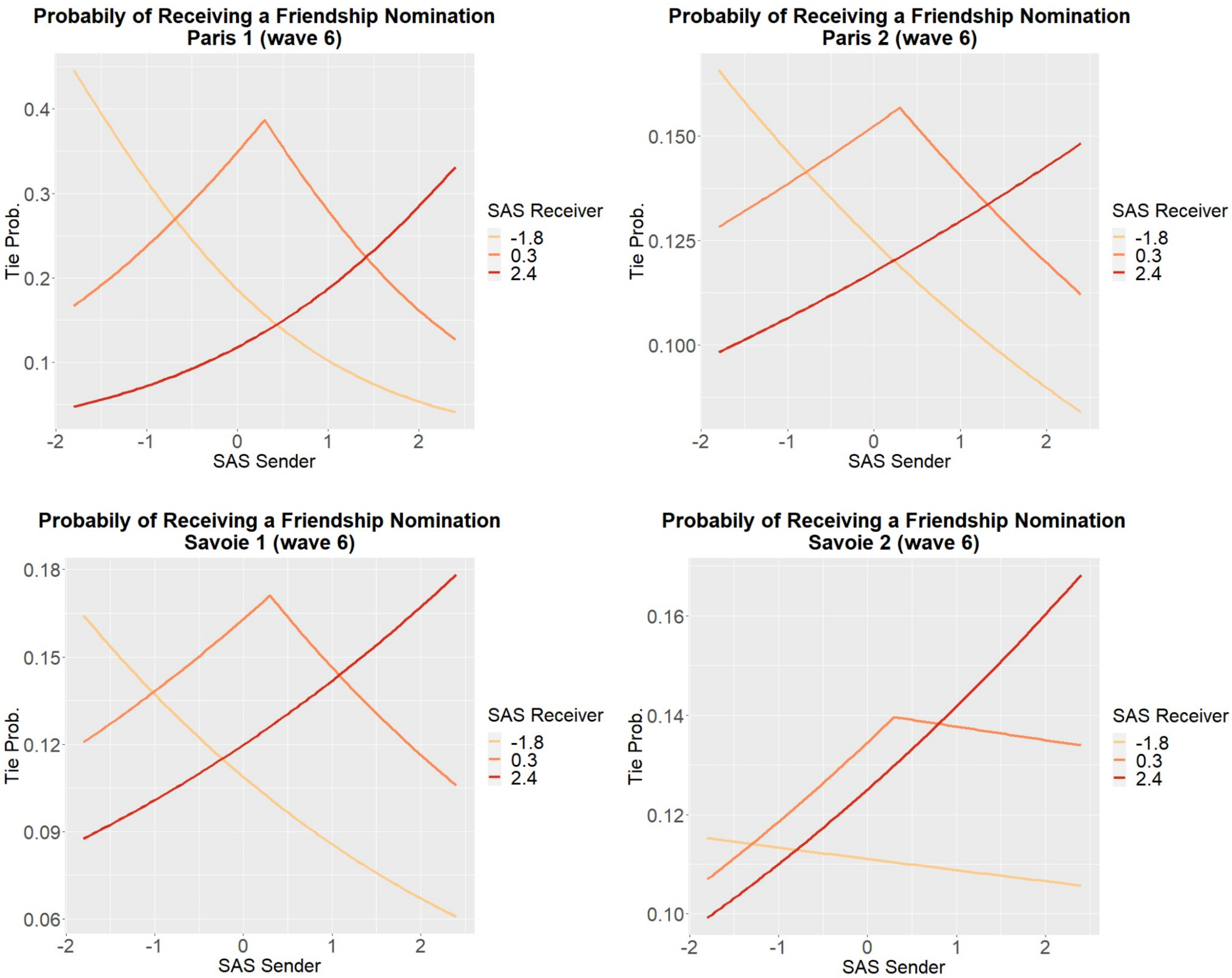


Figure 3: Probability of Receiving a nomination as “Very Good Friend” or “Friend” Depending on the Socioeconomic Background of the Sender and Receiver (ERGM) – Wave 6

Note: predictions derived from the ERGM of Table 2, Appendix 4C (four parameters included: density, SAS homophily, SAS sender and SAS receiver).

The x-axis represents the Socio-Academic Score of the sender of a nomination, and the 3 lines correspond to 3 fictive receivers with different SA Scores (-1.8 and 2.4 are the sample range of SAS, and 0.3 the middle of that range).



Appendix 4F: Supplementary Analyses for the Relationship Between Academic Achievement and Friendship Centrality

In Table 1 below, three nested ERGM specifications are applied to the network of "very good friend" nominations for each school and wave separately. Model 1 takes four parameters: density, homophily of academic achievement, and emission and reception of academic achievement. In Model 2, homophily, emission and reception terms are added for socioeconomic background as well. Finally, in Model 3, an interaction effect is added between the reception effects of academic achievement and socioeconomic background. In essence then, Model 1 is meant to give the "raw" homophily and centrality patterns associated to academic results; Model 2 is meant to show centrality patterns of academic achievement net of socioeconomic background; and Model 3 is meant to check whether the statutory impact of grades differ across socioeconomic groups.

Note that for Paris 2, grades are not available after wave 1, so the models are only estimated for this first wave.

Of particular interest in these models are the points already raised in the body of the text:

(a) There is academic homophily in all the schools.

(b) In Paris 1 and Paris 2, higher grades are associated with less friendship nominations, both received and emitted (Model 1 and 2). However, the negative association of grades and friendship indegree is weaker among upper-background students than among lower-background ones (positive interaction coefficient in Model 3).

(c) In Savoie 1, well-graded students emit less nominations, but do not receive less, or not much (Model 1 and 2). There is no significant interaction effect (Model 3).

(d) In Savoie 2, high grades are associated with more central positions, both in terms of received and emitted nominations (Model 1). This largely overlaps with the centrality of upper-background students: depending on the wave, either academic centrality explains away socioeconomic centrality, or the other way around (Model 2). There is no significant interaction effect (Model 3).

Table 1: Complete ERGMs for the Exploration of the Relationship between Academic Achievement, Socioeconomic Background and Network Centrality – Very Good Friend Networks, All Waves

	Paris 1			Paris 2			Savoie 1			Savoie 2		
	m1	**m2**	**m3**	**m1**	**m2**	**m3**	**m1**	**m2**	**m3**	**m1**	**m2**	**m3**
Wave 1												
Density	-1.914*** (0.075)	-1.699*** (0.093)	-1.722*** (0.093)	-2.490*** (0.025)	-2.311*** (0.031)	-2.311*** (0.031)	-2.626*** (0.048)	-2.421*** (0.060)	-2.419*** (0.060)	-3.034*** (0.044)	-3.041*** (0.055)	-3.043*** (0.055)
Absdiff grades	-0.433*** (0.061)	-0.386*** (0.063)	-0.408*** (0.064)	-0.125*** (0.019)	-0.115*** (0.019)	-0.123*** (0.019)	-0.278*** (0.040)	-0.244*** (0.041)	-0.234*** (0.041)	-0.172*** (0.034)	-0.177*** (0.034)	-0.173*** (0.034)
Grades receiver	-0.334*** (0.048)	-0.360*** (0.056)	-0.359*** (0.057)	-0.070*** (0.015)	-0.061*** (0.016)	-0.073*** (0.016)	-0.057* (0.033)	-0.007 (0.037)	-0.011 (0.037)	0.089*** (0.027)	0.037 (0.029)	0.045 (0.031)
Grades sender	-0.440*** (0.049)	-0.496*** (0.057)	-0.501*** (0.057)	-0.216*** (0.015)	-0.207*** (0.016)	-0.209*** (0.016)	-0.261*** (0.033)	-0.212*** (0.037)	-0.210*** (0.037)	0.088*** (0.027)	0.003 (0.029)	0.003 (0.029)
Absdiff SAS	-	-0.317*** (0.061)	-0.321*** (0.061)	-	-0.184*** (0.021)	-0.189*** (0.021)	-	-0.189*** (0.041)	-0.181*** (0.042)	-	-0.112*** (0.034)	-0.110*** (0.034)
SAS receiver	-	0.102* (0.055)	0.081 (0.056)	-	0.006 (0.017)	-0.002 (0.017)	-	-0.072** (0.036)	-0.056 (0.037)	-	0.148*** (0.029)	0.152*** (0.030)
SAS sender	-	0.166*** (0.055)	0.165*** (0.055)	-	-0.003 (0.017)	-0.002 (0.017)	-	-0.072** (0.037)	-0.072** (0.037)	-	0.243*** (0.029)	0.243*** (0.029)
Grades * SAS receiver	-	-	0.104* (0.054)	-	-	0.069*** (0.017)	-	-	-0.062 (0.039)	-	-	-0.021 (0.029)
Wave 2												
Density	-1.601*** (0.067)	-1.322*** (0.082)	-1.365*** (0.083)	-	-	-	-2.630*** (0.043)	-2.510*** (0.054)	-2.508*** (0.055)	-2.862*** (0.042)	-2.952*** (0.053)	-2.955*** (0.053)
Absdiff grades	-0.519*** (0.057)	-0.451*** (0.059)	-0.463*** (0.059)	-	-	-	-0.097*** (0.032)	-0.081** (0.032)	-0.076** (0.032)	-0.222*** (0.033)	-0.227*** (0.033)	-0.222*** (0.033)
Grades receiver	-0.113** (0.045)	-0.092* (0.052)	-0.132** (0.054)	-	-	-	0.053** (0.027)	0.058** (0.029)	0.059** (0.029)	0.155*** (0.027)	0.141*** (0.028)	0.147*** (0.029)
Grades sender	-0.327*** (0.045)	-0.372*** (0.052)	-0.374*** (0.052)	-	-	-	-0.248*** (0.027)	-0.246*** (0.029)	-0.245*** (0.029)	0.037 (0.027)	0.011 (0.028)	0.011 (0.028)
Absdiff SAS	-	-0.367*** (0.054)	-0.377*** (0.054)	-	-	-	-	-0.137*** (0.035)	-0.130*** (0.035)	-	0.056* (0.032)	0.058* (0.032)
SAS receiver	-	0.011 (0.047)	-0.020 (0.049)	-	-	-	-	0.006 (0.030)	0.019 (0.031)	-	0.036 (0.028)	0.042 (0.028)
SAS sender	-	0.152*** (0.047)	0.151*** (0.048)	-	-	-	-	0.017 (0.031)	0.017 (0.031)	-	0.076*** (0.028)	0.076*** (0.028)
Grades * SAS receiver	-	-	0.133*** (0.046)	-	-	-	-	-	-0.053 (0.035)	-	-	-0.022 (0.029)
Wave 3												
Density	-1.724*** (0.072)	-1.299*** (0.090)	-1.331*** (0.090)	-	-	-	-2.547*** (0.044)	-2.381*** (0.056)	-2.383*** (0.056)	-2.934*** (0.043)	-2.989*** (0.054)	-2.989*** (0.054)
Absdiff grades	-0.548*** (0.063)	-0.488*** (0.064)	-0.515*** (0.065)	-	-	-	-0.292*** (0.036)	-0.269*** (0.036)	-0.275*** (0.036)	-0.235*** (0.034)	-0.239*** (0.034)	-0.235*** (0.034)
Grades receiver	-0.112** (0.049)	-0.105* (0.054)	-0.152*** (0.056)	-	-	-	0.011 (0.029)	0.014 (0.032)	0.011 (0.032)	0.083*** (0.028)	0.063** (0.029)	0.070** (0.030)
Grades sender	-0.408*** (0.050)	-0.394*** (0.054)	-0.401*** (0.055)	-	-	-	-0.192*** (0.029)	-0.162*** (0.032)	-0.163*** (0.032)	-0.027 (0.028)	-0.043 (0.029)	-0.042 (0.029)
Absdiff SAS	-	-0.481*** (0.060)	-0.494*** (0.060)	-	-	-	-	-0.177*** (0.037)	-0.182*** (0.038)	-	0.030 (0.033)	0.033 (0.033)
SAS receiver	-	0.075 (0.050)	0.045 (0.052)	-	-	-	-	0.028 (0.032)	0.018 (0.034)	-	0.052* (0.029)	0.058** (0.029)

SAS sender	-	0.037 (0.051)	0.034 (0.051)	-	-	-	-	-0.047 (0.033)	-0.047 (0.033)	-	0.040 (0.029)	0.040 (0.029)
Grades * SAS receiver	-	-	0.161*** (0.048)	-	-	-	-	-	-0.039 (0.037)	-	-	-0.028 (0.031)
Wave 4												
Density	-1.821*** (0.072)	-1.386*** (0.091)	-1.427*** (0.090)	-	-	-	-2.487*** (0.042)	-2.403*** (0.054)	-2.401*** (0.054)	-2.873*** (0.043)	-2.815*** (0.054)	-2.815*** (0.054)
Absdiff grades	-0.454*** (0.063)	-0.374*** (0.065)	-0.410*** (0.065)	-	-	-	-0.340*** (0.035)	-0.327*** (0.036)	-0.324*** (0.036)	-0.206*** (0.034)	-0.201*** (0.034)	-0.201*** (0.034)
Grades receiver	-0.008 (0.050)	-0.031 (0.056)	-0.054 (0.057)	-	-	-	-0.009 (0.029)	-0.016 (0.032)	-0.015 (0.032)	0.107*** (0.028)	0.090*** (0.029)	0.090*** (0.029)
Grades sender	-0.262*** (0.050)	-0.296*** (0.056)	-0.304*** (0.057)	-	-	-	-0.252*** (0.029)	-0.211*** (0.032)	-0.210*** (0.032)	-0.050* (0.028)	-0.032 (0.029)	-0.032 (0.029)
Absdiff SAS	-	-0.544*** (0.059)	-0.558*** (0.059)	-	-	-	-	-0.080** (0.036)	-0.078** (0.036)	-	-0.067* (0.034)	-0.067* (0.034)
SAS receiver	-	0.094* (0.050)	0.040 (0.053)	-	-	-	-	0.030 (0.032)	0.034 (0.032)	-	0.073** (0.029)	0.073** (0.030)
SAS sender	-	0.108** (0.051)	0.105** (0.051)	-	-	-	-	-0.085*** (0.032)	-0.085*** (0.032)	-	-0.057* (0.029)	-0.057* (0.029)
Grades * SAS receiver	-	-	0.205*** (0.048)	-	-	-	-	-	-0.018 (0.034)	-	-	-0.0001 (0.030)
Wave 5												
Density	-1.984*** (0.080)	-1.495*** (0.095)	-1.504*** (0.096)	-	-	-	-2.571*** (0.047)	-2.386*** (0.060)	-2.368*** (0.061)	-2.949*** (0.048)	-2.883*** (0.060)	-2.884*** (0.060)
Absdiff grades	-0.582*** (0.080)	-0.390*** (0.084)	-0.398*** (0.085)	-	-	-	-0.422*** (0.041)	-0.403*** (0.041)	-0.393*** (0.041)	-0.207*** (0.038)	-0.200*** (0.038)	-0.192*** (0.039)
Grades receiver	-0.099 (0.063)	-0.021 (0.077)	-0.019 (0.077)	-	-	-	0.004 (0.033)	-0.011 (0.035)	0.009 (0.036)	0.036 (0.031)	0.027 (0.033)	0.036 (0.033)
Grades sender	-0.347*** (0.064)	-0.220*** (0.077)	-0.221*** (0.077)	-	-	-	-0.248*** (0.033)	-0.232*** (0.035)	-0.231*** (0.035)	-0.135*** (0.031)	-0.122*** (0.032)	-0.120*** (0.032)
Absdiff SAS	-	-0.597*** (0.071)	-0.598*** (0.071)	-	-	-	-	-0.202*** (0.040)	-0.195*** (0.040)	-	-0.071* (0.039)	-0.069* (0.039)
SAS receiver	-	0.018 (0.063)	0.006 (0.067)	-	-	-	-	0.054 (0.035)	0.059* (0.034)	-	0.040 (0.033)	0.045 (0.033)
SAS sender	-	-0.093 (0.063)	-0.094 (0.063)	-	-	-	-	-0.032 (0.035)	-0.032 (0.035)	-	-0.036 (0.033)	-0.036 (0.033)
Grades * SAS receiver	-	-	0.038 (0.067)	-	-	-	-	-	-0.108*** (0.039)	-	-	-0.041 (0.032)

Appendices for Chapter 5

Appendix 5A: Missing Values per Wave for Individual Attributes

Table 1: Missing Values per Wave for Individual Attributes (% of all the nodes present at this wave)

		Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
Paris 1	SAS	0.02	0.02	0.02	0.01	0.01
	Ethnicity	0	0	0	0	0
	Academic Results	0	0.01	0	0	0.01
	Cultural Status	0	0.09	0	0.06	NA
	Primary School	0.12	0.11	0.11	0.1	0.1
Paris 2	SAS	0.01	0.02	0.02	0	0.02
	Ethnicity	0	0	0	0	0
	Academic Results	0	1	1	1	1
	Cultural Status	0	0.05	0	0.04	NA
	Primary School	0.01	0.01	0.01	0.01	0.01
Savoie 1	SAS	0.02	0.02	0.02	0.01	0.02
	Ethnicity	0	0	0	0	0
	Academic Results	0.03	0.07	0.07	0.07	0.08
	Cultural Status	0.01	0.07	0.02	0.06	NA
	Primary School	0.05	0.05	0.05	0.04	0.04
Savoie 2	SAS	0.01	0.02	0.02	0.01	0.02
	Ethnicity	0.01	0.01	0.01	0.01	0.01
	Academic Results	0.08	0.02	0.01	0.01	0.01
	Cultural Status	0	0.03	0.01	0.03	NA
	Primary School	0.01	0.01	0.01	0.01	0.01

Note: there are no missing values for classrooms, special tracks and gender. The missing values for parental ties are the same as that of the friendship networks from wave 1. See Chapter 3, Table 3-1 for the number of students that answered the questionnaires at each wave (missing values for network variables correspond to the students that did not answer the questionnaires).

Appendix 5B: Complete SAOM Outputs for Waves 1 to 4

Note: these SAOMs correspond to Table 5-3 from the main text.

Table 1: Estimated SAOM Parameters for Paris 1 (very good friend networks, waves 1 to 4)

	Model 1			Model 2			Model 3		
	Estimate	Standard Error	Convergence T-Ratio	Estimate	Standard Error	Convergence T-Ratio	Estimate	Standard Error	Convergence T-Ratio
1. rate constant friendship rate (period 1)	14.32	(1.2100)	0.03	14.62	(1.2635)	-0.02	14.97	(1.2346)	0.04
2. rate constant friendship rate (period 2)	19.03	(1.4910)	-0.05	19.38	(1.5859)	0	19.65	(1.4868)	-0.03
3. rate constant friendship rate (period 3)	13.57	(1.1175)	-0.02	13.53	(1.1690)	-0.03	13.78	(1.1327)	-0.01
4. eval outdegree (density)	-2.84	(0.1203)	-0.01	-2.83	(0.1375)	-0.03	-2.85	(0.1223)	-0.01
5. eval reciprocity	2.36	(0.1891)	0.02	2.34	(0.1852)	0	2.32	(0.1786)	-0.01
6. eval GWESP I -> K -> J (110)	1.14	(0.0609)	0	1.12	(0.0602)	-0.01	1.12	(0.0617)	0.01
7. eval indegree - popularity	-0.03	(0.0111)	-0.01	-0.03	(0.0110)	-0.02	-0.03	(0.0109)	0
8. eval outdegree - popularity	-0.06	(0.0081)	0	-0.06	(0.0077)	-0.01	-0.06	(0.0077)	0
9. eval outdegree - activity	0	(0.0031)	-0.04	0	(0.0032)	-0.02	0	(0.0031)	0
10. eval m_3main_ethn5	0.13	(0.0600)	0.01	0.15	(0.0582)	-0.01	0.11	(0.0571)	-0.02
11. eval m_oib	0.03	(0.1661)	0	0	(0.1756)	-0.04	-0.01	(0.1597)	-0.02
12. eval n_oib	0.52	(0.1365)	0.02	0.45	(0.1328)	0.03	0.5	(0.1341)	0.01
13. eval s_mena_ethn5	-0.23	(0.0848)	0.02	-0.21	(0.0843)	0.01	-0.24	(0.0786)	-0.02
14. eval s_europe_ethn5	0.07	(0.1116)	-0.04	0.05	(0.1171)	-0.01	0.04	(0.1147)	0
15. eval s_subsah_ethn5	-0.04	(0.1036)	-0.03	-0.02	(0.0964)	0	-0.03	(0.0924)	0.02
16. eval s_other_ethn5	0.51	(0.1690)	-0.03	0.57	(0.1621)	0	0.52	(0.1600)	-0.01
17. eval r_mena_ethn5	0.08	(0.0778)	0.03	0.04	(0.0794)	0.03	0.07	(0.0716)	0.01
18. eval r_europe_ethn5	0.05	(0.1039)	0.03	0	(0.1076)	-0.05	0.05	(0.0992)	0.03
19. eval r_subsah_ethn5	0.29	(0.0932)	0.03	0.28	(0.0903)	0	0.31	(0.0878)	-0.01
20. eval r_other_ethn5	0.02	(0.1583)	-0.02	0.08	(0.1634)	0.03	0.04	(0.1519)	0.02
21. eval walk_rg	-0.27	(0.2925)	0.03	-0.1	(0.2745)	0	-0.16	(0.2732)	0
22. eval walk_sq	0.28	(0.2236)	0.03	0.15	(0.2136)	0	0.2	(0.2146)	0.01
23. eval s_5B_classroom	0.13	(0.0774)	0	0.15	(0.0750)	0.02	0.13	(0.0751)	0.04
24. eval s_5C_classroom	0.24	(0.0980)	-0.04	0.28	(0.0912)	-0.02	0.24	(0.0924)	-0.01
25. eval r_5B_classroom	0.02	(0.0867)	0.01	0.04	(0.0798)	0.03	0.03	(0.0805)	0.02
26. eval r_5C_classroom	0.15	(0.0967)	-0.06	0.15	(0.0913)	-0.01	0.15	(0.0907)	-0.02
27. eval s_4B_classroom	0.19	(0.1327)	-0.01	0.19	(0.1302)	0.01	0.2	(0.1314)	-0.01
28. eval s_4C_classroom	0.33	(0.1461)	0.01	0.36	(0.1402)	-0.04	0.33	(0.1392)	-0.01
29. eval r_4B_classroom	-0.14	(0.1448)	0	-0.12	(0.1357)	0	-0.13	(0.1373)	-0.02
30. eval r_4C_classroom	0.44	(0.1370)	0.01	0.47	(0.1339)	-0.03	0.43	(0.1379)	-0.01
31. eval sas alter	0.1	(0.0359)	-0.04	0.11	(0.0356)	-0.06	0.09	(0.0334)	-0.02
32. eval sas ego	0	(0.0383)	0	-0.01	(0.0377)	-0.01	0	(0.0372)	-0.02

33. eval sas similarity	0.36	(0.1353)	-0.01	0.3	(0.1413)	-0.02	0.33	(0.1395)	-0.01
34. eval sex alter	0.02	(0.0577)	-0.03	0	(0.0585)	-0.01	0.02	(0.0567)	-0.01
35. eval sex ego	-0.24	(0.0649)	-0.02	-0.26	(0.0625)	-0.01	-0.25	(0.0603)	-0.01
36. eval sex similarity	0.47	(0.0528)	0.01	0.45	(0.0546)	0.01	0.46	(0.0523)	-0.02
37. eval same primary1	0.03	(0.0635)	0.03	0.03	(0.0601)	-0.02	0	(0.0603)	0.01
38. eval oib ego	-0.12	(0.1930)	-0.01	-0.23	(0.1898)	-0.03	-0.1	(0.1875)	-0.02
39. eval same classroom1	0.17	(0.0654)	0	0.19	(0.0627)	-0.04	0.18	(0.0627)	-0.01
40. eval same classroom2	0.49	(0.0656)	0.01	0.49	(0.0640)	-0.04	0.48	(0.0622)	-0.02
41. eval grades alter	-0.01	(0.0367)	-0.02	0.01	(0.0383)	-0.04	-0.02	(0.0360)	0
42. eval grades ego	-0.03	(0.0364)	0.02	-0.02	(0.0396)	-0.02	-0.03	(0.0373)	-0.03
43. eval grades similarity	0.57	(0.1690)	0.03	0.42	(0.1646)	0	0.54	(0.1620)	-0.01
44. eval timedummy_period2 ego	-0.35	(0.0672)	-0.07	-0.34	(0.0637)	0.01	-0.33	(0.0645)	0
45. eval timedummy_period3 ego	-0.48	(0.1251)	0.02	-0.54	(0.1312)	-0.02	-0.45	(0.1193)	-0.03
46. eval int. reciprocity x GWESP I -> K -> J (110)	-0.49	(0.1116)	0.03	-0.5	(0.1088)	0	-0.47	(0.1065)	0
47. eval match_music	-	-	-	0.12	(0.0322)	0	-	-	-
48. eval match_youtube	-	-	-	0.06	(0.0339)	-0.01	-	-	-
49. eval cult_lvl alter	-	-	-	-0.03	(0.0411)	-0.06	-	-	-
50. eval cult_lvl ego	-	-	-	0.06	(0.0410)	-0.03	-	-	-
51. eval cult_lvl similarity	-	-	-	0.24	(0.1586)	0.02	-	-	-
52. eval w1_parents	-	-	-	-	-	-	0.47	(0.1105)	-0.03
Overall Maximum Convergence Ratio	*0.24*			*0.15*			*0.14*		
Total Number of Iteration Steps	*2954*			*8880*			*8558*		

Note: “s_” and “r_” indicate, respectively, sender and receiver effects. These are essentially equivalent to “ego” and “alter” effects, except that they have been passed through SIENA as a dyadic covariate (effect “X”) rather than applying the ego and alter effect to an individual attribute. “Ethn5” refer to the five-posts ethnicity attribute. For the homophily effect, only the three main ethnic groups (natives, Middle-East and Subsaharian Africa) are counted. “walk_rg” is the linear walking distance between students’ homes, and “walk_sq” the squared distance. “Cult_lvl” indicate the attribute of cultural status, and “match_music” and “match_youtube” the covariates for cultural matching. “w1_parents” is the effect of the parental ties from wave 1. All other parameters have the effect names given by the Rsiena software.
This note applies to all further SAOM tables thereafter.

Table 2: Estimated SAOM Parameters for Paris 2 (very good friend networks, waves 1 to 4)

	Model 1			Model 2			Model 3		
	Estimate	Standard Error	Convergence T-Ratio	Estimate	Standard Error	Convergence T-Ratio	Estimate	Standard Error	Convergence T-Ratio
1. rate constant friendship rate (period 1)	41.47	(1.5134)	0.01	42.16	(1.0684)	0.02	44.05	(1.4125)	-0.01
2. rate constant friendship rate (period 2)	36.23	(1.1350)	0.02	36.61	(1.3340)	0.05	37.69	(1.2225)	0.03
3. rate constant friendship rate (period 3)	33.03	(0.9918)	0.01	33.45	(1.0323)	0.01	34.18	(0.9282)	-0.03
4. eval outdegree (density)	-3.59	(0.0711)	0.06	-3.54	(0.0640)	0.06	-3.63	(0.0709)	0.02
5. eval reciprocity	3.53	(0.1402)	0.06	3.5	(0.1225)	0.03	3.44	(0.1340)	0.03
6. eval GWESP I -> K -> J (69)	1.56	(0.0559)	0.06	1.54	(0.0492)	0.05	1.55	(0.0500)	0.02
7. eval GWESP I <- K <- J (300)	0.21	(0.0138)	0.08	0.21	(0.0118)	0.05	0.2	(0.0154)	0.02
8. eval GWDSP I -> K -> J (69)	0.03	(0.0044)	0.06	0.03	(0.0035)	0.05	0.03	(0.0039)	0.03
9. eval indegree - popularity (sqrt)	0.21	(0.0178)	0.06	0.2	(0.0176)	0.05	0.21	(0.0174)	0.02
10. eval outdegree - popularity (sqrt)	-0.66	(0.0425)	0.07	-0.66	(0.0342)	0.05	-0.64	(0.0417)	0.03
11. eval outdegree - activity	0.01	(0.0006)	0.05	0.01	(0.0005)	0.07	0.01	(0.0005)	0.03
12. eval m_3main_ethn5	0.2	(0.0250)	-0.02	0.19	(0.0248)	0.08	0.18	(0.0254)	-0.01
13. eval s_mena_ethn5	0.07	(0.0390)	0.01	0.09	(0.0392)	0.09	0.07	(0.0386)	0.03
14. eval s_europe_ethn5	0.31	(0.0582)	0	0.27	(0.0571)	-0.05	0.3	(0.0552)	0
15. eval s_subsah_ethn5	0.04	(0.0439)	-0.01	0.03	(0.0451)	0	0.04	(0.0437)	0
16. eval s_other_ethn5	0.66	(0.0583)	0	0.66	(0.0561)	-0.05	0.65	(0.0502)	-0.04
17. eval r_mena_ethn5	0.04	(0.0360)	0	0.02	(0.0342)	0.06	0.04	(0.0334)	-0.05
18. eval r_europe_ethn5	0.24	(0.0530)	0.05	0.2	(0.0507)	0.03	0.24	(0.0515)	0.04
19. eval r_subsah_ethn5	0.27	(0.0426)	0.01	0.23	(0.0391)	-0.03	0.26	(0.0381)	0.06
20. eval r_other_ethn5	0.3	(0.0511)	-0.02	0.31	(0.0510)	0	0.29	(0.0514)	0
21. eval walk_rg	-0.21	(0.0624)	0.02	-0.23	(0.0566)	0.02	-0.12	(0.0585)	0.01
22. eval walk_sq	0.07	(0.0273)	-0.01	0.09	(0.0251)	-0.02	0.04	(0.0267)	0
23. eval sas alter	0.01	(0.0111)	-0.01	0	(0.0117)	-0.04	0.01	(0.0120)	-0.01
24. eval sas ego	-0.04	(0.0126)	-0.06	-0.06	(0.0134)	-0.05	-0.03	(0.0119)	-0.02
25. eval sas similarity	0.11	(0.0585)	0.02	0.09	(0.0583)	0.07	0.09	(0.0571)	0.02
26. eval sex alter	0.07	(0.0252)	0.03	0.05	(0.0262)	0	0.07	(0.0244)	0.01
27. eval sex ego	-0.05	(0.0258)	0.01	-0.1	(0.0279)	-0.02	-0.05	(0.0257)	0
28. eval sex similarity	0.32	(0.0258)	0.03	0.31	(0.0217)	0.05	0.31	(0.0217)	0.01
29. eval same primary1	-0.02	(0.0296)	0.02	-0.03	(0.0281)	-0.01	-0.11	(0.0292)	0.03
30. eval grades alter	-0.03	(0.0120)	0.08	-0.02	(0.0119)	0.01	-0.03	(0.0116)	0
31. eval grades ego	-0.08	(0.0123)	0.03	-0.08	(0.0132)	0	-0.08	(0.0122)	-0.04
32. eval grades similarity	0.28	(0.0661)	0.02	0.25	(0.0677)	-0.05	0.23	(0.0673)	0.01
33. eval same classroom1	0.08	(0.0222)	0.04	0.09	(0.0223)	0	0.1	(0.0235)	0
34. eval same classroom2	0.9	(0.0277)	0.04	0.9	(0.0282)	0.01	0.91	(0.0287)	-0.02

35. eval timedummy_period2 ego	0.2	(0.0285)	-0.01	0.19	(0.0308)	-0.02	0.2	(0.0289)	-0.02
36. eval timedummy_period3 ego	0.07	(0.0268)	0.03	0.05	(0.0269)	0.01	0.08	(0.0268)	-0.01
37. eval int. reciprocity x GWESP I -> K -> J (69)	-1.28	(0.0810)	0.06	-1.27	(0.0757)	0.03	-1.25	(0.0834)	0.03
38. eval timedummy_period3 ego x sas similarity	0.4	(0.1316)	0.02	0.38	(0.1223)	-0.01	0.41	(0.1201)	-0.06
39. eval timedummy_period2 ego x same classroom1	-0.67	(0.0467)	0.03	-0.65	(0.0469)	-0.04	-0.65	(0.0463)	-0.02
40. eval match_music	-	-	-	0.08	(0.0139)	0.02	-	-	-
41. eval match_youtube	-	-	-	0.03	(0.0155)	0.02	-	-	-
42. eval cult_lvl alter	-	-	-	0.02	(0.0118)	0.02	-	-	-
43. eval cult_lvl ego	-	-	-	0.09	(0.0135)	-0.02	-	-	-
44. eval cult_lvl similarity	-	-	-	0.13	(0.0540)	0.02	-	-	-
45. eval w1_parents	-	-	-	-	-	-	0.97	(0.0546)	0
Overall Maximum Convergence Ratio		*0.2*			*0.24*			*0.19*	
Total Number of Iteration Steps		*1797*			*1813*			*2029*	

Table 3: Estimated SAOM Parameters for Savoie 1 (very good friend networks, waves 1 to 4)

		Model 1			**Model 2**			**Model 3**	
	Estimate	Standard Error	Convergence T-Ratio	Estimate	Standard Error	Convergence T-Ratio	Estimate	Standard Error	Convergence T-Ratio
1. rate constant friendship rate (period 1)	25.88	(1.5111)	0.05	25.84	(1.0833)	0.02	28.63	(1.6660)	0.01
2. rate constant friendship rate (period 2)	25.19	(1.0908)	-0.04	25.15	(1.1796)	0.01	26.83	(1.3196)	0
3. rate constant friendship rate (period 3)	20.97	(1.1258)	-0.02	20.99	(0.9330)	-0.04	21.98	(0.9944)	0
4. eval outdegree (density)	-2.86	(0.0513)	-0.05	-2.87	(0.0542)	-0.03	-2.87	(0.0490)	-0.01
5. eval reciprocity	2.39	(0.0819)	-0.04	2.39	(0.0866)	-0.02	2.33	(0.0834)	0
6. eval GWESP I -> K -> J (125)	1.02	(0.0432)	-0.05	1.02	(0.0408)	-0.02	1	(0.0409)	0
7. eval GWESP I <- K <- J (125)	0.15	(0.0437)	-0.05	0.15	(0.0402)	-0.03	0.14	(0.0405)	0
8. eval indegree - popularity	-0.01	(0.0042)	-0.05	-0.01	(0.0041)	-0.03	-0.01	(0.0042)	0
9. eval outdegree - popularity	-0.05	(0.0048)	-0.05	-0.05	(0.0047)	-0.02	-0.05	(0.0047)	0
10. eval outdegree - activity	0	(0.0012)	-0.03	0	(0.0012)	0	0	(0.0012)	-0.01
11. eval m_3main_ethn5	0.19	(0.0455)	0.03	0.18	(0.0485)	0.03	0.15	(0.0458)	-0.03
12. eval s_mena_ethn5	0.06	(0.0534)	-0.03	0.11	(0.0510)	-0.01	0.07	(0.0499)	0.02
13. eval s_europe_ethn5	0.21	(0.0505)	0.03	0.24	(0.0510)	-0.02	0.19	(0.0490)	0.03
14. eval s_subsah_ethn5	-0.33	(0.1244)	-0.09	-0.29	(0.1223)	0.05	-0.3	(0.1242)	0.02
15. eval s_other_ethn5	0.12	(0.1132)	0.08	0.12	(0.1108)	-0.07	0.21	(0.1115)	0.01
16. eval r_mena_ethn5	0.07	(0.0481)	-0.03	0.07	(0.0533)	-0.01	0.08	(0.0480)	-0.01

17. eval r_europe_ethn5	0.15	(0.0464)	-0.02	0.13	(0.0487)	0	0.13	(0.0444)	0.01
18. eval r_subsah_ethn5	0.28	(0.0998)	0.01	0.33	(0.1014)	0.01	0.29	(0.0951)	-0.03
19. eval r_other_ethn5	-0.13	(0.1140)	-0.01	-0.12	(0.1228)	-0.06	-0.06	(0.1108)	0.03
20. eval walk_rg	-0.06	(0.0195)	0.01	-0.06	(0.0218)	0.04	-0.04	(0.0193)	0.02
21. eval walk_sq	0.01	(0.0019)	0	0.01	(0.0020)	0.03	0.01	(0.0018)	0.01
22. eval ma_classroom	-0.19	(0.1650)	0	-0.23	(0.1617)	-0.01	-0.12	(0.1530)	0
23. eval na_classroom	0.41	(0.1108)	-0.02	0.41	(0.1134)	-0.01	0.38	(0.1173)	0
24. eval sas alter	-0.01	(0.0167)	0.01	0.01	(0.0178)	0	-0.01	(0.0172)	-0.01
25. eval sas ego	-0.01	(0.0191)	0.05	-0.04	(0.0197)	-0.02	-0.01	(0.0187)	-0.02
26. eval sas similarity	0.05	(0.0804)	-0.06	0.04	(0.0730)	0.01	0	(0.0721)	0
27. eval sex alter	0.01	(0.0356)	-0.01	0	(0.0352)	0.01	0	(0.0332)	0.02
28. eval sex ego	0.09	(0.0359)	0.02	0.03	(0.0365)	0.02	0.08	(0.0353)	0.02
29. eval sex similarity	0.28	(0.0316)	-0.02	0.27	(0.0322)	0	0.25	(0.0315)	0.01
30. eval same primary1	0.24	(0.0410)	-0.04	0.24	(0.0402)	-0.02	0	(0.0443)	-0.03
31. eval same classroom1	0.18	(0.0343)	-0.04	0.18	(0.0340)	0.02	0.19	(0.0324)	0.02
32. eval same classroom2	0.72	(0.0405)	-0.01	0.71	(0.0387)	-0.02	0.71	(0.0375)	0.01
33. eval grades alter	0.02	(0.0175)	-0.01	0.04	(0.0191)	0	0.01	(0.0175)	0.01
34. eval grades ego	-0.07	(0.0210)	0.02	-0.09	(0.0199)	-0.03	-0.09	(0.0199)	-0.01
35. eval grades similarity	0.18	(0.1116)	-0.01	0.19	(0.1114)	-0.05	0.17	(0.1092)	-0.01
36. eval timedummy_period2 ego	0	(0.0444)	0.03	0	(0.0425)	0.01	0.01	(0.0418)	-0.01
37. eval timedummy_period3 ego	-0.09	(0.0359)	-0.03	-0.12	(0.0357)	0	-0.08	(0.0367)	0.01
38. eval segpa ego	0.61	(0.1553)	0	0.66	(0.1597)	-0.01	0.63	(0.1648)	0.01
39. eval int. reciprocity x GWESP I -> K -> J (125)	-0.59	(0.0495)	-0.05	-0.58	(0.0533)	-0.03	-0.59	(0.0488)	0
40. eval timedummy_period2 ego x same classroom1	-0.54	(0.0712)	0.01	-0.54	(0.0711)	0.02	-0.52	(0.0677)	-0.01
41. eval match_music	-	-	-	0.07	(0.0185)	-0.03	-	-	-
42. eval match_youtube	-	-	-	0.02	(0.0168)	0	-	-	-
43. eval cult_lvl alter	-	-	-	-0.02	(0.0171)	0.01	-	-	-
44. eval cult_lvl ego	-	-	-	0.14	(0.0190)	0.04	-	-	-
45. eval cult_lvl similarity	-	-	-	0.07	(0.0765)	-0.03	-	-	-
46. eval w1_parents	-	-	-	-	-	-	0.69	(0.0502)	-0.03
Overall Maximum Convergence Ratio		*0.23*			*0.22*			*0.1*	
Total Number of Iteration Steps		*1893*			*1719*			*7874*	

Table 4: Estimated SAOM Parameters for Savoie 2 (very good friend networks, waves 1 to 4)

	Model 1			Model 2			Model 3		
	Estimate	Standard Error	Convergence T-Ratio	Estimate	Standard Error	Convergence T-Ratio	Estimate	Standard Error	Convergence T-Ratio
1. rate constant friendship rate (period 1)	23.95	-1.03	0.01	24.07	(1.0332)	0	26.03	(1.2267)	0.03
2. rate constant friendship rate (period 2)	23.19	-1.22	0.04	23.21	(1.0363)	0.03	24.58	(1.3671)	0.05
3. rate constant friendship rate (period 3)	17.93	-0.74	-0.05	17.95	(0.7980)	0.01	18.76	(0.9503)	0.04
4. eval outdegree (density)	-3.12	-0.05	0.03	-3.08	(0.0573)	0	-3.14	(0.0584)	0.05
5. eval reciprocity	2.72	-0.08	0.01	2.72	(0.0928)	0	2.67	(0.0836)	0.03
6. eval GWESP I -> K -> J (85)	1.36	-0.04	0.02	1.36	(0.0359)	0	1.33	(0.0353)	0.05
7. eval indegree - popularity	-0.01	0	0.02	-0.01	(0.0043)	0.02	0	(0.0041)	0.04
8. eval outdegree - popularity	-0.05	0	0.03	-0.05	(0.0038)	0	-0.05	(0.0033)	0.04
9. eval outdegree - activity	0.01	0	0.03	0.01	(0.0012)	0.01	0.01	(0.0011)	0.03
10. eval m_3main_ethn5	0.19	-0.05	0.01	0.18	(0.0499)	0.04	0.16	(0.0478)	-0.02
11. eval s_mena_ethn5	0.11	-0.06	-0.01	0.12	(0.0614)	-0.01	0.1	(0.0563)	-0.01
12. eval s_europe_ethn5	0.18	-0.05	0.01	0.18	(0.0545)	-0.08	0.16	(0.0517)	0.03
13. eval s_subsah_ethn5	0.13	-0.1	0.01	0.11	(0.1046)	0.03	0.08	(0.1007)	-0.01
14. eval s_other_ethn5	0.38	-0.1	-0.06	0.35	(0.1041)	-0.02	0.36	(0.1024)	0
15. eval r_mena_ethn5	0.07	-0.06	0.04	0.1	(0.0619)	0.02	0.07	(0.0578)	0.01
16. eval r_europe_ethn5	0.21	-0.05	0.01	0.2	(0.0484)	-0.04	0.2	(0.0480)	0
17. eval r_subsah_ethn5	0.27	-0.1	-0.01	0.27	(0.0929)	0.1	0.26	(0.0950)	0.04
18. eval r_other_ethn5	0.31	-0.09	-0.05	0.34	(0.0958)	0.02	0.28	(0.0896)	0
19. eval walk_rg	-0.05	-0.03	-0.01	-0.05	(0.0257)	-0.02	-0.03	(0.0268)	0.01
20. eval walk_sq	0.01	0	-0.03	0.01	(0.0044)	0	0	(0.0045)	0.02
21. eval sas alter	0.02	-0.02	0	0.02	(0.0160)	-0.01	0	(0.0160)	-0.02
22. eval sas ego	-0.09	-0.02	-0.05	-0.1	(0.0173)	0.08	-0.1	(0.0167)	-0.04
23. eval sas similarity	-0.05	-0.07	0.05	-0.07	(0.0734)	0.01	-0.08	(0.0735)	0.04
24. eval sex alter	0.02	-0.03	-0.03	0.03	(0.0330)	0	0.03	(0.0304)	-0.01
25. eval sex ego	-0.05	-0.03	-0.01	-0.02	(0.0355)	0.01	-0.02	(0.0330)	-0.02
26. eval sex similarity	0.33	-0.03	-0.02	0.32	(0.0296)	0	0.31	(0.0287)	0.03
27. eval same primary1	0.17	-0.04	0.01	0.16	(0.0364)	0.05	0.03	(0.0387)	0.05
28. eval same classroom1	0.1	-0.03	0.02	0.1	(0.0307)	-0.03	0.13	(0.0320)	0.02
29. eval same classroom2	1.09	-0.04	0.03	1.09	(0.0406)	-0.01	1.08	(0.0402)	0.04
30. eval grades alter	0.01	-0.02	0.04	0.01	(0.0160)	0	0.01	(0.0153)	-0.02
31. eval grades ego	-0.06	-0.02	0.03	-0.07	(0.0182)	0.01	-0.07	(0.0167)	-0.06
32. eval grades similarity	0.09	-0.1	0.09	0.09	(0.0891)	0.01	0.08	(0.0872)	0
33. eval timedummy_period2 ego	-0.06	-0.04	0.06	-0.07	(0.0381)	0.02	-0.04	(0.0392)	-0.05
34. eval timedummy_period3 ego	-0.05	-0.03	-0.05	-0.06	(0.0355)	0.01	-0.02	(0.0353)	0.05
35. eval int. reciprocity x GWESP I -> K -> J (85)	-0.8	-0.07	0	-0.8	(0.0774)	-0.02	-0.81	(0.0719)	0.04

36. eval timedummy_period2 ego x same classroom1	-0.6	-0.07	0.06	-0.6	(0.0735)	0.01	-0.59	(0.0745)	0.01
37. eval match_music	-	-	-	0.09	(0.0196)	0.03	-	-	-
38. eval match_youtube	-	-	-	-0.02	(0.0179)	-0.01	-	-	-
39. eval cult_lvl alter	-	-	-	0	(0.0181)	0.05	-	-	-
40. eval cult_lvl ego	-	-	-	0.05	(0.0179)	0.02	-	-	-
41. eval cult_lvl similarity	-	-	-	0.21	(0.0725)	-0.02	-	-	-
42. eval w1_parents	-	-	-	-	-	-	0.66	(0.0499)	0
Overall Maximum Convergence Ratio		*0.22*			*0.23*			*0.18*	
Total Number of Iteration Steps		*2337*			*1919*			*1881*	

Figure 1: Goodness-of-fit Plots for Model 1 – Paris 1

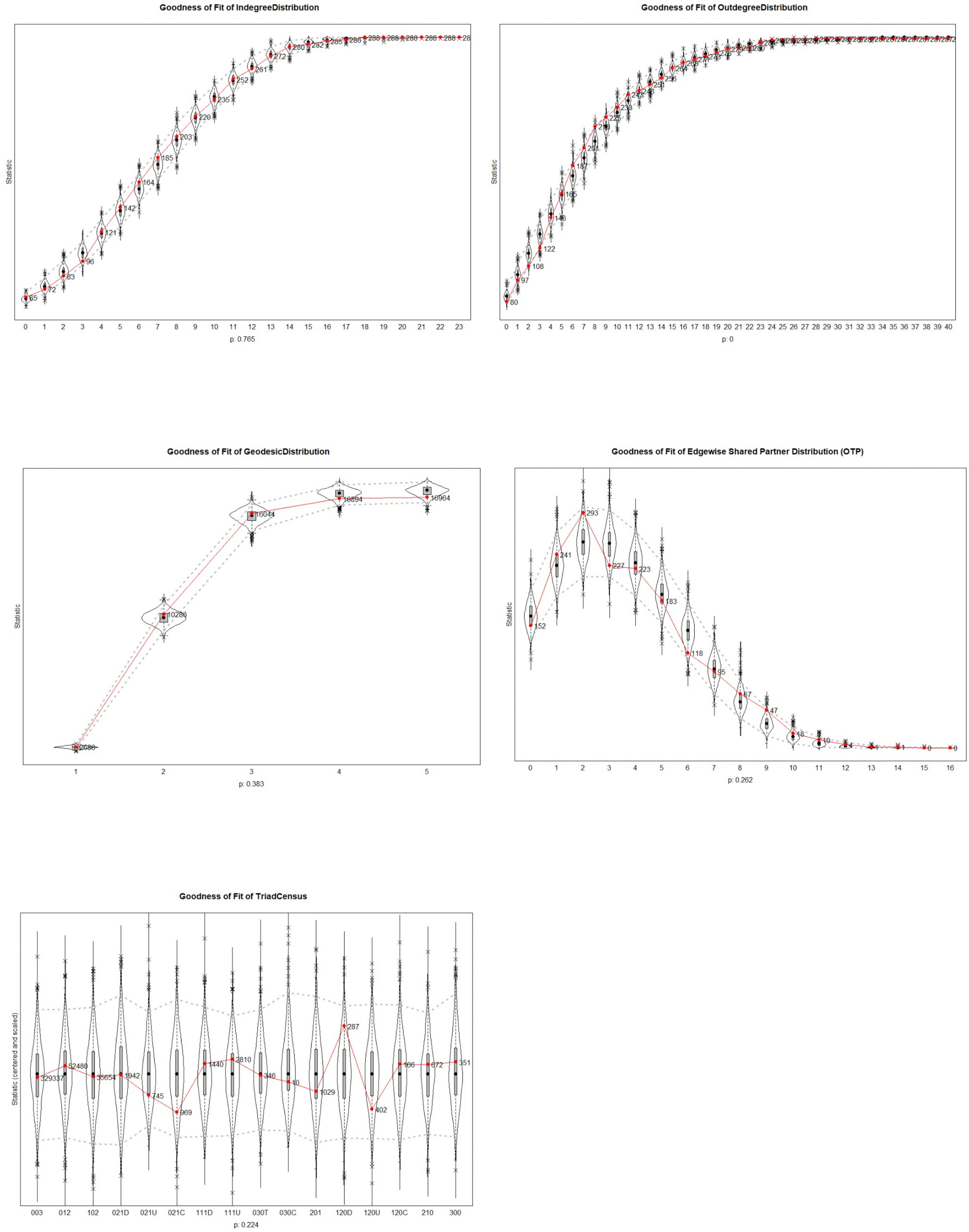


Figure 2: Goodness-of-fit Plots for Model 1 – Paris 2

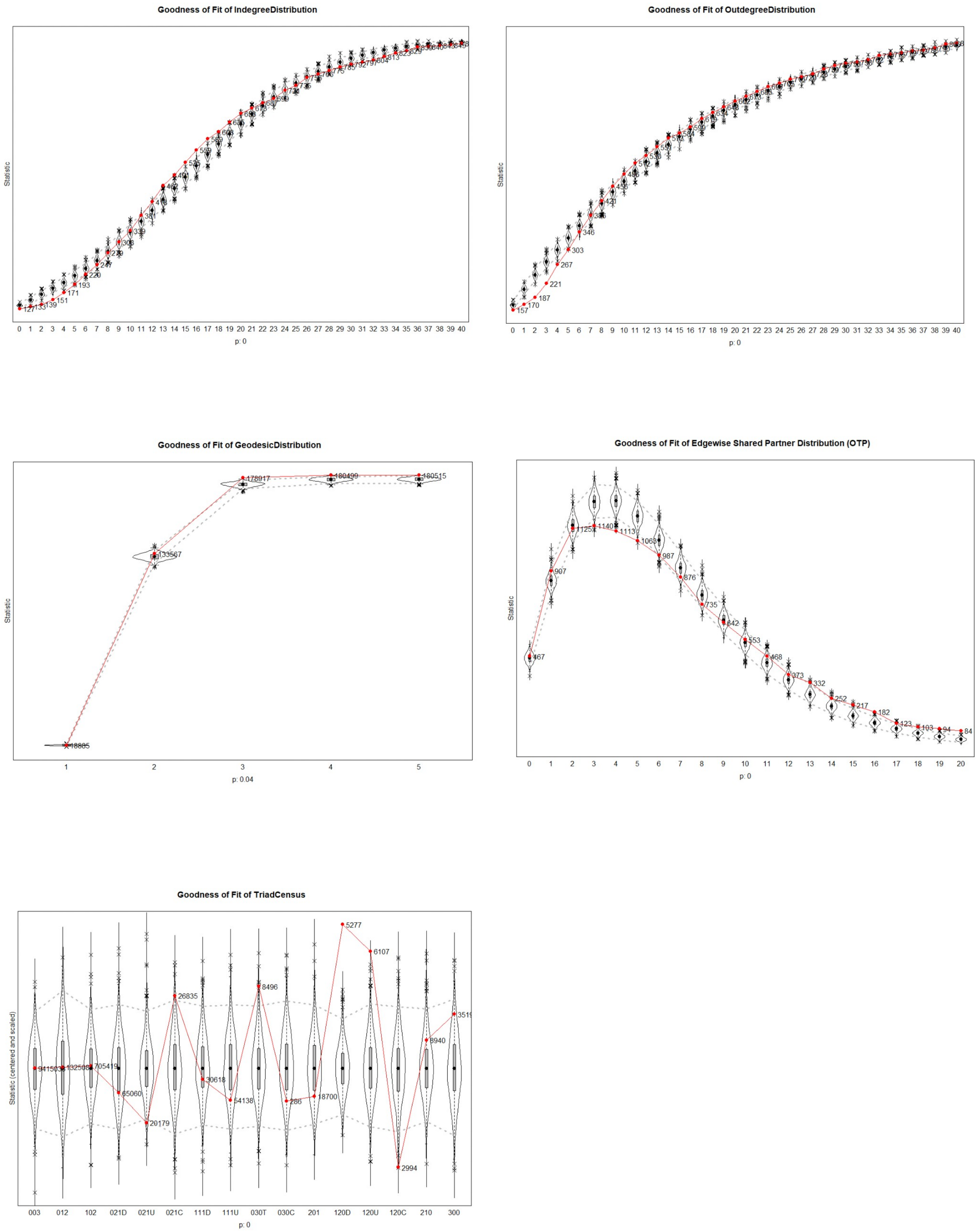


Figure 3: Goodness-of-fit Plots for Model 1 – Savoie 1

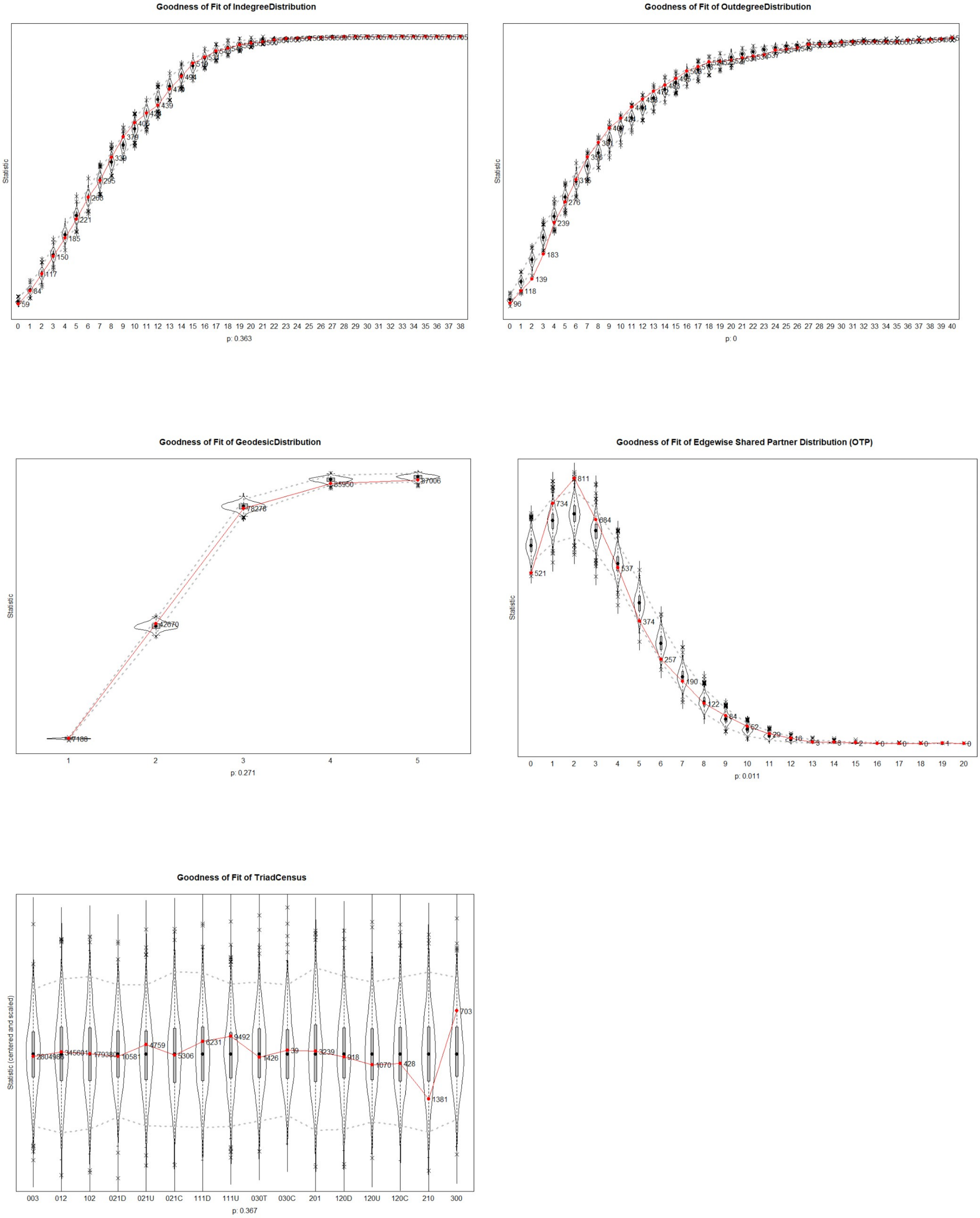
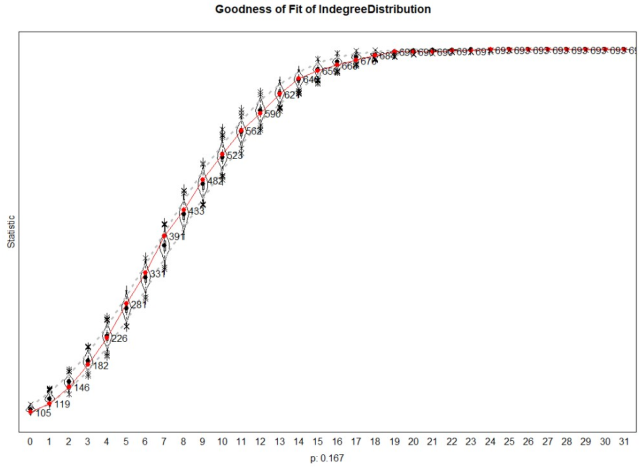
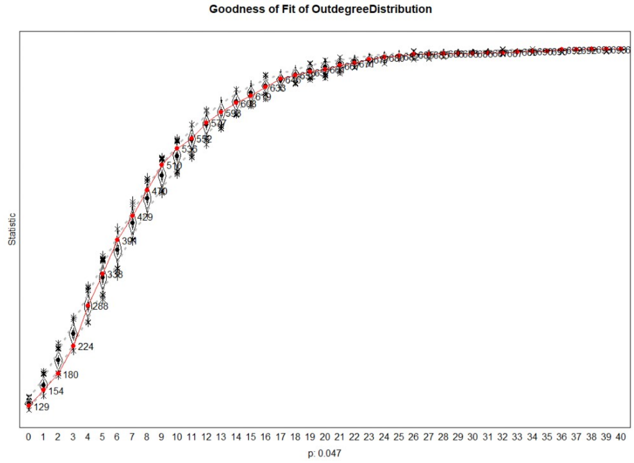
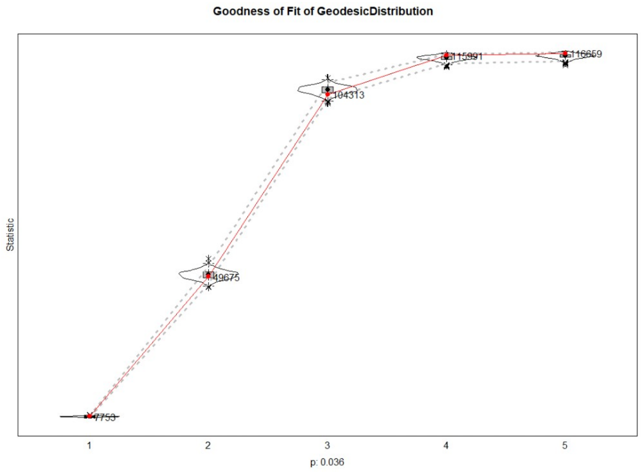
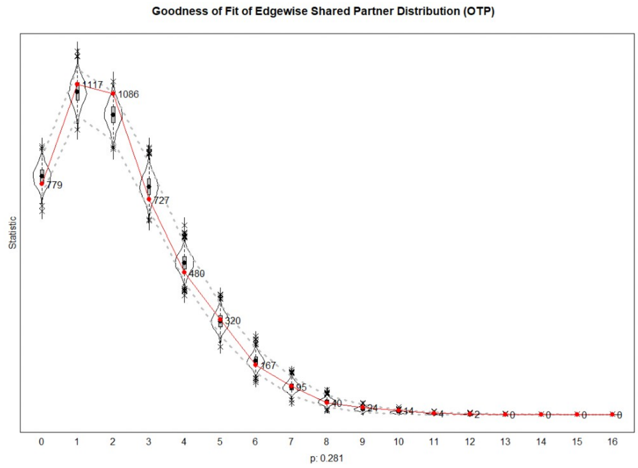


Figure 4: Goodness-of-fit Plots for Model 1 – Savoie 2









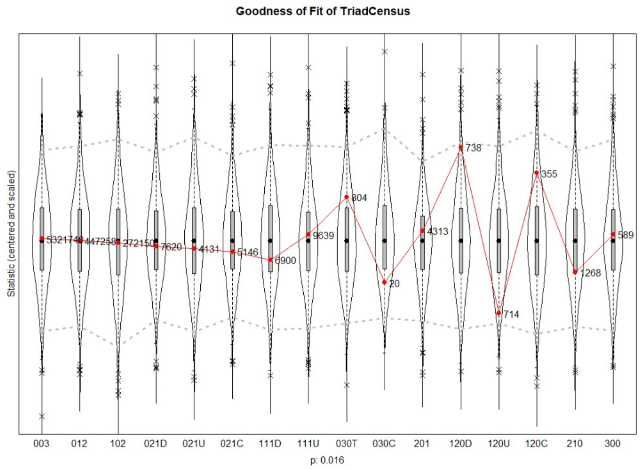


Table 5: Time Heterogeneity Tests for Model 1 (Effect Test)

	Paris 1			Paris 2			Savoie 1			Savoie 2		
	chi-sq.	df	p-value	chi-sq.	df	p-value	chi-sq.	df	p-value	chi-sq.	df	p-value
outdegree (density)	0	2	1	0	2	1	0	2	1	0	2	1
reciprocity	0.21	2	0.9	4.2	2	0.12	12.99	2	0	1.31	2	0.52
GWESP I -> K -> J	0.41	2	0.82	0.08	2	0.96	0.32	2	0.85	0.94	2	0.63
GWESP I <- K <- J	-	-	-	1.59	2	0.45	3.61	2	0.16	-	-	-
GWDSP I -> K -> J	-	-	-	1.16	2	0.56	-	-	-	-	-	-
indegree - popularity	0.05	2	0.98	0.12	2	0.94	0.7	2	0.71	1.19	2	0.55
outdegree - popularity	0.38	2	0.83	0.62	2	0.73	0.3	2	0.86	1.92	2	0.38
outdegree - activity	0.08	2	0.96	4.23	2	0.12	5.32	2	0.07	1.77	2	0.41
same ethnicity	1.38	2	0.5	0.71	2	0.7	1.1	2	0.58	10.54	2	0.01
s_mena_ethn5	3.8	2	0.15	5.62	2	0.06	1.85	2	0.4	0.6	2	0.74
s_europe_ethn5	5.45	2	0.07	25.3	2	0	2.1	2	0.35	19.31	2	0
s_subsah_ethn5	2.69	2	0.26	7.55	2	0.02	1.51	2	0.47	0.13	2	0.94
s_other_ethn5	3.79	2	0.15	13.2	2	0	0.66	2	0.72	0.74	2	0.69
r_mena_ethn5	2.34	2	0.31	1.52	2	0.47	3.02	2	0.22	0.08	2	0.96
r_europe_ethn5	0.03	2	0.99	0.88	2	0.64	7.73	2	0.02	9.88	2	0.01
r_subsah_ethn5	2.09	2	0.35	8.17	2	0.02	0.03	2	0.99	0.56	2	0.76
r_other_ethn5	3.02	2	0.22	1.63	2	0.44	1.14	2	0.57	1.07	2	0.59
walk_regular	0.29	2	0.87	5.33	2	0.07	18.72	2	0	7.05	2	0.03
walk_squared	0.38	2	0.83	0.56	2	0.76	27.29	2	0	6.83	2	0.03
sas alter	0.05	2	0.98	4.9	2	0.09	4.74	2	0.09	1.12	2	0.57
sas ego	0.81	2	0.67	2.5	2	0.29	7	2	0.03	4.28	2	0.12
sas similarity	0.41	2	0.82	1.14	2	0.57	2.32	2	0.31	6.35	2	0.04
sex alter	0.72	2	0.7	29.67	2	0	0.28	2	0.87	4.87	2	0.09
sex ego	6.46	2	0.04	87.68	2	0	2.58	2	0.28	3.27	2	0.2
sex similarity	2.43	2	0.3	2.94	2	0.23	10.75	2	0.01	4.52	2	0.1
same primary1	0.02	2	0.99	1.88	2	0.39	0.57	2	0.75	0.62	2	0.73
same special track	0.16	2	0.92	-	-	-	0.23	2	0.89	-	-	-
same regular track	1.95	2	0.38	-	-	-	0.63	2	0.73	-	-	-
special track ego	1.04	2	0.6	-	-	-	1	2	0.61	-	-	-
same classroom	0.06	2	0.97	9.5	2	0.01	2.08	2	0.35	1.56	2	0.46
grades alter	0.53	2	0.77	2.77	2	0.25	3.74	2	0.15	2.07	2	0.36
grades ego	0.73	2	0.69	18.64	2	0	0.58	2	0.75	0.56	2	0.76
grades similarity	0.17	2	0.92	10.91	2	0	6.13	2	0.05	4.18	2	0.12
int. reciprocity x GWESP I -> K -> J	0.5	2	0.78	3.8	2	0.15	8.92	2	0.01	0.61	2	0.74

Joint Significance Test of Time Heterogeneity:

Paris 1: chi-squared = 95.69, d.f. = 64, p= 0.0063

Paris 2: chi-squared = 312.70, d.f. = 62, p < 0.0001

Savoie 1: chi-squared = 205.57, d.f. = 66, $p < 0.0001$
Savoie 2: chi-squared = 122.46, d.f. = 58, $p < 0.0001$

NB: note that the joint test is statistically significant for all four schools, meaning that there is still time heterogeneity in the models (the probability that *all* the interaction dummies are 0 is extremely small). However, to keep the models relatively parsimonious, I only relied on the p-value specific to each effect (i.e. those of Table 5), that I tried to keep above the 0.05 threshold. Moreover, for a few effects, I did not add time dummies even though the p-value is under 0.05, because these effects are not very important theoretically, and they have very little impact on the structure of the networks overall (this is particularly the case of certain emission and reception terms for ethnic categories, which only concern part of the student population).

For the interpretation of time heterogeneity tests, see Lospinoso et al. (2011) and Ripley and Snijders (2020).

Appendix 5C: SAOM Estimation for the Networks of 5-Friend Nominations in Paris 2

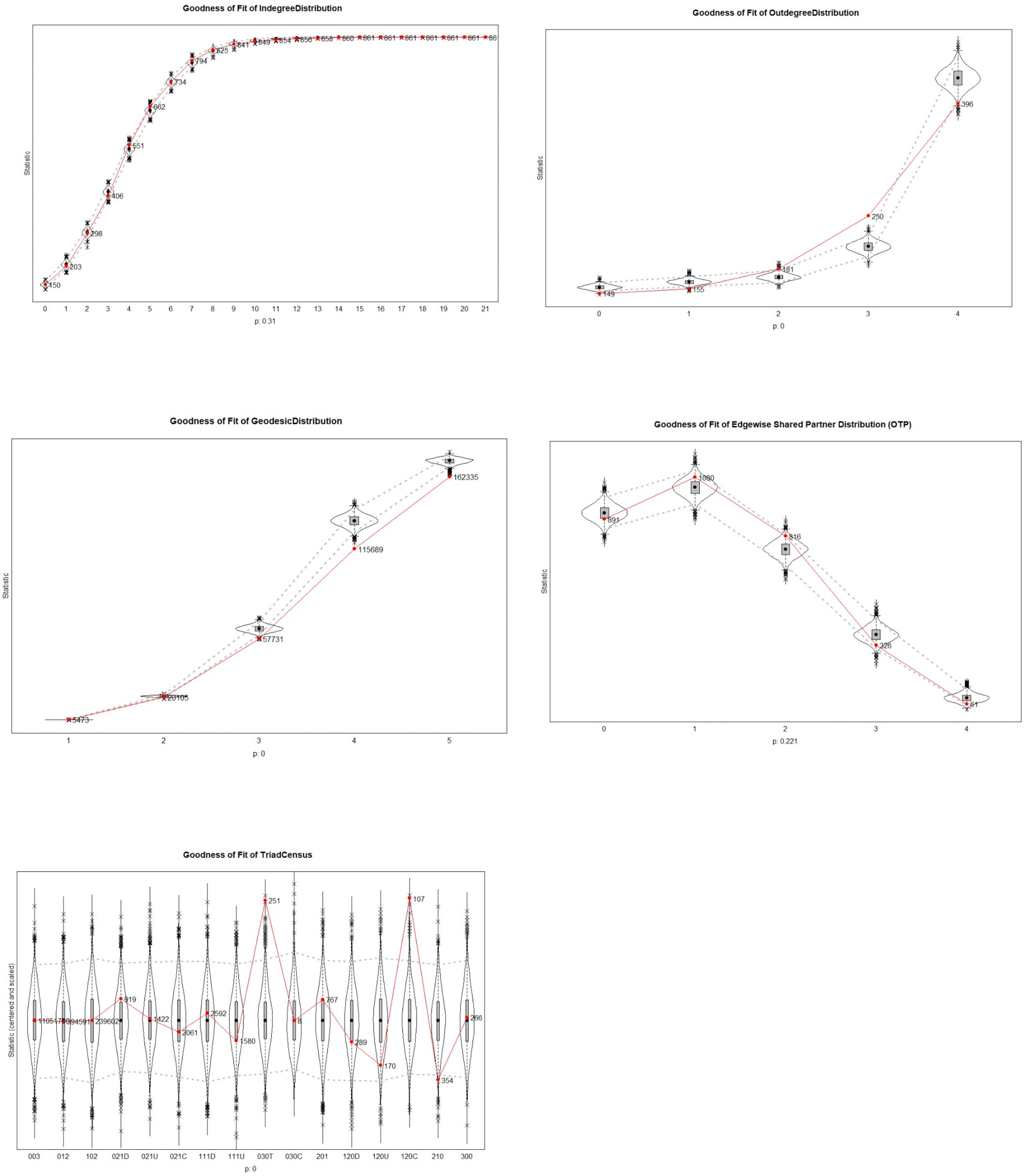
Table 1: Estimated SAOM Parameters for Paris 2 (5-Friend Networks, Waves 1 to 4)

	Estimate	Standard Error	Convergence T-Ratio
1. rate constant friendship rate (period 1)	28.39	(7.95)	-0.02
2. rate constant friendship rate (period 2)	33.74	(7.82)	0.02
3. rate constant friendship rate (period 3)	28.56	(4.2)	0
4. eval outdegree (density)	-5.38	(0.65)	0.01
5. eval reciprocity	2.33	(0.13)	0.02
6. eval GWESP I -> K -> J (69)	1.94	(0.11)	0.02
7. eval GWESP I <- K <- J (300)	-0.01	(0.05)	0.02
8. eval GWDSP I -> K -> J (69)	-0.01	(0.07)	0
9. eval indegree - popularity (sqrt)	-0.19	(0.05)	0
10. eval outdegree - popularity (sqrt)	-0.18	(0.18)	0.01
11. eval outdegree - activity	0.66	(0.11)	0.01
12. eval m_3main_ethn5	0.28	(0.04)	0.04
13. eval s_mena_ethn5	0.37	(0.25)	0.01
14. eval s_europe_ethn5	1.48	(1.43)	0.04
15. eval s_subsah_ethn5	1.27	(1.22)	0
16. eval s_other_ethn5	0.22	(0.34)	-0.01
17. eval r_mena_ethn5	-0.06	(0.06)	0.03
18. eval r_europe_ethn5	0.18	(0.08)	-0.03
19. eval r_subsah_ethn5	0.27	(0.07)	0.01
20. eval r_other_ethn5	0.04	(0.08)	-0.02
21. eval walk_rg	-0.27	(0.1)	0
22. eval walk_sq	0.12	(0.06)	0
23. eval sas alter	0	(0.02)	0.02
24. eval sas ego	0	(0.09)	0
25. eval sas similarity	0.19	(0.1)	0
26. eval sex alter	0.03	(0.05)	-0.04
27. eval sex ego	-0.08	(0.26)	-0.01
28. eval sex similarity	0.69	(0.05)	0.01
29. eval same primary1	0.18	(0.05)	0.01
30. eval grades alter	0.01	(0.02)	-0.02
31. eval grades ego	-0.24	(0.13)	0
32. eval grades similarity	0.27	(0.16)	-0.03
33. eval same classroom1	-0.15	(0.05)	0
34. eval same classroom2	1.24	(0.05)	0.02

35. eval timedummy_period2 ego	-0.31	(0.35)	0.01
36. eval timedummy_period3 ego	-0.11	(0.47)	-0.02
37. eval int. reciprocity x GWESP I -> K -> J (69)	-0.91	(0.12)	0.01
38. eval timedummy_period3 ego x sas similarity	0.49	(0.22)	0
39. eval timedummy_period2 ego x same classroom1	-0.29	(0.08)	0.02
Overall Maximum Convergence Ratio	*0.12*		
Total Number of Iteration Steps	*7913*		
Degrees Constrained to Maximum Values	*Friendship = 5*		

Note: “s_” and “r_” indicate, respectively, sender and receiver effects. These are essentially equivalent to “ego” and “alter” effects, except that they have been passed through SIENA as a dyadic covariate (effect “X”) rather than applying the ego and alter effect to an individual attribute. “Ethn5” refer to the five-posts ethnicity attribute. For the homophily effect, only the three main ethnic groups (natives, Middle-East and Subsaharian Africa) are counted. “walk_rg” is the linear walking distance between students’ homes, and “walk_sq” the squared distance. All other parameters have the effect names given by the Rsiena software.

Figure 1: Goodness-of-Fit Plots for the SAOM of Table 1 (Paris 2, 5-Friend Networks)



Appendix 5D: SAOM Estimation for the Periods Between Waves 2, 4 and 6

Table 1: Estimated SAOM Parameters (Very Good Friend Networks, Waves 2, 4 and 6)

	Paris 1		Paris 2		Savoie 1		Savoie 2	
	Estimate (Standard Error)	Convergence T-Ratio	Estimate (Standard Error)	Convergence T-Ratio	Estimate (Standard Error)	Convergence T-Ratio	Estimate (Standard Error)	Convergence T-Ratio
1. rate constant friendship rate (period 1)	23.44 (2.040)	0.03	48.85 (1.5916)	-0.02	33.44 (1.8686)	-0.03	35.05 (1.7739)	0
2. rate constant friendship rate (period 2)	14.66 (1.5925)	-0.03	37.95 (1.5344)	0.01	26.61 (1.6549)	-0.03	28.74 (1.5358)	-0.01
3. eval outdegree (density)	-2.82 (0.1885)	0.04	-3.31 (0.0895)	-0.02	-2.78 (0.0707)	0.03	-2.91 (0.0728)	0
4. eval reciprocity	2.52 (0.2339)	0.06	3.52 (0.1658)	-0.02	2.64 (0.0985)	0.02	2.73 (0.1096)	0.01
5. eval GWESP I -> K -> J (110 / 69 / 125 / 85)	1.23 (0.0858)	0.04	1.62 (0.0718)	-0.03	1.03 (0.0652)	0.05	1.38 (0.0486)	-0.01
6. eval GWESP I <- K <- J (. / 300 / 125 / .)	-	-	0.15 (0.0281)	-0.05	0.17 (0.0648)	0.03	-	-
7. eval GWDSP I -> K -> J (. / 69 / . / .)	-	-	0.01 (0.0079)	-0.01	-	-	-	-
6. eval indegree - popularity	-0.04 (0.0169)	0.04	0.07 (0.0349)	-0.04	-0.01 (0.0062)	0.07	0 (0.0061)	-0.01
7. eval outdegree - popularity	-0.1 (0.0142)	0.03	-0.46 (0.0781)	-0.03	-0.07 (0.0086)	0.04	-0.08 (0.0068)	0
8. eval outdegree - activity	0 (0.0047)	0.02	0 (0.0011)	-0.05	0 (0.0023)	0.04	0 (0.0019)	-0.01
9. eval same ethnicity	0.17 (0.0698)	-0.03	0.24 (0.0364)	-0.09	0.17 (0.0546)	-0.01	0.2 (0.0604)	0.02
12. eval s_mena_ethn5	-0.05 (0.1064)	-0.01	0.11 (0.0516)	-0.04	0.04 (0.0713)	0.03	0.05 (0.0719)	0.01
13. eval s_europe_ethn5	0.36 (0.1477)	0.01	0.16 (0.0761)	-0.01	0.22 (0.0602)	0.01	0.15 (0.0622)	0
14. eval s_subsah_ethn5	0.09 (0.1160)	-0.05	0.02 (0.0566)	0.01	-0.35 (0.1582)	0.07	0.18 (0.1440)	-0.02
15. eval s_other_ethn5	0.73 (0.1962)	0.04	0.63 (0.0801)	0.09	0.14 (0.1317)	-0.01	0.35 (0.1236)	-0.02
16. eval r_mena_ethn5	0.11 (0.0967)	0.01	-0.06 (0.0477)	-0.07	0.18 (0.0618)	0.01	0.1 (0.0671)	0
17. eval r_europe_ethn5	0.15 (0.1217)	0	0.02 (0.0705)	0	0.19 (0.0568)	-0.01	0.29 (0.0598)	0
18. eval r_subsah_ethn5	0.41 (0.1278)	-0.03	0.14 (0.0588)	0	0.19 (0.1255)	0.06	0.08 (0.1281)	-0.01
19. eval r_other_ethn5	0.01 (0.2128)	0	0.17 (0.0798)	0.1	0.1 (0.1437)	0.01	0.21 (0.1144)	-0.01
20. eval walk_rg	0.37 (0.3489)	0	-0.03 (0.0674)	0.04	-0.02 (0.0263)	-0.01	-0.06 (0.0320)	0.01
21. eval walk_sq	-0.15 (0.2602)	0.01	0.01 (0.0280)	0.03	0 (0.0039)	0.04	0 (0.0052)	0.01
22. eval s_4B_classroom	0.32 (0.1278)	-0.01	-	-	-	-	-	-
23. eval s_4C_classroom	0.57 (0.1356)	0.01	-	-	-	-	-	-
24. eval r_4B_classroom	0.17 (0.1214)	0	-	-	-	-	-	-
25. eval r_4C_classroom	0.21 (0.1343)	0.02	-	-	-	-	-	-
26. eval s_3B_classroom	-0.26 (0.1318)	-0.02	-	-	-	-	-	-

27. eval s_3C_classroom	0.08 (0.1455)	-0.01	-	-	-	-	-	-
28. eval r_3B_classroom	-0.11 (0.1310)	-0.03	-	-	-	-	-	-
29. eval r_3C_classroom	0.11 (0.1461)	-0.01	-	-	-	-	-	-
30. eval sas alter	0.18 (0.0479)	0.01	0 (0.0144)	-0.07	-0.01 (0.0216)	-0.01	0.03 (0.0192)	-0.03
31. eval sas ego	-0.06 (0.0489)	0.01	-0.01 (0.0180)	0.03	-0.03 (0.0244)	0	-0.09 (0.0203)	-0.02
32. eval sas similarity	0.9 (0.1805)	0.03	0.06 (0.0672)	-0.02	0.14 (0.0965)	-0.03	0.09 (0.0903)	0.02
33. eval sex alter	-0.05 (0.0698)	0.01	0.04 (0.0313)	-0.01	-0.01 (0.0406)	0.01	-0.03 (0.0368)	-0.03
34. eval sex ego	-0.01 (0.0764)	0.01	-0.14 (0.0353)	0.02	0.07 (0.0452)	0.02	0.03 (0.0401)	-0.02
35. eval sex similarity	0.41 (0.0654)	0.03	0.22 (0.0268)	0.01	0.16 (0.0393)	0.01	0.24 (0.0343)	-0.01
36. eval same primary1	0.05 (0.0736)	-0.04	0.06 (0.0351)	-0.04	0.14 (0.0576)	0.02	0.07 (0.0450)	-0.01
10. eval same special track	-0.61 (0.1879)	0.03	-	-	-0.12 (0.2378)	-0.05	-	-
11. eval same regular track	0.7 (0.1740)	-0.02	-	-	0.26 (0.1595)	0.05	-	-
37. eval special track ego	0.33 (0.2230)	0.04	-	-	0.1 (0.2822)	-0.05	-	-
38. eval same classroom1	0.45 (0.0726)	0.05	-0.04 (0.0166)	-0.06	0.29 (0.0361)	-0.01	0.37 (0.0323)	0.02
39. eval same classroom2	0.39 (0.0868)	0.05	-0.07 (0.0176)	-0.01	0.6 (0.0400)	-0.03	0.75 (0.0365)	0
40. eval grades alter	-0.08 (0.0438)	0.03	0.32 (0.0887)	0.04	0 (0.0218)	0.03	-0.01 (0.0188)	0
41. eval grades ego	-0.11 (0.0497)	0.06	0.23 (0.0307)	-0.02	-0.08 (0.0229)	-0.04	-0.08 (0.0192)	-0.01
42. eval grades similarity	0.39 (0.1942)	0.06	0.64 (0.0307)	-0.05	0.57 (0.1394)	-0.04	0.01 (0.1110)	0.01
43. eval timedummy_period2 ego	0.37 (0.1425)	0	0.1 (0.0526)	0.03	-0.14 (0.0363)	-0.02	-0.04 (0.0395)	-0.01
44. eval int. reciprocity x GWESP I -> K -> J (110)	-0.46 (0.1531)	0.06	-1.32 (0.1063)	-0.02	-0.6 (0.0598)	0.03	-0.62 (0.0941)	0
45. eval int. timedummy_period2 ego x sas ego	-0.24 (0.0720)	-0.01	-	-	0.15 (0.0448)	-0.03	-0.26 (0.0653)	-0.02
36. eval timedummy_period2 ego x outdegree - activity	-	-	-0.01 (0.0012)	0.05	-	-	-	-
38. eval timedummy_period2 ego x sas similarity	-	-	-	-	0.47 (0.1901)	0.02	-	-
Overall Maximum Convergence Ratio	*0.19*		*0.25*		*0.24*		*0.08*	
Total Nb of Iteration Steps	*2066*		*2272*		*2009*		*7309*	

Note: "s_" and "r_" indicate, respectively, sender and receiver effects. These are equivalent to "ego" and "alter" effects, except that they have been passed to SIENA as a dyadic covariate (effect "X"). "Ethn5" refer to the five-posts ethnicity attribute. For the homophily effect, only the three main ethnic groups are counted. "walk_rg" is the linear walking distance between students' homes, and "walk_sq" the squared distance.

Figure 1: Goodness-of-Fit Plots for the SAOM of Table 1 – Paris 1, Very Good Friend Networks, Waves 2,4 and 6

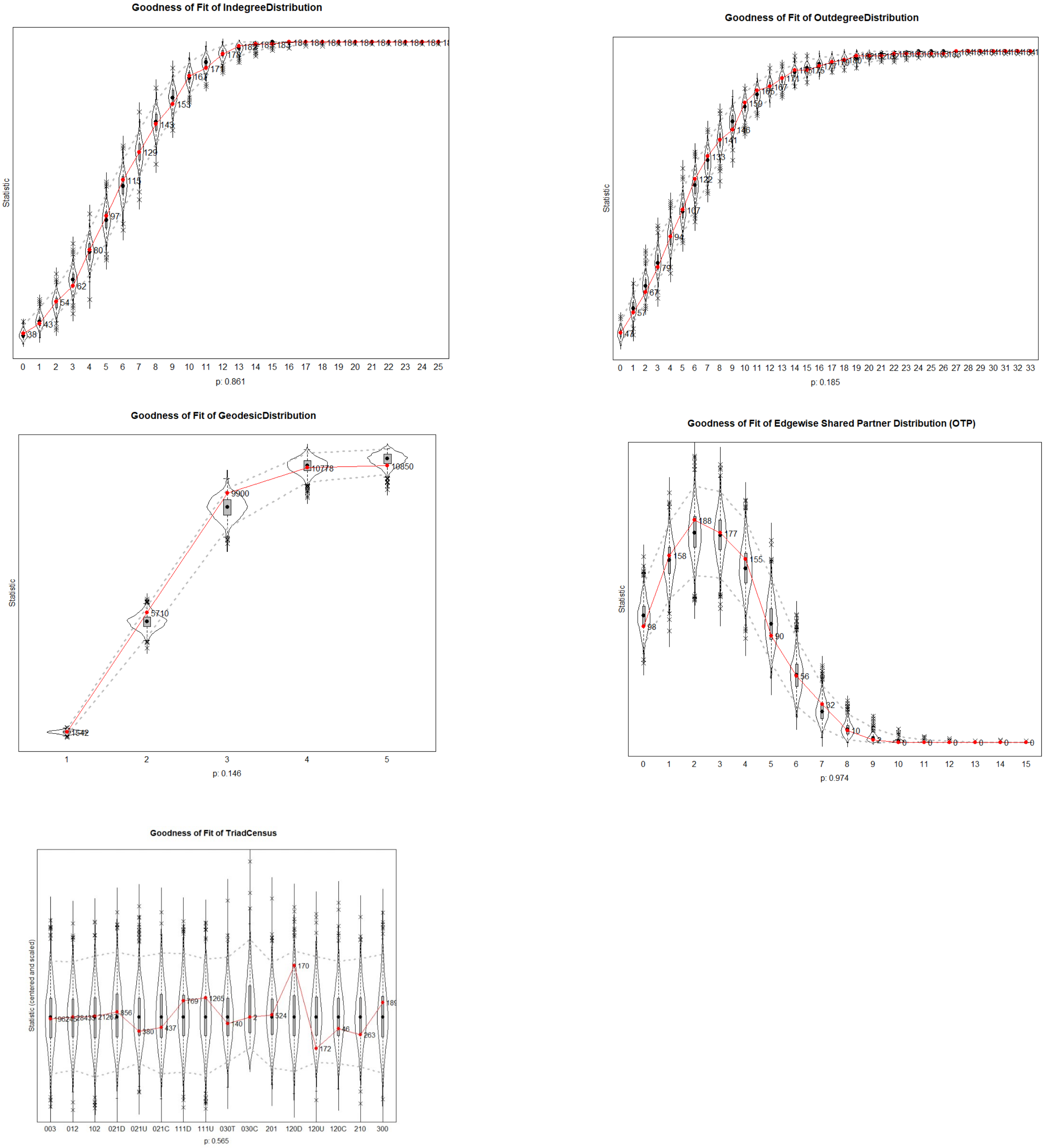


Figure 2: Goodness-of-Fit Plots for the SAOM of Table 1 – Paris 2, Very Good Friend Networks, Waves 2,4 and 6

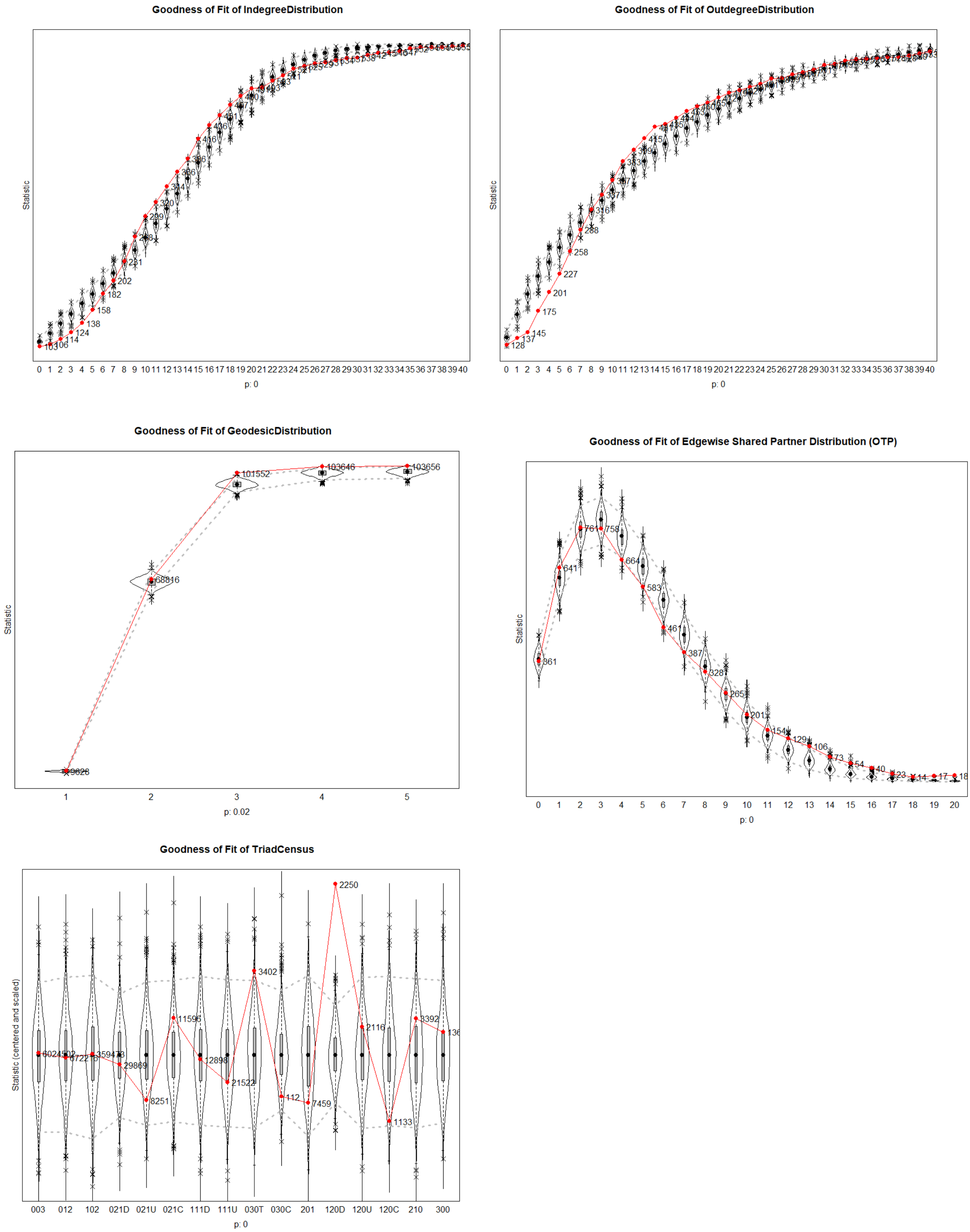


Figure 3: Goodness-of-Fit Plots for the SAOM of Table 1 – Savoie 1, Very Good Friend Networks, Waves 2,4 and 6

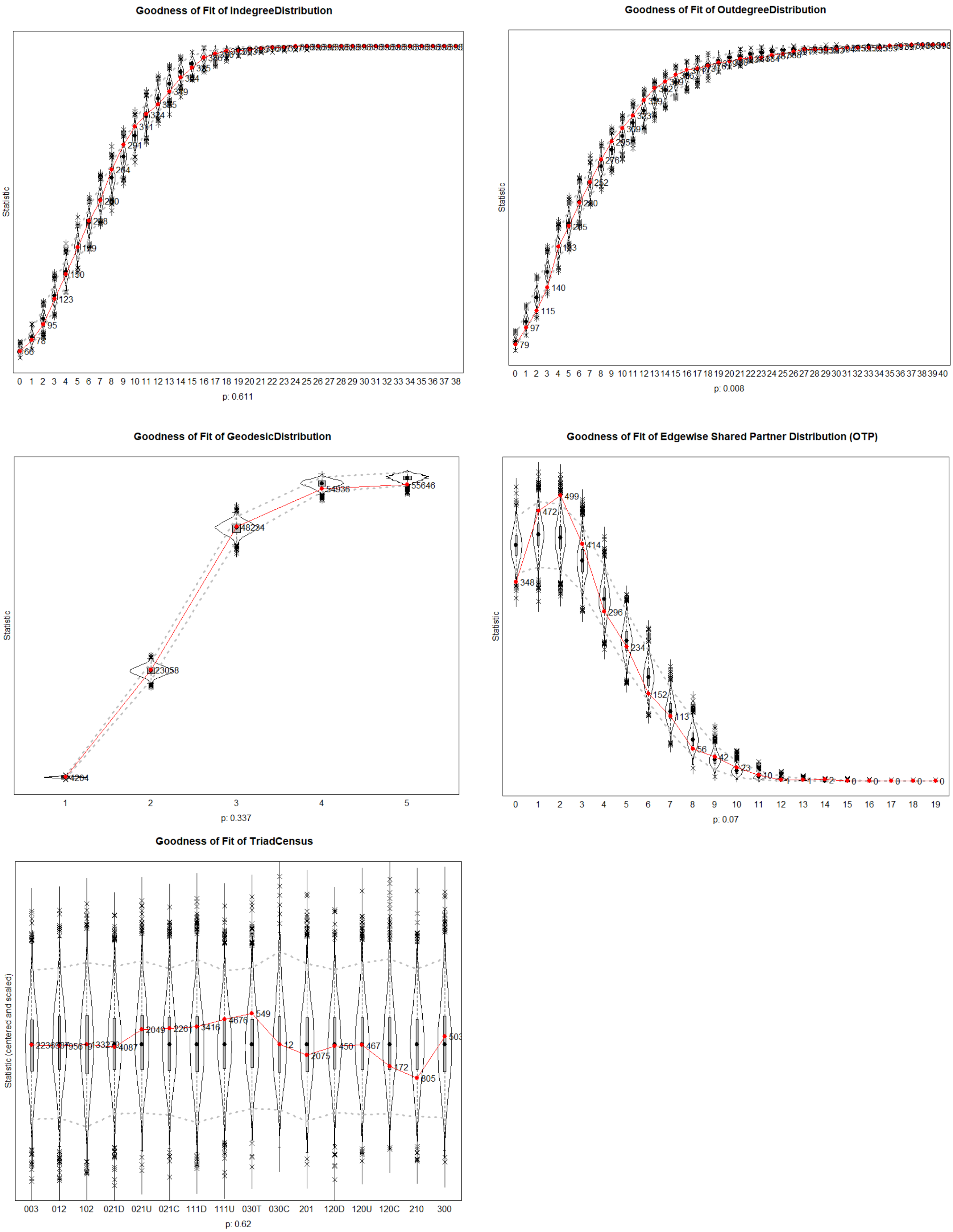


Figure 4: Goodness-of-Fit Plots for the SAOM of Table 1 – Savoie 2, Very Good Friend Networks, Waves 2,4 and 6

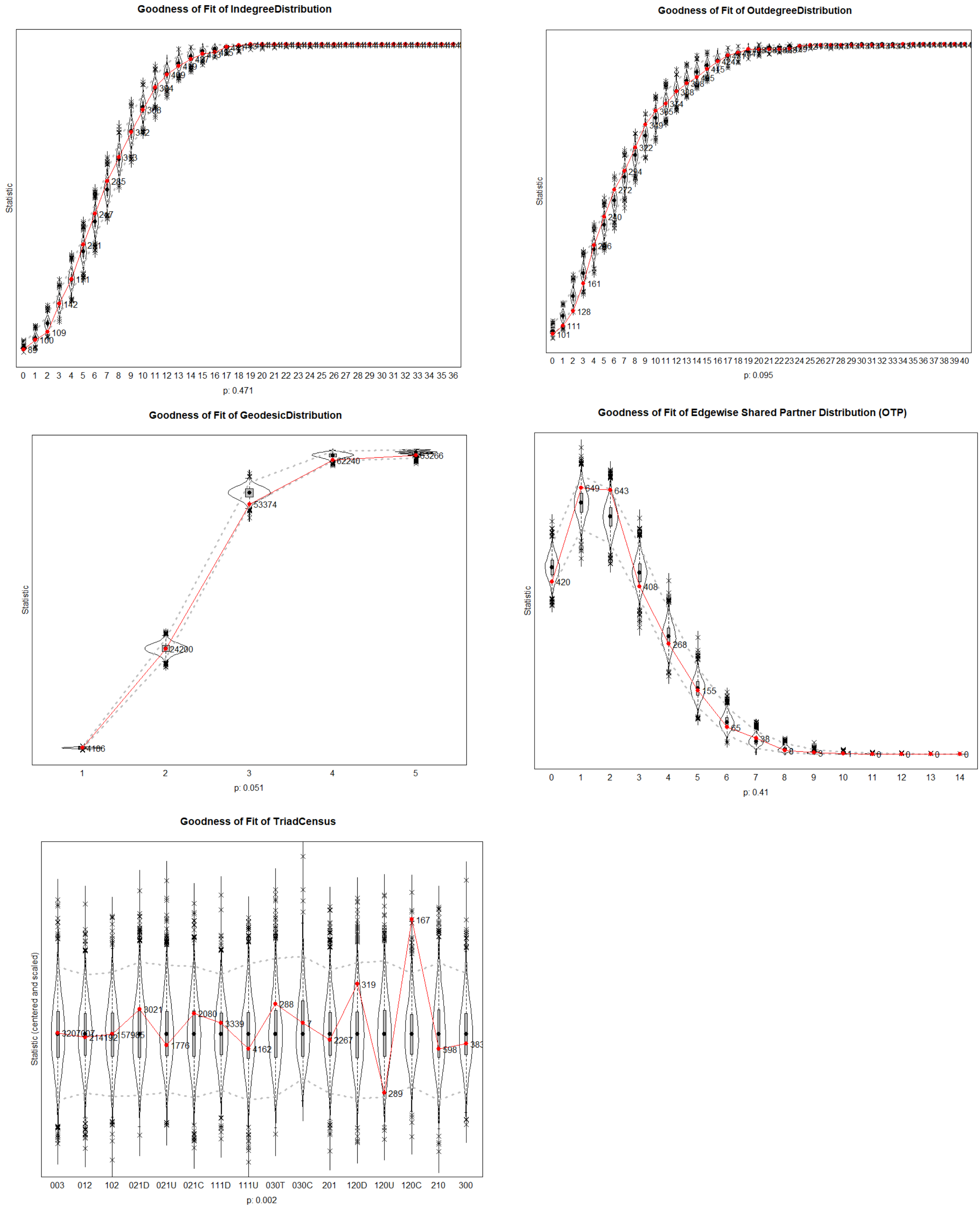


Table 2: Time Heterogeneity Tests for the SAOM of Table 1 (Very Good Friend Networks, Waves 2, 4 and 6) (Effect Test)

	Paris 1			Paris 2			Savoie 1			Savoie 2		
	chi-sq.	df	p-value	chi-sq.	df	p-value	chi-sq.	df	p-value	chi-sq.	df	p-value
reciprocity	0.24	1	0.62	0.27	1	0.6	0.16	1	0.69	0.26	1	0.61
GWESP I -> K -> J	0.04	1	0.84	0.45	1	0.5	1.28	1	0.26	0.44	1	0.51
GWESP I <- K <- J				4.93	1	0.03	1.71	1	0.19			
GWDSP I -> K -> J				0.65	1	0.42						
indegree - popularity	0.19	1	0.66	0.43	1	0.51	0.76	1	0.38	0.01	1	0.92
outdegree - popularity	0.15	1	0.7	0.39	1	0.53	1.84	1	0.18	0.02	1	0.89
outdegree - activity	0.11	1	0.74	0	1	1	1.1	1	0.29	0	1	1
same ethnicity	0.29	1	0.59	0.87	1	0.35	0.81	1	0.37	0.01	1	0.92
s_mena_ethn5	0.82	1	0.37	0.13	1	0.72	0.06	1	0.81	5.88	1	0.02
s_europe_ethn5	0.2	1	0.66	0.54	1	0.46	12.84	1	0	3.41	1	0.07
s_subsah_ethn5	0.2	1	0.66	0.78	1	0.38	0.37	1	0.54	1.34	1	0.25
s_other_ethn5	0.36	1	0.55	5.81	1	0.02	1.12	1	0.29	1.33	1	0.25
r_mena_ethn5	0.15	1	0.7	0.07	1	0.79	1.72	1	0.19	0.14	1	0.71
r_europe_ethn5	1.59	1	0.21	1.48	1	0.22	0.24	1	0.62	0.68	1	0.41
r_subsah_ethn5	1.61	1	0.2	0.03	1	0.86	0.92	1	0.34	0.43	1	0.51
r_other_ethn5	0.01	1	0.92	1.04	1	0.31	5.25	1	0.02	7.03	1	0.01
walk_regular	0.78	1	0.38	0.52	1	0.47	0	1	1	0.77	1	0.38
walk_squared	0.21	1	0.65	0.6	1	0.44	0	1	1	0.01	1	0.92
sas alter	0.86	1	0.35	0.58	1	0.45	1.43	1	0.23	2.47	1	0.12
sas ego	0	1	1	0.3	1	0.58	0	1	1	0.43	1	0.51
sas similarity	0.09	1	0.76	0.05	1	0.82	0	1	1	0.04	1	0.84
sex alter	0.17	1	0.68	0.13	1	0.72	1.74	1	0.19	0.34	1	0.56
sex ego	0.84	1	0.36	0.1	1	0.75	0.17	1	0.68	2.72	1	0.1
sex similarity	1.36	1	0.24	0.85	1	0.36	1.15	1	0.28	0.14	1	0.71
same primary1	0.03	1	0.86	1.27	1	0.26	1.53	1	0.22	1.78	1	0.18
same special track	2.41	1	0.12	-	-	-	0.08	1	0.78			
same regular track	0	1	1	-	-	-	2.38	1	0.12			
special track ego	0.18	1	0.67	-	-	-	1.14	1	0.29			
same classroom	0.02	1	0.89	2.99	1	0.08	2.23	1	0.14	0	1	1
grades alter	0.32	1	0.57	0.96	1	0.33	2.49	1	0.12	0.12	1	0.73
grades ego	0.53	1	0.47	3.29	1	0.07	2.93	1	0.09	0.45	1	0.5
grades similarity	0.7	1	0.4	3.99	1	0.05	0.33	1	0.57	3.4	1	0.07
int. reciprocity x GWESP I -> K -> J	0	1	1	1.22	1	0.27	0.55	1	0.46	0.8	1	0.37

Joint Significance Test of Time Heterogeneity:

Paris 1: chi-squared = 39.24, d.f. = 31, p= 0.1471

Paris 2: chi-squared = 73.30, d.f. = 30, p < 0.0001

Savoie 1: chi-squared = 63.06, d.f. = 32, p = 0.0009
Savoie 2: chi-squared = 49.56, d.f. = 28, p = 0.0072

Appendix 5E: ERGM Estimation for Wave 1

Table 1: Estimated ERGM Parameters (Very Good Friend Networks, Wave 1)

	Paris 1	Paris 2	Savoie 1	Savoie 2
edges	constrained	constrained	constrained	constrained
mutual	2.147*** (0.218)	2.074*** (0.066)	2.275*** (0.151)	2.686*** (0.117)
gwesp.OTP.fixed.0.69	1.060*** (0.104)	2.002*** (0.045)	1.240*** (0.059)	1.206*** (0.046)
gwesp.ITP.fixed.0.69	-0.159** (0.062)	-0.305*** (0.020)	-0.283*** (0.040)	-0.268*** (0.033)
gwesp.OSP.fixed.0.69	-	0.004 (0.028)	-	-
gwdsp.OTP.fixed.0.69	-0.104*** (0.013)	-	-0.062*** (0.008)	-0.061*** (0.005)
gwodeg.fixed.1.5	0.026 (0.277)	-	-0.115 (0.151)	-0.257** (0.120)
gwideg.fixed.1.5	0.981*** (0.340)	-	0.815*** (0.214)	0.773*** (0.168)
gwodeg.fixed.0.69	-	0.813*** (0.234)	-	-
gwideg.fixed.0.69	-	2.535*** (0.630)	-	-
nodematch.sex	0.477*** (0.054)	0.692*** (0.015)	0.670*** (0.051)	0.477*** (0.031)
nodeifactor.sex.M	0.192** (0.097)	-0.044* (0.025)	0.017 (0.073)	0.069 (0.050)
nodeofactor.sex.M	-0.183** (0.087)	0.051** (0.022)	0.041 (0.062)	-0.054 (0.043)
nodematch.classroom	0.502*** (0.062)	0.942*** (0.017)	0.715*** (0.042)	0.642*** (0.027)
absdiff.sas	-0.083* (0.050)	-0.052*** (0.014)	-0.028 (0.033)	-0.060** (0.025)
nodeicov.sas	0.135** (0.064)	-0.016 (0.014)	-0.016 (0.035)	0.064** (0.026)
nodeocov.sas	0.149*** (0.053)	0.007 (0.012)	0.020 (0.028)	0.077*** (0.020)
absdiff.grade	-0.173*** (0.051)	-0.024*** (0.009)	-0.092*** (0.030)	-0.029 (0.021)
nodeicov.grade	-0.096 (0.064)	0.045*** (0.013)	0.068* (0.036)	0.024 (0.024)
nodeocov.grade	-0.173*** (0.055)	-0.069*** (0.012)	-0.083*** (0.029)	-0.028 (0.018)
nodematch.primary1	0.444*** (0.077)	0.286*** (0.022)	0.801*** (0.049)	0.367*** (0.030)
edgecov.walk_rg	-0.546** (0.221)	-0.023 (0.022)	-0.248*** (0.071)	-0.103*** (0.031)
edgecov.walk_sq	0.374*** (0.112)	0.006 (0.007)	0.068*** (0.021)	0.016*** (0.005)
nodematch.ethn5	0.284*** (0.081)	0.300*** (0.032)	0.323*** (0.072)	0.156* (0.094)
nodeofactor.ethn5.europe	-0.301 (0.193)	-0.072 (0.056)	0.244*** (0.074)	0.055 (0.083)
nodeofactor.ethn5.mena	0.181* (0.109)	0.005 (0.034)	0.221*** (0.074)	0.048 (0.088)
nodeofactor.ethn5.other	-0.003 (0.300)	0.241*** (0.053)	-0.262 (0.211)	-0.003 (0.156)
nodeofactor.ethn5.subsah	0.040 (0.129)	0.240*** (0.042)	0.186 (0.135)	0.053 (0.157)
nodeifactor.ethn5.europe	0.407** (0.192)	0.011 (0.061)	0.222*** (0.085)	0.253*** (0.094)
nodeifactor.ethn5.mena	0.105 (0.133)	-0.137*** (0.037)	0.084 (0.097)	0.123 (0.105)
nodeifactor.ethn5.other	0.071 (0.284)	-0.202*** (0.061)	0.043 (0.215)	-0.039 (0.190)
nodeifactor.ethn5.subsah	0.256 (0.159)	-0.097** (0.044)	0.596*** (0.162)	0.354** (0.175)
nodematch.special track.0	0.443** (0.201)	-	-	-
nodematch.special track.1	0.187 (0.234)	-	-	-

nodeofactor.special track.1	0.247 (0.313)	-	-	-
Akaike Inf. Crit.	*-1,365.165*	*-11,881.480*	*-3,519.905*	*-5,408.742*
Bayesian Inf. Crit.	*-1,159.070*	*-11,625.750*	*-3,294.752*	*-5,163.339*

Note: all models are estimated with a constraint on the number of edges.

Structural terms (gwesp, gwidegree, etc) were chosen to improve the model fit in each school. In Paris 2, there is no two-path term (gwdsp) because it is too demanding in terms of computing power (Paris 2's networks are both larger and denser, which makes the estimation harder there – this is also reflected in the fact that the fit of the model is less good in this school; see below).

"Ethn5" refer to the five-posts ethnicity attribute. For the homophily effect, only the three main ethnic groups (natives, Middle-East and Subsaharian Africa) are counted. "walk_rg" is the linear walking distance between students' homes, and "walk_sq" the squared distance. All other parameters have the effect names given by the statnet R package.

Figure 1: Goodness-of-Fit Plots for the ERGM of Table 1 – Paris 1

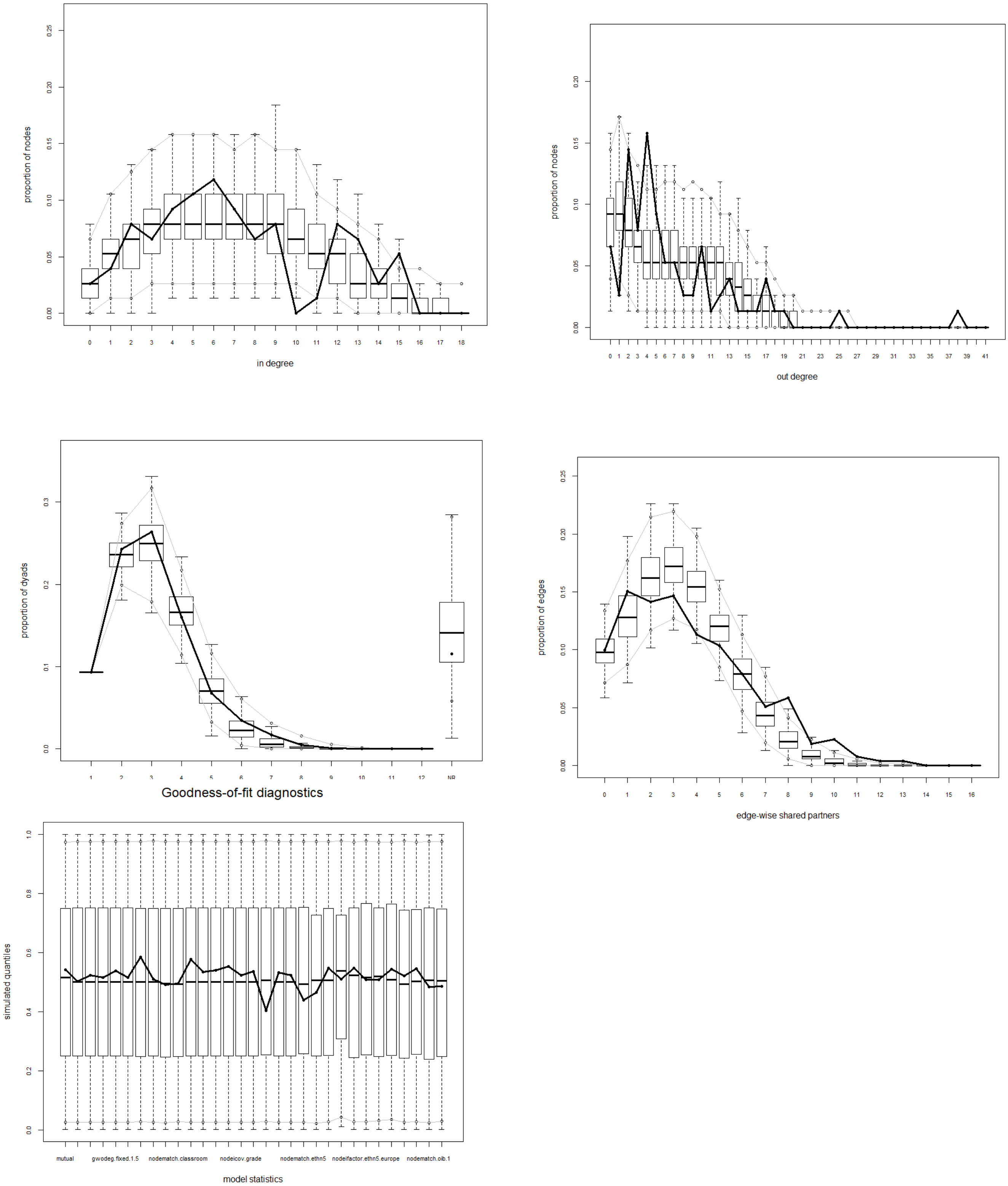


Figure 2: Goodness-of-Fit Plots for the ERGM of Table 1 – Paris 2

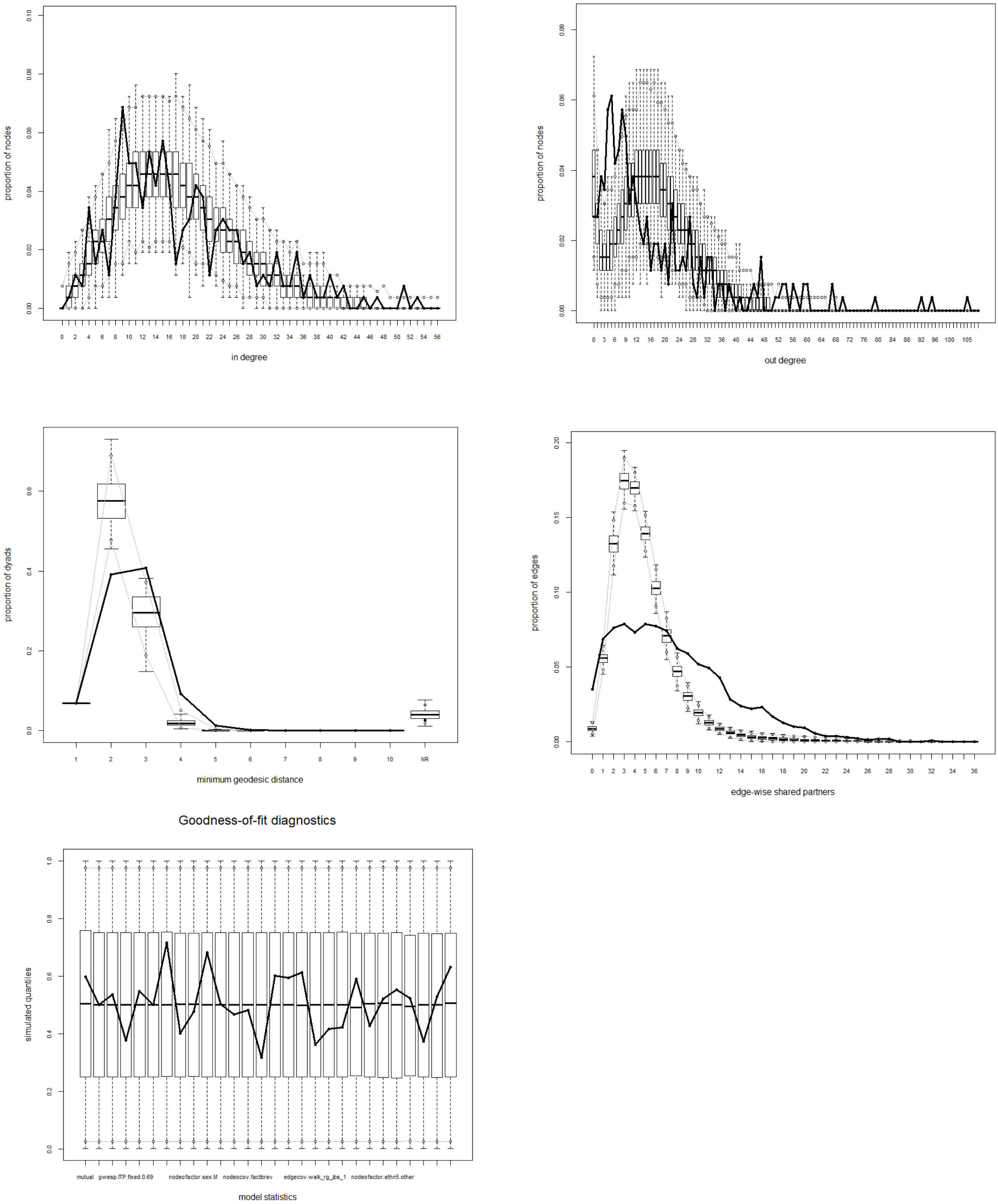


Figure 3: Goodness-of-Fit Plots for the ERGM of Table 1 – Savoie 1

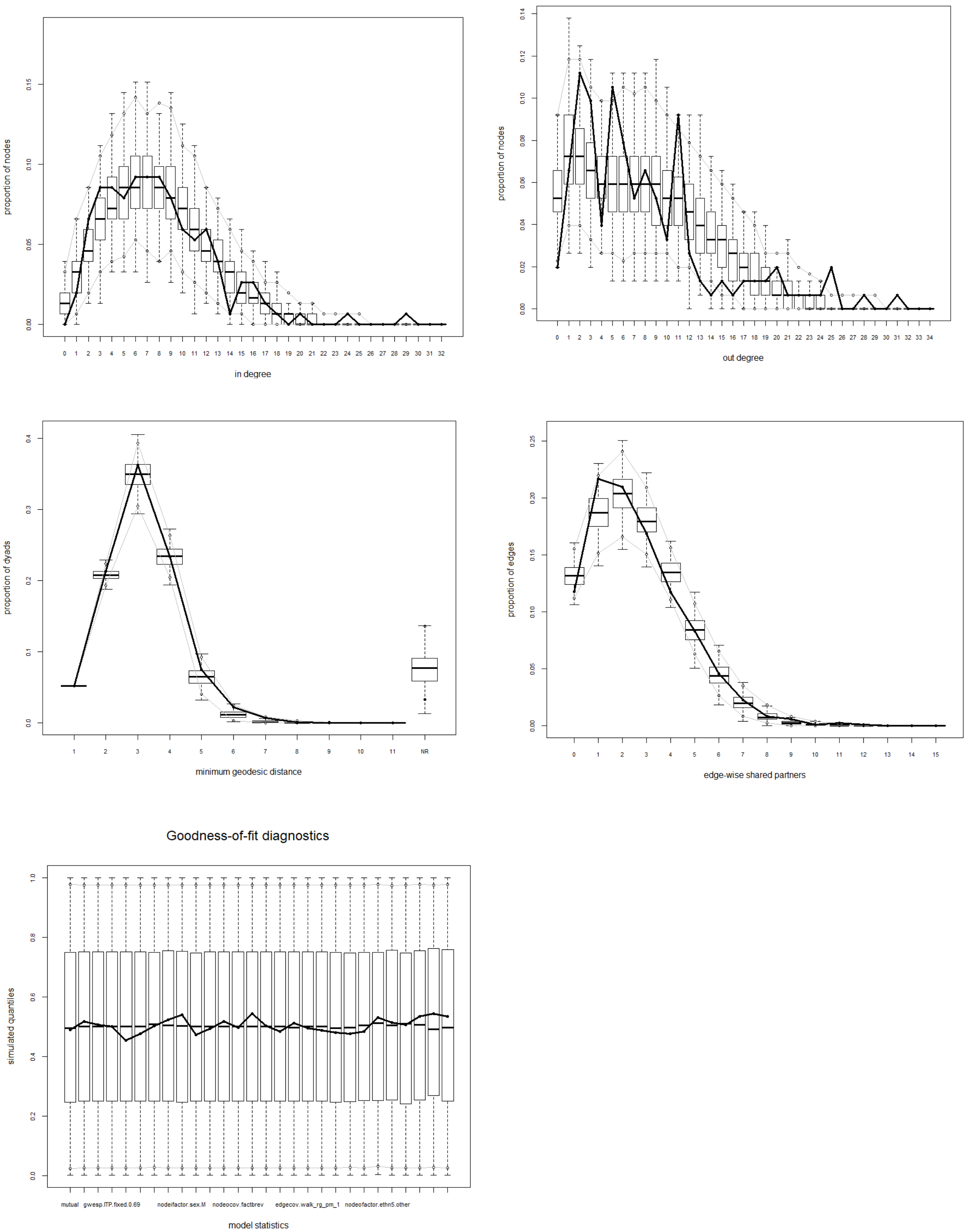
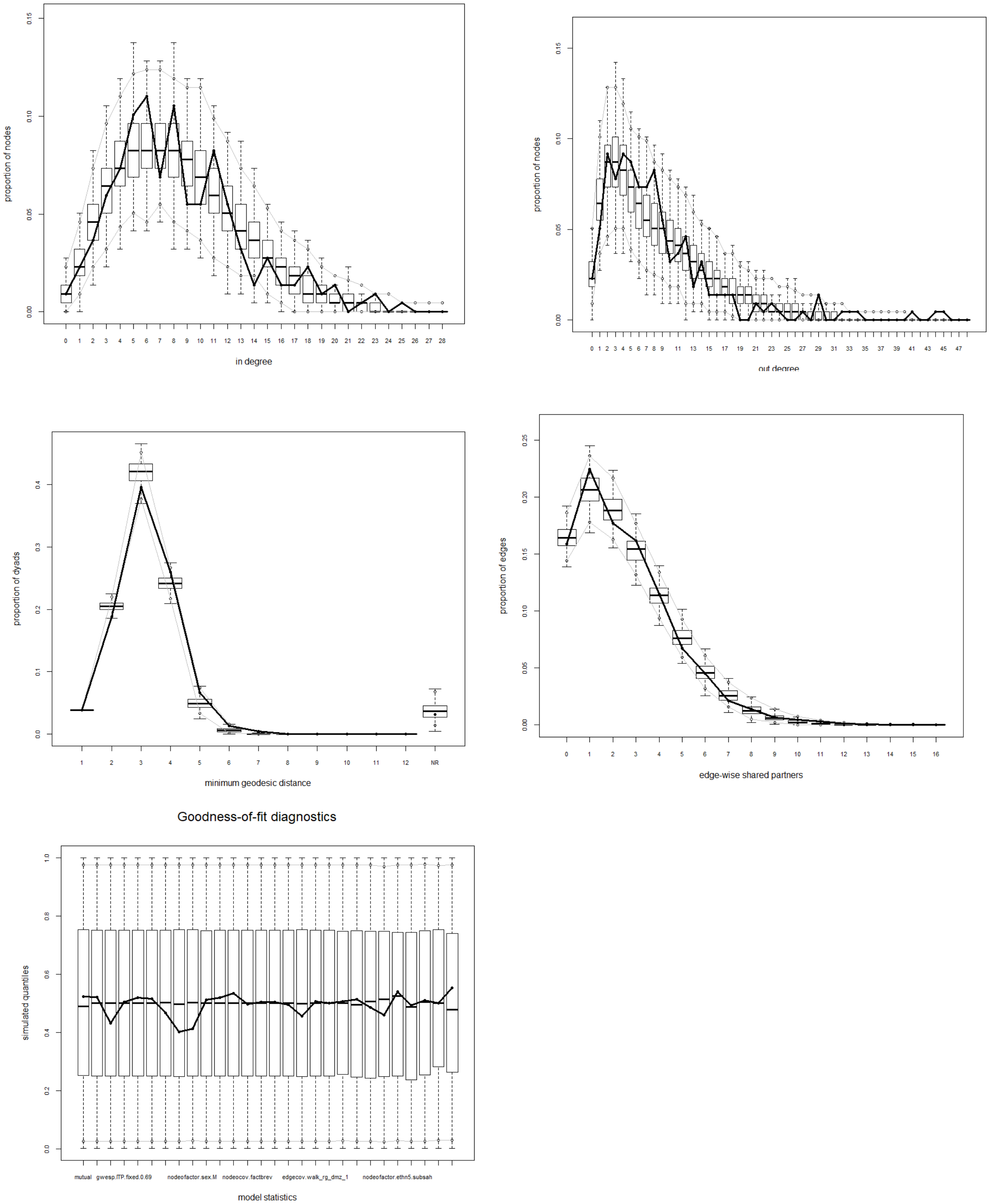


Figure 4: Goodness-of-Fit Plots for the ERGM of Table 1 – Savoie 2



Appendix 5F: Homophily of Cultural Status

Table 1: ERGM coefficients for Cultural Status Homophily, Emission and Reception – Very Good Friend Networks, per School and Wave (Dyad-Independent ERGM)

		Wave 1	Wave 2	Wave 3	Wave 4	Wave 6
Paris 1	Density	-1.066*** (0.131)	-1.011*** (0.120)	-1.301*** (0.105)	-1.850*** (0.115)	-1.578*** (0.125)
	Cultural Status Homophily	-0.674*** (0.069)	-0.724*** (0.063)	-0.803*** (0.061)	-0.874*** (0.064)	-0.840*** (0.074)
	Cultural Status Emission	-0.132** (0.054)	-0.083* (0.049)	0.073 (0.048)	0.164*** (0.050)	0.098* (0.058)
	Cultural Status Reception	-0.114** (0.054)	-0.053 (0.049)	-0.109** (0.048)	0.082 (0.050)	-0.053 (0.058)
Paris 2	Density	-1.461*** (0.051)	-1.606*** (0.055)	-1.611*** (0.047)	-1.694*** (0.051)	-2.152*** (0.064)
	Cultural Status Homophily	-0.265*** (0.021)	-0.221*** (0.022)	-0.265*** (0.021)	-0.339*** (0.023)	-0.329*** (0.028)
	Cultural Status Emission	-0.163*** (0.017)	-0.121*** (0.018)	-0.138*** (0.017)	-0.081*** (0.018)	-0.058** (0.023)
	Cultural Status Reception	-0.225*** (0.017)	-0.221*** (0.018)	-0.201*** (0.017)	-0.153*** (0.018)	-0.121*** (0.023)
Savoie 1	Density	-2.090*** (0.094)	-2.055*** (0.087)	-2.209*** (0.082)	-2.197*** (0.083)	-2.410*** (0.095)
	Cultural Status Homophily	-0.274*** (0.039)	-0.341*** (0.036)	-0.189*** (0.033)	-0.227*** (0.033)	-0.243*** (0.038)
	Cultural Status Emission	-0.146*** (0.032)	-0.051* (0.029)	-0.068** (0.027)	-0.098*** (0.027)	-0.064** (0.031)
	Cultural Status Reception	-0.093*** (0.032)	-0.103*** (0.029)	-0.151*** (0.027)	-0.092*** (0.027)	-0.079** (0.031)
Savoie 2	Density	-3.256*** (0.094)	-2.912*** (0.098)	-2.958*** (0.095)	-2.670*** (0.097)	-2.544*** (0.114)
	Cultural Status Homophily	-0.140*** (0.034)	-0.065* (0.035)	-0.120*** (0.036)	-0.173*** (0.038)	-0.243*** (0.046)
	Cultural Status Emission	-0.007 (0.028)	-0.026 (0.029)	-0.013 (0.030)	-0.012 (0.031)	-0.014 (0.038)
	Cultural Status Reception	0.082*** (0.028)	-0.008 (0.029)	-0.035 (0.030)	-0.093*** (0.031)	-0.152*** (0.038)

Appendix 5G: Contribution Scores for Ethnic Homophily and Gender Homophily

Contribution scores are computed exactly like for socioeconomic homophily, and based on the same estimated SAOM, except that the network-level statistic considered is the number of same-gender or same-ethnicity ties (in this later case, only considering the three largest ethnic groups, i.e. natives, Middle-East and Subsaharian Africa). Of particular interest is the case of Savoie 1 for ethnicity: there is almost no ethnic homophily in the evolution of the network after wave 1, which is the reason why contribution scores appear absurdly large (since they are expressed in proportion of the total homophily). However, we can see on Figure 1 that setting some parameters to 0 still change the predicted amount of the ethnic-matching statistic. In particular, setting the SAOM parameter for ethnic homophily to 0 results in the predictive networks exhibiting very high levels of ethnic *heterophily*. Therefore, it appears that, according to the SAOM, there is indeed some ethnic homophilic selection among students, but which gets counter-balanced by other relational processes, resulting in a null homophily at the aggregate level. This further implies that, in Savoie 1, most relational processes tend to induce ethnic *heterophily* in the network. To a lesser extent, there is a similar pattern for ethnic homophily in Paris 2 and gender homophily in Paris 1 (this is shown by contribution scores of the corresponding parameter being inferior to -1). Beyond these specific cases, we see that direct homophilic selection clearly has the largest contribution score across schools, both for ethnicity and gender. This sharply contrasts with the pattern found for socioeconomic homophily, where direct selection played a minor role and where a large portion of the total homophily was indirectly induced by other relational processes.

Table 1: Contribution Scores for Ethnic Homophily

			Predicted % change in ethnic homophily		
Model	**Scenario Name**	**Parameter(s) set to 0**	Paris 1	Paris 2	Savoie 1
Model 1	trans	transitivity (gwesp)	-0.87	-0.75	-41.03
	SAS	SAS similarity	-0.06	-0.01	5.70 (n. sign.)
	ethn	same ethnicity	-0.61	-1.87	147.92
	grades	grades similarity	-0.02	0.02	10.60
	sex	same sex	0.20	0.05	-1.75
	classroom	current year classroom + past year classroom	-0.07	0.47	1.84
	track	same * special track + same * normal track	-0.02	x	3.28
	resid	same primary school + residential distance + residential distance squared	-0.04 (n. sign.)	-0.03	10.23
Model 2	cult	cult_lvl similarity + music matching + youtube matching	-0.01	-0.36	-12.91
Model 3	parents	w1_parents	-0.19	-0.10	20.26

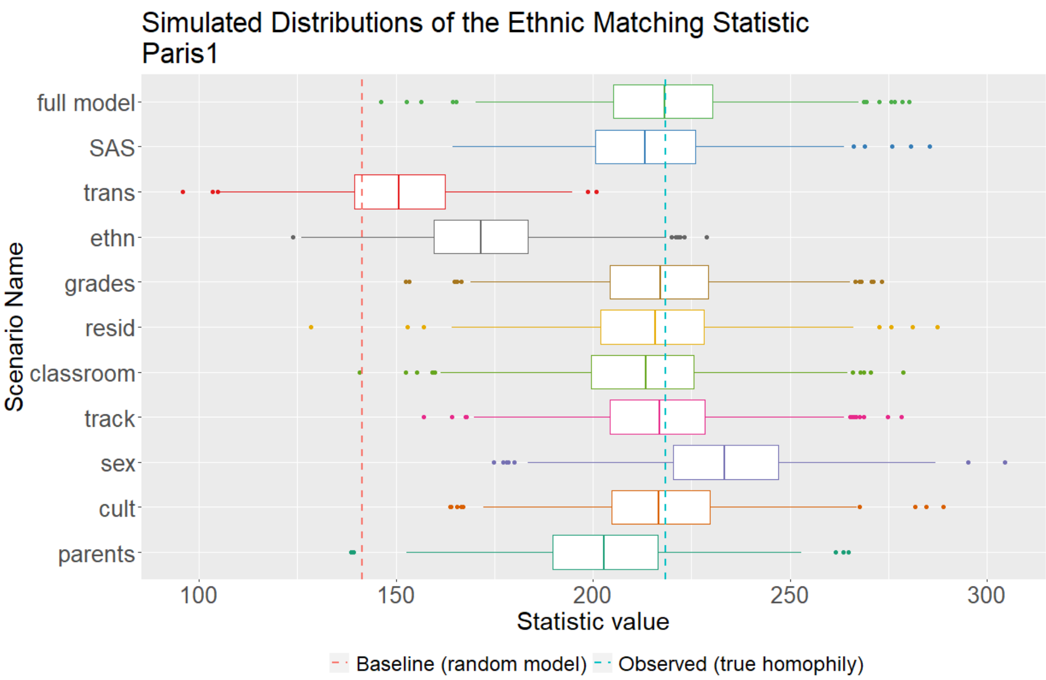
“n. sign.” indicates that the corresponding parameter(s) in the SAOM are not significantly different from 0 at the 10% threshold (see Table 5-5).

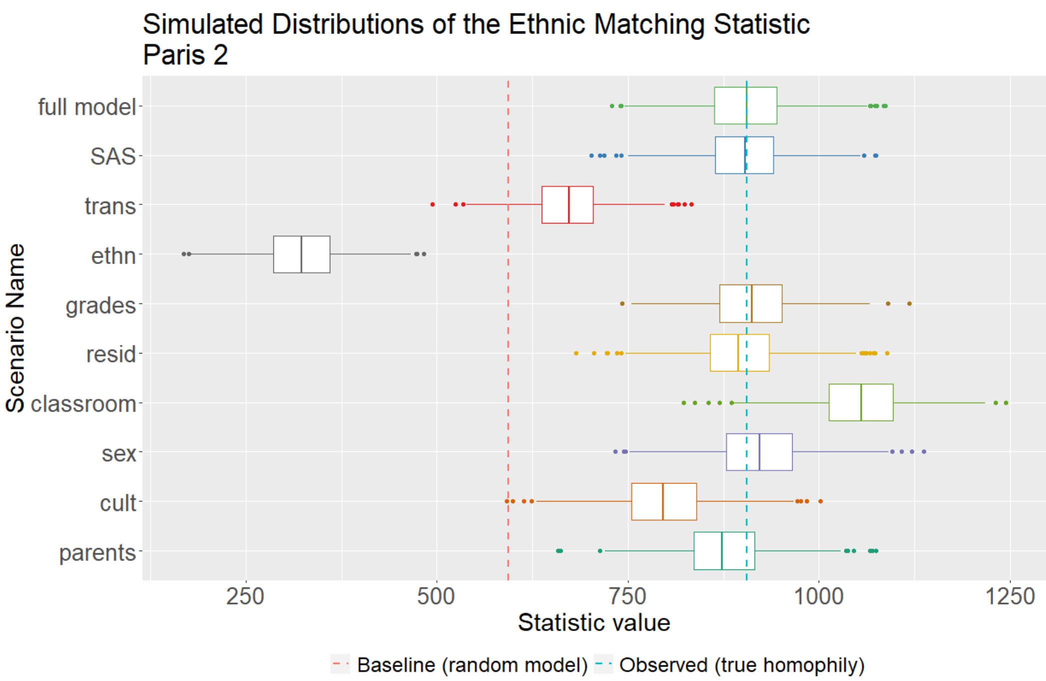
Table 2: Contribution Scores for Gender Homophily

Model	Scenario Name	Parameter(s) set to 0	Predicted % change in ethnic homophily		
			Paris 1	Paris 2	Savoie 1
Model 1	trans	transitivity (gwesp)	-0.09	-0.35	-0.55
	SAS	SAS similarity	0.02	0.00	-0.08 (n. sign.)
	ethn	same ethnicity	0.03	0.00	-0.08
	grades	grades similarity	-0.01	-0.02	-0.08
	sex	same sex	-1.21	-0.73	-0.75
	classroom	current year classroom + past year classroom	0.07	0.01	-0.05
	track	same * special track + same * normal track	-0.16	x	-0.08
	resid	same primary school + residential distance + residential distance squared	-0.01 (n. sign.)	0.00	-0.04
Model 2	cult	cult_lvl similarity + music matching + youtube matching	0.01	0.01	-0.02
Model 3	parents	w1_parents	0.01	-0.05	-0.07

Figure 1: Distributions of Simulated Statistics for Ethnic Homophily (same ethnicity) under the Different Counterfactual Scenarios

Note: boxes represent the values of the SAS statistic in the fictive networks. The blue vertical line corresponds to the true observed value of the statistic, and the red line to the average value observed under the null distribution. Therefore, the total amount of socioeconomic homophily observed empirically is represented by the distance between the red and blue lines, and the reduction in the SAS similarity statistic is the distance between the blue line and the center of the box. Contribution scores are defined as the reduction in the statistic divided by the total homophily, so they are represented by the proportion of the red/blue distance that stands right to the center of the box.





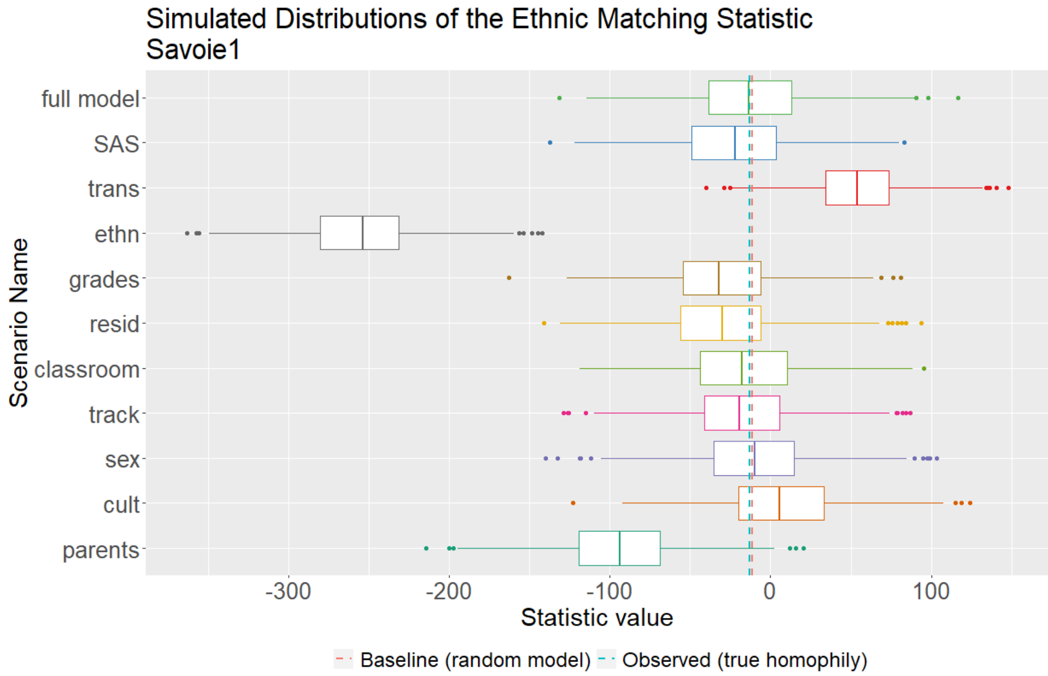
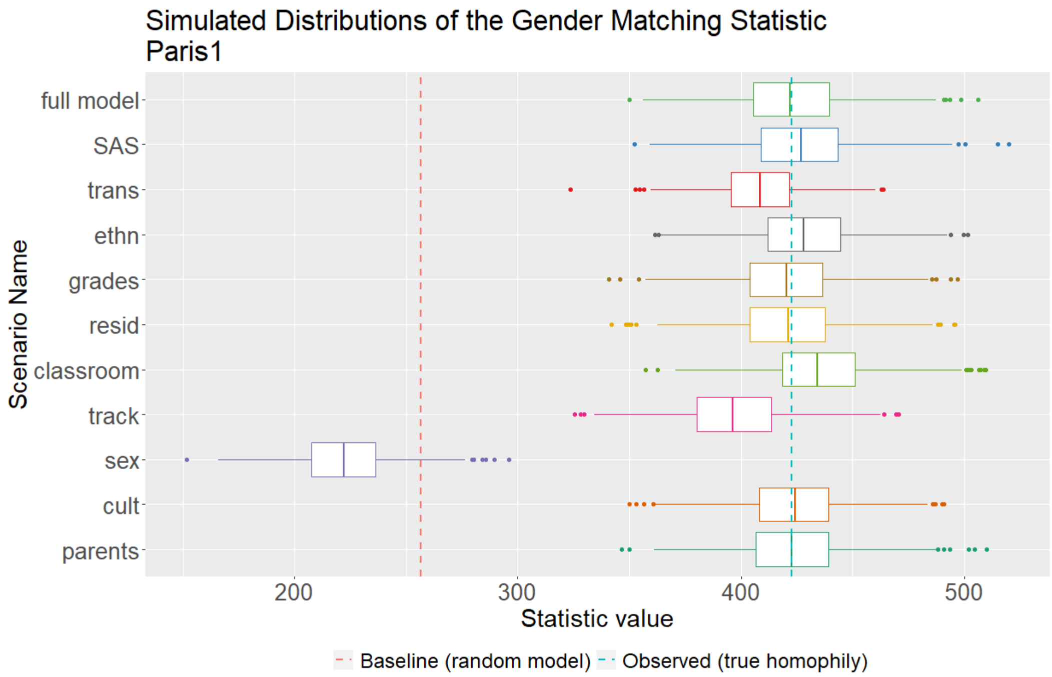
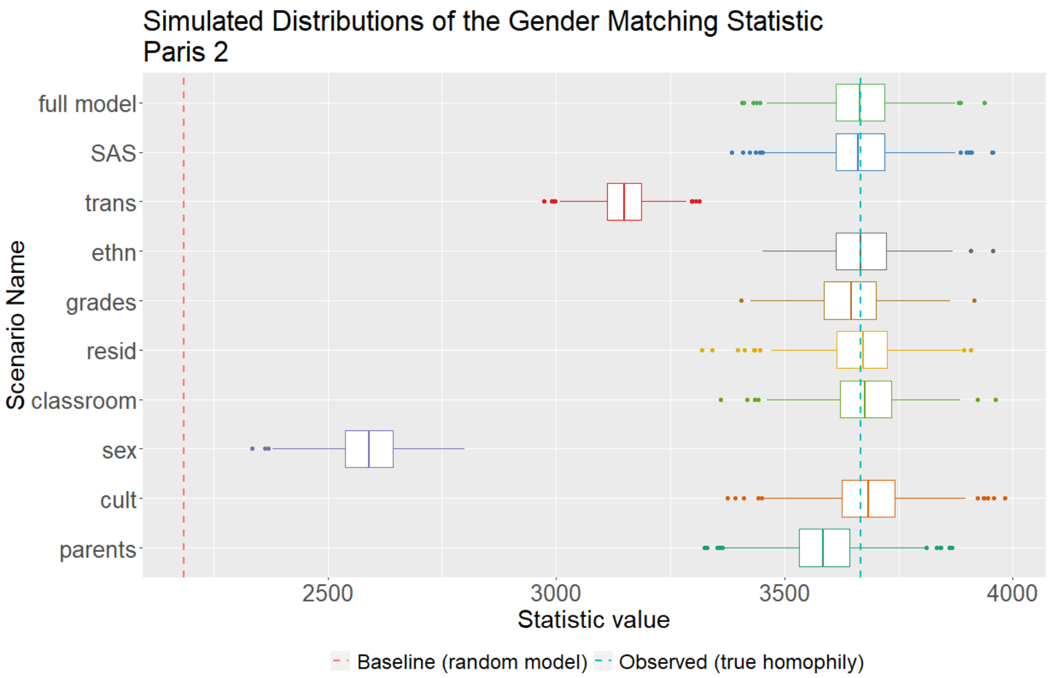


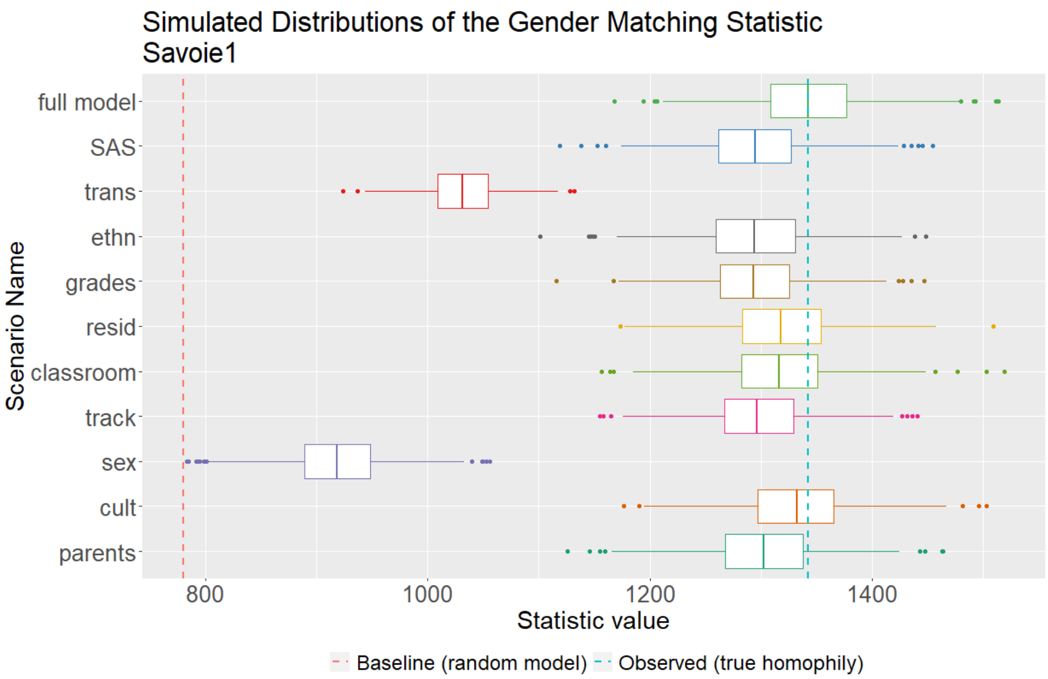
Figure 2: Distributions of Simulated Statistics for Gender Homophily (same gender) under the Different Counterfactual Scenarios

Note: boxes represent the values of the SAS statistic in the fictive networks. The blue vertical line corresponds to the true observed value of the statistic, and the red line to the average value observed under the null distribution. Therefore, the total amount of socioeconomic homophily observed empirically is represented by the distance between the red and blue lines, and the reduction in the SAS similarity statistic is the distance between the blue line and the center of the box. Contribution scores are defined as the reduction in the statistic divided by the total homophily, so they are represented by the proportion of the red/blue distance that stands right to the center of the box.





Simulated Distributions of the Gender Matching Statistic
Paris 2
full model
SAS
trans
ethn
grades
resid
classroom
sex
cult
parents
Scenario Name
2500
3000
3500
4000
Statistic value
Baseline (random model)
Observed (true homophily)



Simulated Distributions of the Gender Matching Statistic
Savoie1
full model
SAS
trans
ethn
grades
resid
classroom
track
sex
cult
parents
Scenario Name
800
1000
1200
1400
Statistic value
Baseline (random model)
Observed (true homophily)

Appendices for Chapter 6

Appendix 6A: Sociodemographic Information about the Interviewed Students

Table 1: Information about the Interviewed Students in Phase 1 (wave 2/3)

Name (Anonymized)	Coded PCS for the Father (lvl 3)	Coded Pcs for the Mother (lvl 3)	Household Head PCS (lvl 1)	Socio-Academic Score	Grade obtained at the 2nd Trimester of the 2nd Year (Trimester General Mean)	School
DAHLI	21=63	81	8	-1.72	-0.6	Savoie 2
ASSA	68	82	6	-1.08	-0.51	Paris 1
RAYENE	67;?	99	6	-1.08	-0.68	Savoie 2
NOEMIE	21=63	67=62	6	-0.94	-0.53	Savoie 2
JADE	62=67	99	NA	-0.94	0.56	Savoie 2
KAIS	68=63	82	6	-0.94	-0.87	Savoie 2
IRIS	55	56	5	-0.69	-0.7	Paris 1
LAURA	22	67	2	-0.67	0.17	Savoie 1
SYLVAIN	62=48	56	5	-0.65	-0.34	Savoie 2
ILAN	65	52	5	-0.6	-1.62	Savoie 2
HUGO	63=47	56	5	-0.54	0.23	Savoie 2
AHMED	68	54	5	-0.44	0.56	Paris 1
INES	68=63;47	54	5	-0.39	0.4	Paris 1
ALY	?	52	5	-0.3	0.4	Paris 1
MAILYS	10	52	5	-0.3	2.02	Savoie 2
MATHILDE	53	82	5	-0.15	-0.59	Savoie 2
ASAN	46=37	55;46	4	0.29	1.57	Savoie 2
LAURENCE	81	46=54	NA	0.33	0.6	Paris 1
QUENTIN	23	82	2	0.35	0.22	Savoie 1
CHARLOTTE	46=54	54	5	0.35	-1.41	Savoie 1
YANN	62=67	31=37	3	0.38	-1	Savoie 1
THEOPHILE	62	33	3	0.53	0.78	Savoie 1
KEVIN	45;33	82	4	0.66	-1.16	Paris 1
BARNABE	45;33	54;46	4	0.76	-0.8	Paris 1
BRIEUC	21=63	42	4	0.77	0.25	Savoie 1
MARION	43	82	4	0.87	0.89	Paris 1
JULIE	35	?	3	1.03	-0.95	Savoie 2
KEZIAH	21	34	3	1.19	-0.6	Paris 1
MANON	44	44=34	4	1.36	1.6	Paris 1
JASMINE	22	34	3	1.43	0.56	Paris 1
EMMA	38	37=46	3	1.51	1.94	Savoie 2

NINON	54	33	3	1.55	0.31	Savoie 2
REMI	23	42	2	1.56	1.58	Savoie 2
JEAN	46=37	42	4	1.65	1.12	Savoie 1
PAUL	42;43	33	3	1.67	1.24	Paris 1
CELINE	53	31=34	5	1.85	0.74	Savoie 1
AURIANNE	37	31	3	1.91	1	Savoie 2
HARUHI	33;46	31	3	2.14	1.44	Savoie 1
ANNA	34	33=34	3	2.23	1.44	Paris 1
LOU	34	34	3	2.43	1.6	Paris 1

Note: Grades are standardized per school. For the PCS, the sign “=” indicates that the two categories have been deemed equally possible when coding the declared occupation, whereas “;” signifies that the left-hand category has been deemed more likely, but the right-hand one is possible nevertheless. The household head is defined as the parent living with the student that has the highest PCS.

Appendix 6B: Probability of Having Smartphone depending on Socioeconomic Background

Table 1: Logit Model for the Possession of a Smartphone in Wave 2 (all schools, including controls for school and sex)

	Estimate	Standard Error
Intercept	2.04 ***	0.24
Sex – Female	ref.	-
Sex – Male	-0.29	0.21
School – Paris 1	0.31	0.41
School – Paris 2	0.01	0.27
School – Savoie 1	-0.60 **	0.27
School – Savoie 2	ref.	-
SAS	-0.29 ***	0.11

***: p<0.01; **: p<0.05; *: p<0.1

Appendices for Chapter 7

Appendix 7A: Complete Stergm Output (Table 7-2 from Main Text)

Table 1: STERGM Estimation (Paris 1, Very Good Friend Networks)

	w1 → w2	w2 → w3	w3 → w4	w4 → w6
Density	-2.36*** (0.11)	-2.61*** (0.11)	-2.63*** (0.13)	-2.49*** (0.12)
SAS abs. difference	-0.32*** (0.09)	-0.37*** (0.09)	-0.65*** (0.10)	-0.68*** (0.11)
SAS Reception	-0.09 (0.07)	-0.03 (0.07)	0.18** (0.08)	-0.02 (0.09)
SAS Emission	-0.14* (0.07)	-0.13* (0.07)	0.06 (0.08)	-0.21** (0.09)
Density	0.87*** (0.17)	0.72*** (0.14)	0.56*** (0.15)	0.56*** (0.16)
SAS abs. difference	-0.02 (0.15)	-0.49*** (0.12)	-0.20 (0.13)	-0.23 (0.15)
SAS Reception	-0.07 (0.11)	0.16* (0.09)	0.004 (0.09)	0.03 (0.10)
SAS Emission	0.07 (0.12)	-0.11 (0.09)	0.26*** (0.10)	-0.38*** (0.10)

***: p<0.01; **: p<0.05; *: p<0.1

Appendix 7B: Socioeconomic Homophily and Primary School Homophily

Table 1: ERGM Estimation (Very Good Friend Networks, Wave 1)

	Paris 1	Paris 2	Savoie 1	Savoie 2
Density	-2.04*** (0.321)	-6.201*** (1.011)	-2.677*** (0.400)	-3.174*** (0.278)
SAS abs. difference	-0.378*** (0.071)	-0.163*** (0.024)	-0.199*** (0.048)	-0.109*** (0.037)
SAS Reception	0.035 (0.058)	0.005 (0.018)	-0.034 (0.037)	0.153*** (0.027)
SAS Emission	0.096* (0.058)	-0.056*** (0.018)	-0.116*** (0.038)	0.253*** (0.027)
Same primary school	0.809*** (0.184)	0.953*** (0.068)	2.043*** (0.110)	1.307*** (0.097)
Prima. Sch. Emission	not shown	not shown	not shown	not shown
Prima. Sch. Emission	not shown	not shown	not shown	not shown
SAS abs. diff. * same primary school	0.305** (0.139)	-0.115** (0.048)	-0.089 (0.091)	-0.062 (0.071)

***: p<0.01; **: p<0.05; *: p<0.1
Note: Emission and reception terms for the former primary school are not shown, as these are very long (there is one parameter per school, meaning over 30 in some schools).

Appendix 7C: Quadratic Effect for Socioeconomic Selection in Paris 1

The following SAOM is similar to "Model 1" from Table 5-3 (Main Text), except that the "SAS similarity" parameter has been replaced by two terms. One is the absolute difference in SAS (equivalent to the similarity score, except that it is not rescaled on [0,1]) and the second is the squared difference in SAS. Only this latter coefficient is statistically significant, indicating a quadratic pattern indeed: SAS difference seems to increase the odds of tie formation (or maintenance) only for sufficiently large differences.

Table 1: SAOM Estimation for the Quadratic Effect of Socioeconomic Distance (Paris 1, Very Good Friend Networks)

	Estimate	Standard Error	Convergence t-ratio
1. rate constant friendship rate (period 1)	14.37	(1.1272)	-0.03
2. rate constant friendship rate (period 2)	19	(1.2993)	-0.01
3. rate constant friendship rate (period 3)	13.55	(1.1850)	0.02
4. eval outdegree (density)	-2.82	(0.1357)	0.02
5. eval reciprocity	2.36	(0.1810)	0.01
6. eval GWESP I -> K -> J (110)	1.14	(0.0536)	0.02
7. eval indegree - popularity	-0.03	(0.0110)	0.02
8. eval outdegree - popularity	-0.06	(0.0077)	0.02
9. eval outdegree - activity	0	(0.0032)	0.02
10. eval m_3main_ethn5	0.13	(0.0601)	0.03
11. eval same track – special track	0.04	(0.1667)	0.03
12. eval same track – regular track	0.51	(0.1357)	-0.02
13. eval s_mena_ethn5	-0.23	(0.0772)	0
14. eval s_europe_ethn5	0.08	(0.1195)	-0.01
15. eval s_subsah_ethn5	-0.04	(0.0965)	0.01
16. eval s_other_ethn5	0.5	(0.1726)	-0.02
17. eval r_mena_ethn5	0.08	(0.0714)	0.02
18. eval r_europe_ethn5	0.05	(0.1032)	0.03
19. eval r_subsah_ethn5	0.3	(0.0895)	-0.02
20. eval r_other_ethn5	0.01	(0.1546)	-0.03
21. eval walk_rg	-0.24	(0.2955)	-0.03
22. eval walk_sq	0.26	(0.2266)	-0.04
23. eval s_5B_classroom	0.13	(0.0797)	-0.01
24. eval s_5C_classroom	0.24	(0.0936)	0.01
25. eval r_5B_classroom	0.03	(0.0820)	-0.01
26. eval r_5C_classroom	0.15	(0.0915)	0.03
27. eval s_4B_classroom	0.19	(0.1326)	0
28. eval s_4C_classroom	0.32	(0.1349)	0.01

29. eval r_4B_classroom	-0.14	(0.1384)	-0.03
30. eval r_4C_classroom	0.44	(0.1350)	0.03
31. eval sas alter	0.1	(0.0345)	0
32. eval sas ego	-0.01	(0.0374)	0
33. eval sas abs. difference	0.14	(0.1262)	0.03
34. eval sas diff. squared	-0.09	(0.0473)	0.03
35. eval sex alter	0.02	(0.0588)	0
36. eval sex ego	-0.23	(0.0610)	0.01
37. eval sex similarity	0.46	(0.0522)	0
38. eval same primary1	0.04	(0.0633)	0.09
39. eval oib ego	-0.12	(0.1880)	0.02
40. eval same classroom1	0.17	(0.0627)	0.02
41. eval same classroom2	0.49	(0.0630)	0
42. eval grades alter	-0.01	(0.0371)	0
43. eval grades ego	-0.03	(0.0395)	-0.03
44. eval grades similarity	0.58	(0.1659)	-0.04
45. eval timedummy_period2 ego	-0.35	(0.0662)	-0.01
46. eval timedummy_period3 ego	-0.48	(0.1260)	-0.01
47. eval int. reciprocity x GWESP I -> K -> J (110)	-0.49	(0.1046)	0.02
Overall Maximum Convergence Ratio		*0.2*	
Total Number of Iteration Steps		*2218*	

Note: “s_” and “r_” indicate, respectively, sender and receiver effects. These are essentially equivalent to “ego” and “alter” effects, except that they have been passed through SIENA as a dyadic covariate (effect “X”) rather than applying the ego and alter effect to an individual attribute. “Ethn5” refer to the five-posts ethnicity attribute. For the homophily effect, only the three main ethnic groups (natives, Middle-East and Subsaharian Africa) are counted. “walk_rg” is the linear walking distance between students’ homes, and “walk_sq” the squared distance. “Cult_lvl” indicate the attribute of cultural status, and “match_music” and “match_youtube” the covariates for cultural matching. “w1_parents” is the effect of the parental ties from wave 1. All other parameters have the effect names given by the Rsiena software.
This note applies to all further SAOM tables thereafter.

Appendix 7D: SAOM Model for the Interaction of Socioeconomic Similarity and Track Homophily

The estimated SAOM is similar to "Model 1" in chapter 5 (see Table 5-3 of the main text), except that an interaction effect is added between the parameter of socioeconomic similarity and the parameter of track matching for the regular track. Therefore, the parameter of socioeconomic similarity now represents socioeconomic selection among two types of dyads: mixed dyads (i.e. the nodes are from different tracks) and dyads from the international section. The interaction effect indicates the change in the coefficient of socioeconomic similarity that needs to be applied for regular track dyads (i.e. neither of the nodes is from the international section).

Note that, compared to the analyses of the main text, I do not differentiate between mixed dyads and international-to-international dyads: this would make for an exceedingly complex model, with very large standard errors on the coefficients. In practice, the parameter mostly reflects what happens in mixed dyads, as these are much more numerous.

Results show a statistically significant effect of socioeconomic similarity in mixed and international dyads. Since we know that socioeconomic factors hardly matter among students from the international section, this effect can safely be considered as reflecting the effect among mixed dyads only (in fact, it is probably under-estimated due to the inclusion of international dyads). Therefore, there seems to be socioeconomic selection among mixed dyads; in practice, this means that students from the international section are much more likely to send and receive ties from upper-class non-international students, than from working-class non-international ones (controlling for all other model parameters, which strongly suggests that this is due to dispositional selection).

On the other hand, the interaction parameter is not statistically significant, which makes a definite conclusion impossible. Nevertheless, based on the sign and size of the parameter, there seems to be virtually no effect of socioeconomic similarity among non-international students: the main coefficient is of 0.41 and the interaction effect of -0.40, so the resulting coefficient of socioeconomic similarity among regular-track dyads is 0.01. Again, the estimate is very imprecise – given the number of parameters in the model and the relatively small sample size of Paris 1, the model might face over-fitting issues at this point. At the very least, though, this seems coherent with the idea that dispositional selection mostly operates among mixed dyads – suggesting that a good part of the socioeconomic homophily of the network is due to upper-class non-international students “crossing over” the track difference to befriend international students, whereas lower-class non-international ones do not. This resonates with the hypothesis put forward in the main text that the international section may act as a kind of anchor point for upper-class non-international students, indirectly pulling them away from their working-class peers (see chapter 7 section 2.2).

Finally, time heterogeneity tests find no time heterogeneity in the parameters of interest (socioeconomic similarity and interaction with track homophily). This supports the conclusion of the main text that socioeconomic selection has not increased between waves 1 and 4 in Paris 1 (or not much).

Table 1: SAOM Estimation for the Interaction of Socioeconomic Similarity and Track Homophily (Paris 1, Very Good Friend Networks)

	Estimate	Standard Error	Convergence t-ratio
1. rate constant friendship rate (period 1)	14.37	-1.17	-0.04
2. rate constant friendship rate (period 2)	18.97	-1.32	0.05
3. rate constant friendship rate (period 3)	13.53	-1.12	0.05
4. eval outdegree (density)	-2.84	-0.13	0.03
5. eval reciprocity	2.36	-0.18	0.03
6. eval GWESP I -> K -> J (110)	1.14	-0.06	0.04
7. eval indegree - popularity	-0.03	-0.01	0.02
8. eval outdegree - popularity	-0.06	-0.01	0.03
9. eval outdegree - activity	0	0	0.04
10. eval m_3main_ethn5	0.13	-0.06	0.05
11. eval same special track	-0.01	-0.16	-0.01
12. eval same regular track	0.51	-0.14	0
13. eval s_mena_ethn5	-0.23	-0.08	-0.03
14. eval s_europe_ethn5	0.06	-0.12	-0.07
15. eval s_subsah_ethn5	-0.04	-0.1	0.05
16. eval s_other_ethn5	0.51	-0.16	-0.01
17. eval r_mena_ethn5	0.07	-0.07	0
18. eval r_europe_ethn5	0.05	-0.1	0.02
19. eval r_subsah_ethn5	0.29	-0.09	0.03
20. eval r_other_ethn5	0.01	-0.17	0.01
21. eval walk_rg	-0.25	-0.29	-0.01
22. eval walk_sq	0.27	-0.22	-0.01
23. eval s_5B_classroom	0.12	-0.08	0.03
24. eval s_5C_classroom	0.24	-0.1	-0.03
25. eval r_5B_classroom	0.02	-0.08	-0.02
26. eval r_5C_classroom	0.15	-0.1	0.02
27. eval s_4B_classroom	0.19	-0.13	0.02

28. eval s_4C_classroom	0.33	-0.14	0.05
29. eval r_4B_classroom	-0.13	-0.13	0.03
30. eval r_4C_classroom	0.44	-0.13	0.03
31. eval sas alter	0.1	-0.04	-0.02
32. eval sas ego	0	-0.04	0
33. eval sas similarity	0.41	-0.15	0.01
34. eval sas similarity * same regular track	-0.4	-0.33	0.02
35. eval sex alter	0.02	-0.06	0
36. eval sex ego	-0.24	-0.06	0.03
37. eval sex similarity	0.46	-0.05	0.02
38. eval same primary1	0.04	-0.06	0.04
39. eval oib ego	-0.13	-0.2	-0.01
40. eval same classroom1	0.17	-0.06	0.04
41. eval same classroom2	0.49	-0.07	0.02
42. eval grades alter	-0.01	-0.04	-0.04
43. eval grades ego	-0.03	-0.04	0.01
44. eval grades similarity	0.57	-0.17	0.01
45. eval timedummy_period2 ego	-0.34	-0.06	0.01
46. eval timedummy_period3 ego	-0.48	-0.11	0.03
47. eval int. reciprocity x GWESP I -> K -> J (110)	-0.49	-0.11	0.04
Overall Maximum Convergence Ratio		*0.24*	
Total Number of Iteration Steps		*1949*	

Note: “s_” and “r_” indicate, respectively, sender and receiver effects. These are essentially equivalent to “ego” and “alter” effects, except that they have been passed through SIENA as a dyadic covariate (effect “X”) rather than applying the ego and alter effect to an individual attribute. “Ethn5” refer to the five-posts ethnicity attribute. For the homophily effect, only the three main ethnic groups (natives, Middle-East and Subsaharian Africa) are counted. “walk_rg” is the linear walking distance between students’ homes, and “walk_sq” the squared distance. All other parameters have the effect names given by the Rsiena software.

Table 2: Time Heterogeneity Tests for the SAOM of Table 1 (Effect Test)

	Paris 1				Paris 1 (continued)		
	chi-sq.	df	p-value		chi-sq.	df	p-value
reciprocity	0.24	2	0.89	sas alter	0.05	2	0.98
GWESP I -> K -> J	0.42	2	0.81	sas ego	0.81	2	0.67
indegree - popularity	0.08	2	0.96	sas similarity	0.3	2	0.86
outdegree - popularity	0.46	2	0.8	sas similarity * same regular track	0.41	2	0.82
outdegree - activity	0.06	2	0.97	sex alter	0.74	2	0.07
same ethnicity	1.37	2	0.5	sex ego	5.41	2	0.36
s_mena_ethn5	4.1	2	0.13	sex similarity	2.07	2	0.99
s_europe_ethn5	5.38	2	0.07	same primary1	0.02	2	0.66
s_subsah_ethn5	2.71	2	0.26	same special track	0.12	2	0.94
s_other_ethn5	3.77	2	0.15	same regular track	1.69	2	0.43
r_mena_ethn5	2.47	2	0.29	special track ego	0.83	2	0.66
r_europe_ethn5	0.03	2	0.99	same classroom	0.08	2	0..941
r_subsah_ethn5	1.83	2	0.4	grades alter	0.54	2	0.76
r_other_ethn5	3.39	2	0.18	grades ego	0.77	2	0.68
walk_regular	0.3	2	0.86	grades similarity	0.16	2	0.92
walk_squared	0.37	2	0.83	int. reciprocity x GWESP I -> K -> J	0.47	2	0.79

Joint Significance Test of Time Heterogeneity: chi-squared = 96.17, d.f. = 66, p= 0.0091.

Appendices for Chapter 8

Appendix 8A: Homophily of Past Friendships

In the body of the article, we did permutation tests for the network of previous friendships, using all nodes present in the camp (see section 4.1.1). This gives us an indication about the homophily present at the start of the stay, but not about the homophily present in previous camps. For this, it is necessary to take into account the meeting opportunities adolescents had during these previous camps.

For all adolescents of the 2019 camp, we have two types of information: the previous camps they have attended – thus the peers they already met –, and the peers whom they declared to have been friend with at the time. We therefore define two networks: the network of past friendship ties (same as in the body of the article), and the network of past meeting ties, where a meeting tie is defined as having met at least once in the past four years (i.e. having been in the same camp). Importantly, the first network (friendship) is a strict subset of the second (meeting). Then, we repeatedly draw random samples of meeting ties that have the same density as the network of friendship ties, and compute for each the sum of absolute differences in ISEI scores among connected nodes. By comparing mean ISEI differences in the friendship network with this fictive distribution, we can therefore assess whether there is socioeconomic homophily in past friendships net of meeting opportunities (note that this test controls for network density, but not for network structure, as the randomly drawn samples do not have the same structure as the observed friendship network).

The test indicates that the previous friendships have lower ISEI differences than if ties had been formed at random among the previous meetings (observed mean of ISEI differences: 16.43, mean in the random samples: 23.52, p-value: 0). Also, note that the mean ISEI in past meetings (23.52) is practically identical to the mean ISEI among all dyads in the camp, including null ones (23.80). What this means is that past meeting opportunities among the camp's attendees were not segregated by socioeconomic background whatsoever. Therefore, the socioeconomic homophily observed in past friendships cannot be explained by meeting opportunities alone.

Appendix 8B: Selection Biases

In this appendix, we report the information used to assess the presence of selection biases in the sample of participants. Two elements are considered: participants' ISEI (which is directly observed) and alleged homophilic dispositions (which are not observed).

Distribution of ISEI Scores

We use the PISA 2018 sample for France as a reference, as it is a nationally representative sample of French adolescents of age 14. Looking at the highest ISEI score out of both parents, the average value reported in the PISA study is 51.5, and the median 54.5. In our sample, the average score for the highest ISEI is 62.00 and the median 72.24 (respectively 55.30 and 58.69 for the average of both parents' ISEI, which we could not find for PISA). This indicates a clear over-representation of high ISEI scores relative to the national population.

Figure 1: Histograms for the Average of Parents' ISEI (left) and the highest ISEI out of Both Parents (right) in the 2019 Dataset

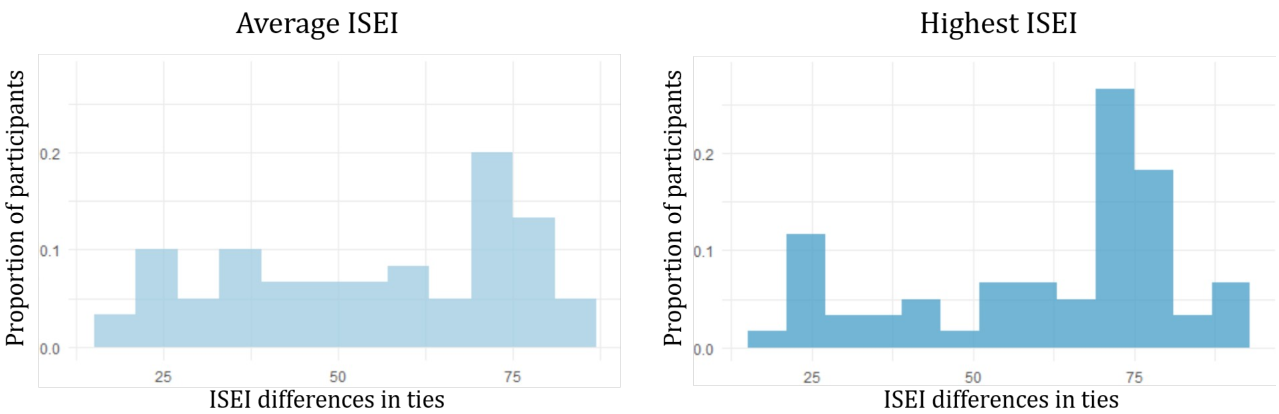


Table 1: Comparison of Quantile Values for the Highest ISEI of Parents in the PISA 2018 Dataset and the 2019 Camp Dataset

	Quantiles for Highest ISEI						
	5	**10**	**25**	**50**	**75**	**90**	**95**
PISA dataset	17.0	22.8	28.5	54.5	72.3	80.5	82.4
Camp dataset	25.0	26.6	49.4	72.2	75.5	80.7	88.7

To investigate the reasons for this over-representation, we divide the 60 respondents of our sample in three groups, based on which camp they attended in the summer 2018 (i.e. one year before the present study). 16 teenagers had attended the same camp (i.e. in the same facility as in 2019), 27 a different camp, and 17 no camp (meaning that the 2019 camp is the first they ever attended). These three groups should be subject to three different levels of selection biases. The respondents that were in the same camp in 2018 may have self-selected into the 2019 camp based on whether they enjoyed their 2018 stay, which means that they are subject to self-selection due to properties of this particular facility. The ones that attended another camp could self-select depending on their experience of these summer camps in general, but not of this particular facility. Finally, the ones who attended a camp for the first time in 2019 could not self-select based on their experience of previous camps, but only through anticipation of what this summer camp might be (thus with very limited information).

Table 2 shows the mean and median ISEI (highest parent and average of both parents) for these three groups separately. It clearly appears that ISEI scores are higher for the adolescents who attended this same facility in 2018.

Table 2: Means and medians of ISEI Distributions Depending on Adolescents' Facility of Attendance on the Previous Year

	Average ISEI		**Highest ISEI**	
	Mean	Median	Mean	Median
Same Camp	73.16	73.81	65.49	72.58
Other Camp	56.34	57.99	51.32	51.67
No Camp	60.48	65.01	52.06	50.42

Note: "same camp" indicates teenagers who already came in this same facility in 2018, "other camp" those that went to a different facility, and "no camp" those that attended no camp in 2018.

Therefore, even though all groups have means and medians somehow higher than the reference PISA sample, most of the selection effect for higher-background adolescents seems to rest on those that came back from one year to the next into this same facility.

In theory, this over-representation of high ISEI scores among adolescents that came back from 2018 could be due to three factors. First, the 2018 camp itself may have had an over-representation of high ISEIs, for various reasons. Second, higher-background attendees may have enjoyed the 2018 stay more, thus desiring to come back more than lower-background ones did. Third, higher-background attendees, given the desire to come back to the same camp, may be more successful in doing so. For example, upper-class parents may be more used to filling administrative files properly and in time, in order to register for the particular camp that their children wished for. Informal discussion with members of the organizing institution suggests that this is the case indeed: the facility we studied is one of the most demanded by families, and available spots are mostly attributed on a "first come first served" basis.

Reasons to Participate

In order to further investigate selection biases in the sample, the 60 respondents were asked, during the camp, the reason for which they had registered to this particular camp. Answers were then grouped into 5 categories: (a) location in the mountains (as the camp took place in the Alpes) (b) to meet with friends from previous summers (c) because they knew this particular facility and liked it (d) practical reasons (dates of the stay were convenient, or the camp was closer from home than other facilities were) and (e) all the reasons that implied the child *not* to choose this camp in particular (chose at random, parents chose in his/her stead, followed the recommendation from the doctor, got affected here by the institution, had attended other facilities before and wanted to try a new one). Note that each adolescent could provide several answers. Table 3 shows the number of respondents that gave each type of answer, depending on the quartile of ISEI score they belong to.

Table 3: Reasons for Participating to the Camp, by ISEI Quartiles (2019 Dataset)

		Location in the Mountains	**Meeting with Friends**	**Knew the Facility**	**Practical Reasons**	**Did Not Choose this Camp in Particular**
Quartile of ISEI	1st	**9**	4	1	0	9
	2nd	**7**	5	5	1	6
	3rd	**5**	5	7	2	3
	4th	**7**	4	5	0	9

Performing a chi-squared test on Table 3, we find no significant differences across quartiles in the reasons declared for coming to the camp. This suggests that, at the very least, respondents from different backgrounds did not have widely different motivations for coming to the camp.

Differences in Levels of Homophily Depending on Attendees' Situation on the Previous Year

That attendees to the camp were selected on their socioeconomic background – as shown by the over-representation of high ISEI scores – is not problematic in itself, since there is still enough socioeconomic diversity in the camp for questions of homophily to be meaningful. Rather, the selection bias most decisive for our research design is that pertaining to adolescents' intrinsic preferences toward socioeconomic homophily. By definition, these preferences are not directly observed. However, if our sample is made of individuals that have few homophilic preferences and/or are particularly inclined to heterophilic mating, this would severely limit the external validity of our conclusions.

To test for this, we kept the three groups defined above based on individuals' situation during the summer 2018. We reason that whatever selection biases are at work, they should be most pronounced for the first group (attendance to this very same camp in 2018), then for the second (attendance to any other camp) and finally the third (newcomers). This is further supported by our finding that the first group has an average ISEI significantly larger than the other two (the difference between groups 2 and 3 was not significant). Therefore, if it is the case that adolescents attending the camp were selected on weak homophilic inclinations, then we would expect socioeconomic homophily to be least pronounced in the friendship nominations emitted by the first group, then the second and finally the third. For e.g., one potential source of selection would be that only those adolescents that are well-disposed toward social mixing liked the camps they attended in the past (as these were mixed as well), and therefore came back from one year to the next. Clearly, such a bias cannot be at work for newcomers.

We carried out permutation tests on the average ISEI differences in friendship ties at T1 and T2 reported by respondents in these three categories.

Table 4: Permutation Tests for ISEI Homophily (Average Absolute Difference in Ties) for the Friendship Networks at Time 1 and 2, Depending on Individuals' Situation in the Previous Year

	Middle of the Stay (T1)			End of the Stay (T2)		
	Obs.	Av. Perm.	P-val.	Obs.	Av. Perm.	P-val.
Same Camp	18.14	**23.81**	0.007	18.06	23.88	0
Other Camp	26.73	**23.82**	0.09	25.94	23.89	0.15
No Camp	21.41	**23.88**	0.14	22.22	23.86	0.41

Note: "same camp" indicates teenagers who already came in this same facility in 2018, "other camp" those that went to a different facility, and "no camp" those that attended no camp in 2018. For each category of respondents, we consider the ties *emitted* by them.

These tests show that socioeconomic homophily is actually higher among the ties emitted by the adolescents who already attended the same camp in 2018: the mean ISEI difference is significantly lower than the mean of permutations (at the 90% threshold for x1 and the 99% threshold for x2), unlike the other two categories or the whole of friendship ties (cf. section 4.1.1 of the main text). What this suggests is that adolescents that came back from the previous year were *more* homophilic than newcomers or those that attended other camps.

Of course, these tests only pertain to the observed levels of homophily, and they are not necessarily indicative of individuals' intrinsic preferences. To a large extent, this over-representation of homophilic ties among group 1 is due to the past friendships inherited from last year's camp, which we discussed above. Still, this supports the idea that attendees to the camp were not selected on positive attitudes toward socioeconomic mixing and heterophilic mating; if anything, they might have been selected on the opposite. This comforts us in assuming that the lack of socioeconomic homophily on the camp cannot be explained by a selection bias alone.

It also makes the lack of socioeconomic homophily that we found overall even more surprising, given the fact that the structure of past friendships and of previous attendance to this same camp clearly favored homophilic bonding. As we discuss in section 4.3. of the main text, the socioeconomic homophily of past friendships got diluted in all the friendships formed during the camp. Similarly, the socioeconomic homophily of the adolescents that attended the same camp in 2018 is not enough to "pull" the entire network toward a significant homophily at the aggregate level.

Appendix 8C: Complete Permutation Tests

We report here the full results of the permutation tests carried out on friendship networks and meal partitions. These tests were first carried out on the average ISEI score of participants' parents, which is the variable used in the main text to capture the socioeconomic background of the adolescents. The same tests were also carried out for the maximal ISEI score of the parents, as a robustness check. This second variable is used based on the idea that the highest socioeconomic position of the parents might better represent the living conditions of a household \citep{erikson1984social}. All results are consistent to what is reported in the results section.

Table 1: Permutation Tests for ISEI Homophily in Friendship Networks (Average and Highest ISEI, All Networks)

	Anterior to the Stay			**Middle of the Stay**			**End of the Stay**		
	Obs.	Av. Perm.	P-val.	Obs.	Av. Perm.	P-val.	Obs.	Av. Perm.	P-val.
Average ISEI	16.43	23.76	0.01	22.95	23.81	0.37	22.84	23.81	0.26
Highest ISEI	14.94	22.27	0.02	21.97	22.31	0.76	21.95	22.3	0.75

Figure 1: Histograms of the Average ISEI Absolute Difference in Ties in the Permuted Networks

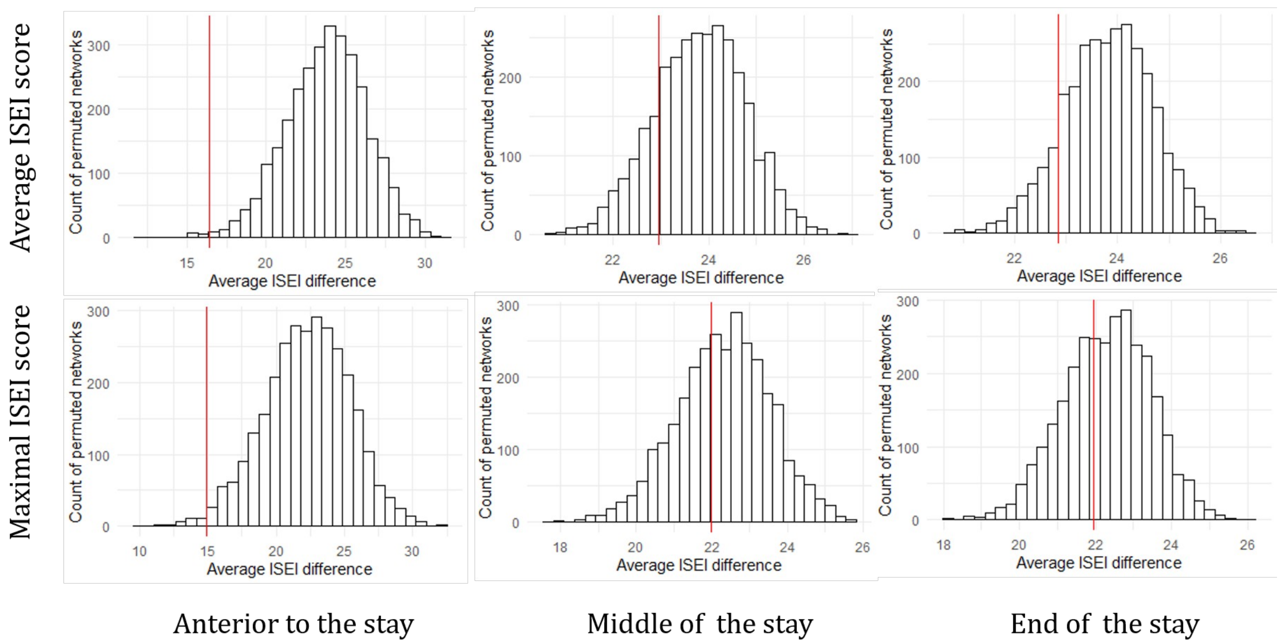


Table 2: Permutation Tests for ISEI Homophily in Meal Partitions (Average and Highest ISEI, Average Differences and Average Ranges of the Groups) – All Meals Together

	Average ISEI Differences			Average ISEI Ranges		
	Obs.	Av. Perm.	P-val.	Obs.	Av. Perm.	P-val.
Average ISEI	21.43	23.74	0.04	29.83	31.19	0.34
Highest ISEI	20.16	22.22	0.07	29.29	30.02	0.62

Figure 2: Histograms of the Average ISEI Differences and Average ISEI Ranges in Groups in the Permuted Partitions

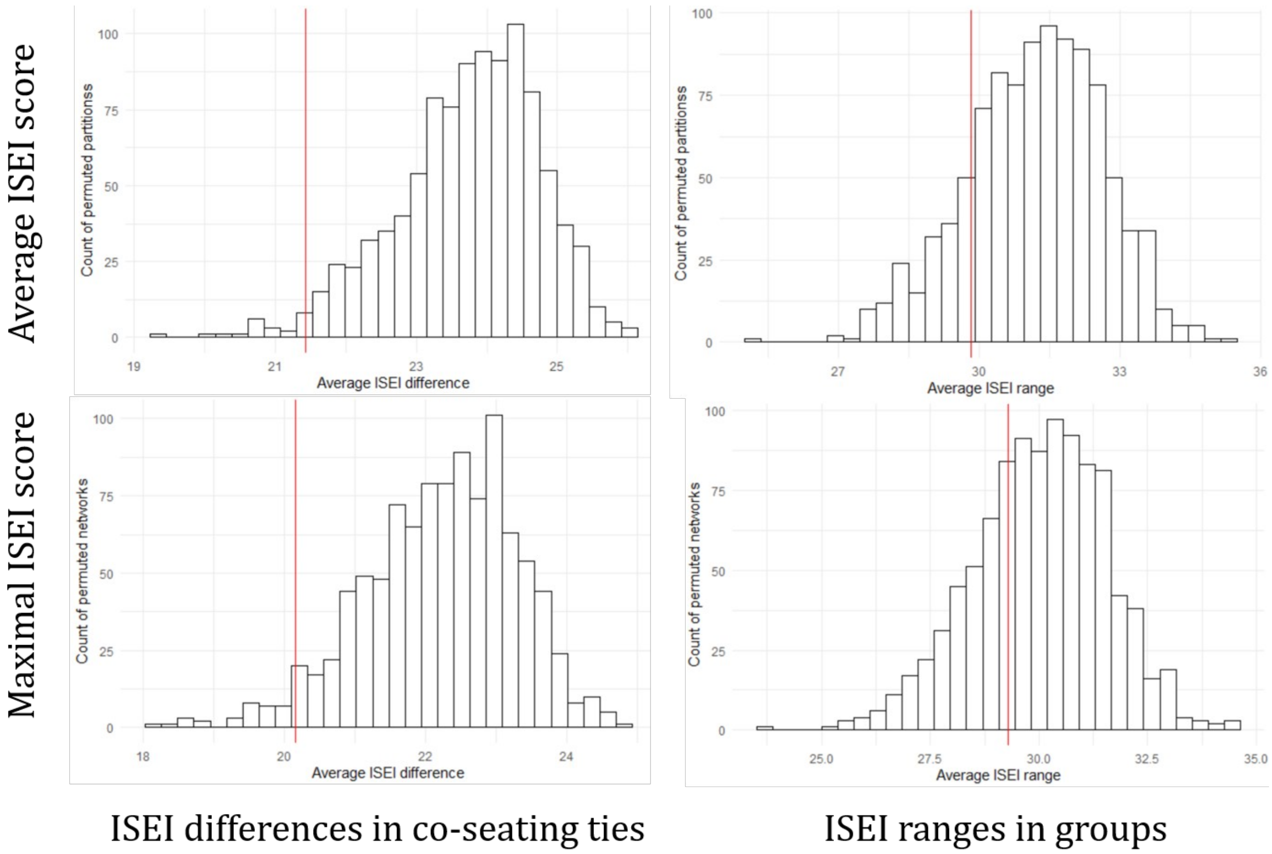
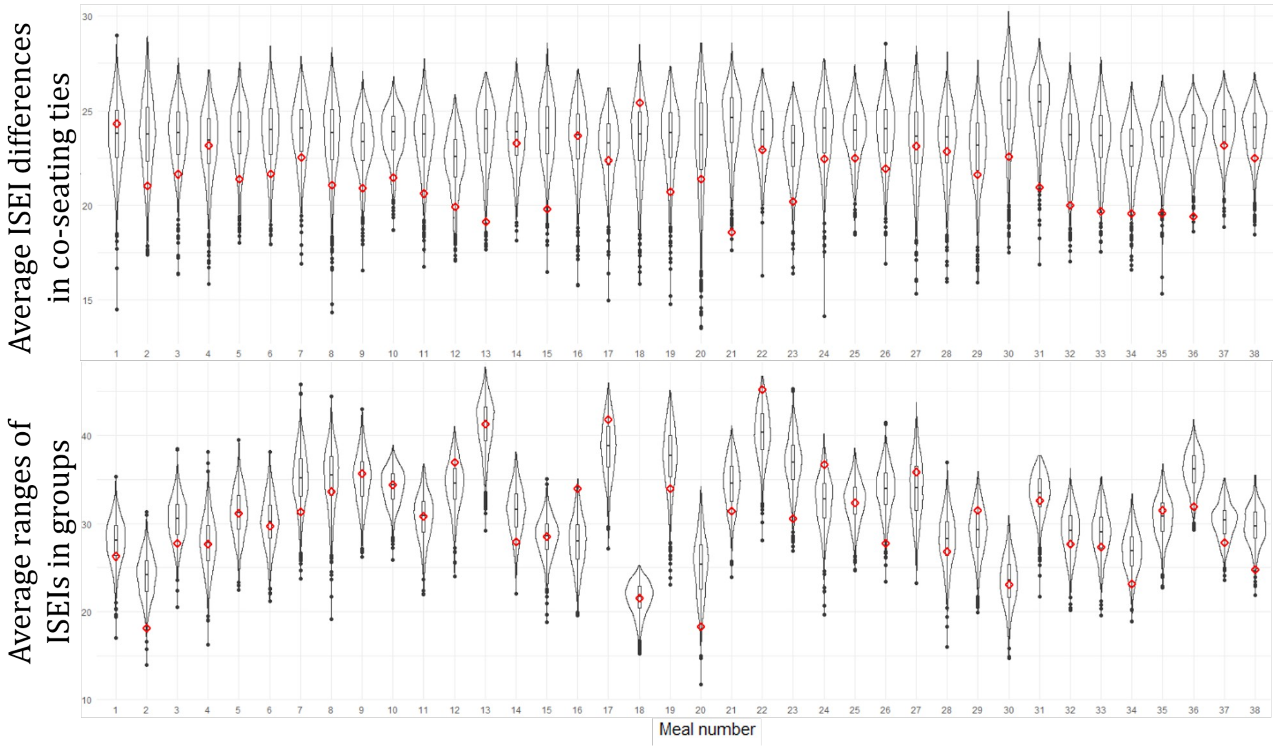


Table 3: Permutation Tests for ISEI Homophily in Meal Partitions, (Average and Highest ISEI, Average Differences and Average Ranges of the Groups) – Permutations of the 38 Meals Separately

	Observed			Average in Permutations			Nb of Obs. < av. perm.	Mean p-value	Nb of p-val. < 0.05
	min	mean	max	min	mean	max			
Average ISEI Differences in Co-Sitting Ties									
Average ISEI	18.56	21.54	25.41	22.36	23.67	25.28	35	0.31	6
Highest ISEI	16.03	20.33	23.81	20.52	22.3	25.48	33	0.37	6
Average ISEI Ranges in Groups									
Average ISEI	18.09	30.48	45.16	21.5	31.54	41.14	24	0.48	4
Highest ISEI	18.86	29.83	43.42	20.44	30.5	38.85	21	0.46	1

Note: Reported numbers are the observed values (min, mean, max), the average of values in permuted partitions (min, mean, max), the number of observations below the average of permutations, the average of empirical p-values, and the number of p-values below 0.05.

Figure 3: Violin Plots for Permuted Partitions at Every Meal Compared to Observations (Red Square)



Appendix 8D: Complete SAOM Results

Table 1: Estimated SAOM Parameters for Models 1 to 3 (Friendship Networks)

	Model 1			Model 2			Model 3		
	est.	sig.	s.e.	est.	sig.	s.e.	est.	sig.	s.e.
Rate period 1	18.12	***	(1.67)	18.76	***	(1.83)	20.47	***	(3.02)
Rate period 2	7.8	***	(0.48)	7.81	***	(0.5)	8.31	***	(0.56)
Density	7.94		(5.62)	7.63		(4.68)	7.7	*	(3.87)
Reciprocity	4.73	***	(1.04)	4.8	***	(0.83)	4.41	***	(0.88)
GWESP I→K→J	0.39		(0.47)	0.37		(0.38)	0.56		(0.34)
GWESP I→K←J	3.14	**	(1.18)	3.17	***	(0.82)	2.67	***	(0.74)
Recip x GWESP I→K→J	-2.3	***	(0.64)	-2.35	***	(0.5)	-2.1	***	(0.62)
Indegree Popularity	0.14	**	(0.05)	0.14	**	(0.04)	0.13	***	(0.04)
Outdegree Popualrity (sqrt)	-1.67	**	(0.51)	-1.69	***	(0.37)	-1.51	***	(0.32)
Indegree Activity	0.45		(0.32)	0.44		(0.24)	0.55	**	(0.21)
Gender receiver	-0.22		(0.12)	-0.22		(0.11)	-0.16		(0.11)
Gender sender	0.36		(0.31)	0.36		(0.29)	0.52		(0.31)
Same gender	0.67	***	(0.12)	0.67	***	(0.11)	0.48	***	(0.11)
Age receiver	0.01		(0.06)	0.01		(0.06)	0		(0.05)
Age sender	-1.03		(0.56)	-1.01	*	(0.44)	-1.07	**	(0.39)
Similarity Age	0.43		(0.23)	0.44		(0.23)	0.39		(0.23)
ISEI receiver	0		(0)	0		(0)	0		(0)
ISEI sender	-0.02		(0.01)	-0.02	*	(0.01)	-0.02	*	(0.01)
Similarity ISEI	0.24		(0.2)	0.11		(0.27)	0.31		(0.2)
Similarity ISEI (endowmenty)				0.99		(1.1)			
Same bedroom							1.32	***	(0.24)
Same activities (period 2)							0.29		(0.18)
Time dummy density	-1.31		(1.03)	-1.24		(1.16)	-0.95		(1.36)
Time dummy indegree activity	-0.46		(0.42)	-0.45		(0.35)	-0.58	*	(0.29)
Time dummy gender sender	1.9		(1.10)	1.91		(0.96)	1.82	*	(0.83)
Maximum convergence ratio		0.17			0.19			0.21	
Joint test time heterogeneity		p = 0.056			p = 0.046			p = 0.027	

***: p<0.01; **: p<0.05; *: p<0.1

Figure 1: Goodness-of-fit for Model 1

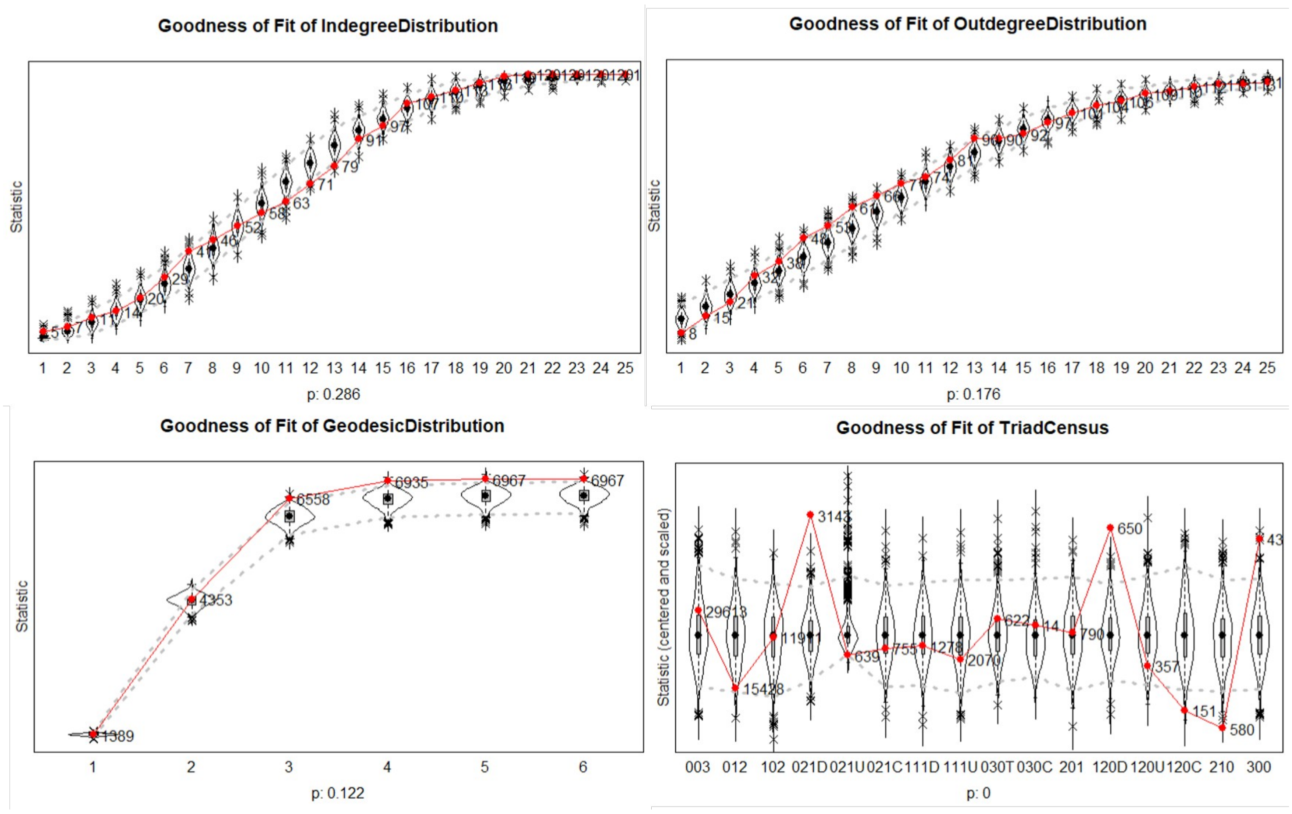


Figure 2: Goodness-of-fit for Model 2

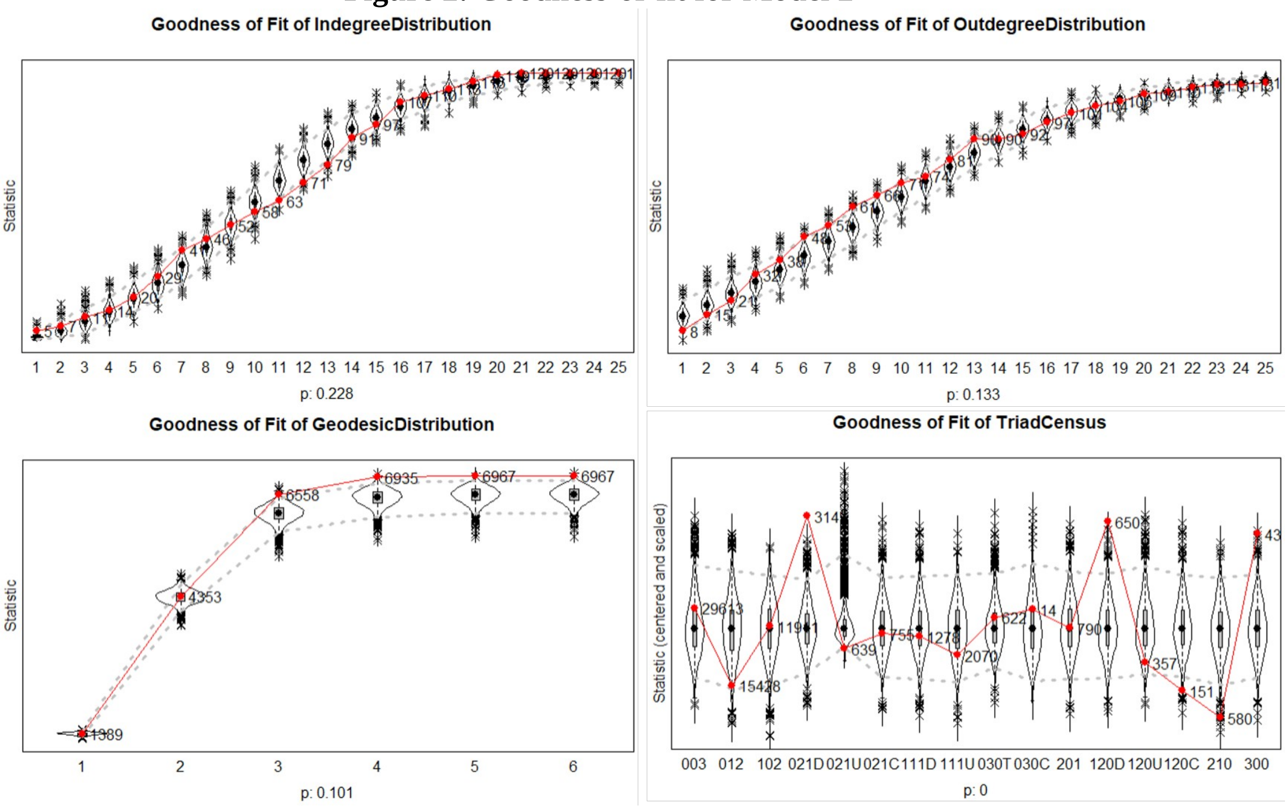
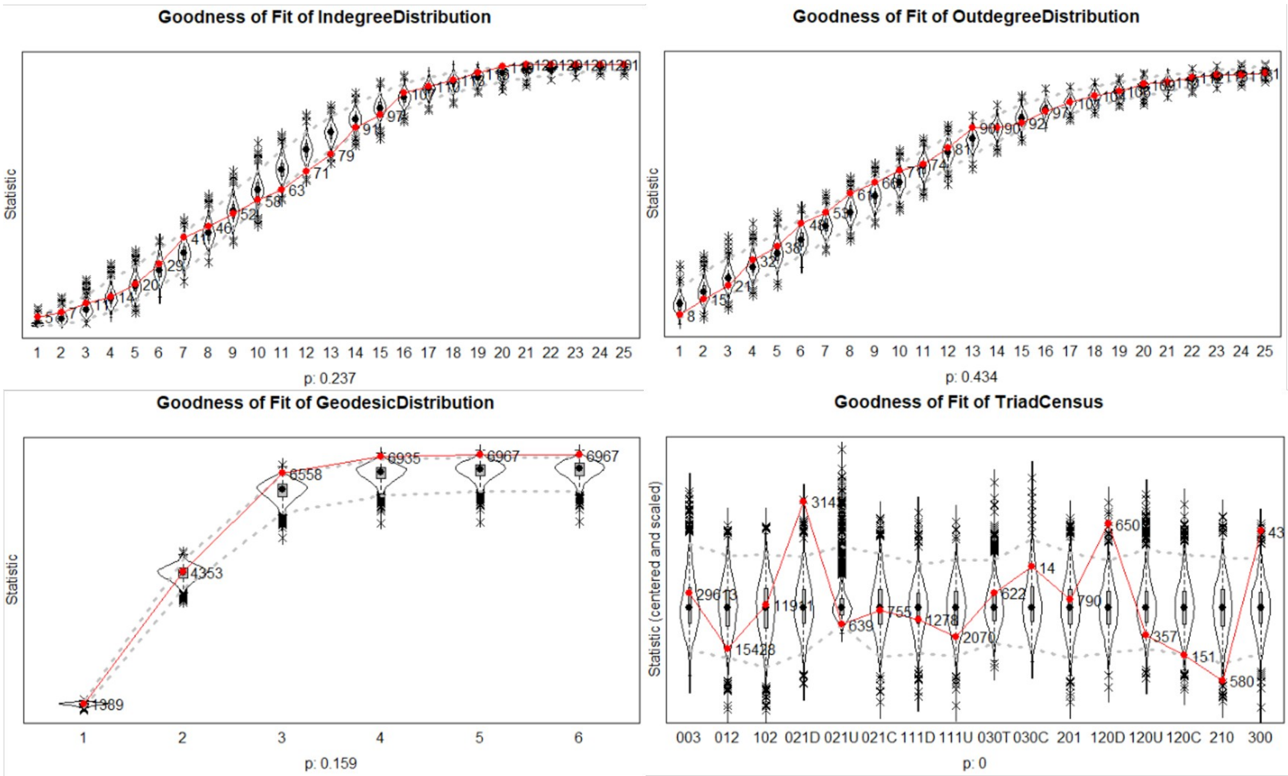


Figure 3: Goodness-of-fit for Model 3



Appendix 8E: Complete ERPM Results

Table 1: Estimated ERPM Parameters for Model 1 (Meal Partitions)

	est.	sig.	s.e.	Obs. Statistic	Exp. Statistic	Conv. Ratio
Number of groups	-6.39	***	(1.11)	87	71.01	0.0023
Sum squared sizes	-0.1		(0.33)	1717	1712.36	-0.082
Number isolates	4.14	***	(0.87)	20	19.02	-0.22
Girls popularity	0.05		(0.05)	986	991.04	0.14
Same gender ties	-0.23	*	(0.09)	609	605.15	-0.15
Same gender groups	3.49	***	(0.61)	56	55.68	-0.073
Higher age popularity	0		(0.024)	21414	21412.91	-0.0016
Age differences	-0.08		(0.024)	691	692.22	0.031
Age ranges	-0.48	*	(0.24)	117	116.68	-0.050
Higher ISEI popularity	0		(0.001)	94054.17	94396.66	0.11
ISEI differences	0		(0.003)	14834.86	14868	0.041
ISEI ranges	-0.02		(0.013)	2462.76	2654.96	0.079
Number meals	0.11	***	(0.01)	4859	4922.90	0.43
Anterior friendship ties	0.32		(0.17)	89.5	94.64	0.69

Table 2: Estimated ERPM Parameters for Model 2 (Meal Partitions)

	est.	sig.	s.e.	Obs. Statistic	Exp. Statistic	Conv. Ratio
Number of groups	-6.26	***	1.07	87	87.22	0.051
Sum squared sizes	-0.08		0.34	1717	1710.28	0.12
Number isolates	4.05	***	0.85	20	19	-0.23
Girls popularity	0.04		0.05	986	990.26	0.10
Same gender ties	-0.22	*	0.10	609	603.71	-0.21
Same gender groups	3.46	***	0.61	56	55.89	-0.024
Higher age popularity	0		0.03	21414	21391.55	-0.031
Age differences	-0.07		0.06	691	691.65	0.016
Age ranges	-0.48	*	0.24	117	117.04	-0.0057
Higher ISEI popularity	0		0.002	94054.17	94243.14	0.059
ISEI differences	0		0.003	14834.86	14816.23	-0.022
ISEI ranges	-0.02		0.01	2462.76	2648.03	0.046
Number meals	0.09	***	0.02	4859	4910.92	0.35
Anterior friendship ties	0.35	*	0.17	89.5	95.11	0.75
Number meals * ISEI diff.	0		0.00008	98473.81	99666.42	0.34

Table 3: Estimated ERPM Parameters for Model 3 (Meal Partitions)

	est.	sig.	s.e.	Obs. Statistic	Exp. Statistic	Conv. Ratio

Number of groups	-5.07	***	1.09	87	86.87	-0.030
Sum squared sizes	-0.08		0.33	1717	1716.82	-0.029
Number isolates	3.48	***	0.86	20	19.14	-0.20
Girls popularity	0.07		0.05	986	990.61	0.10
Same gender ties	-0.24	**	0.09	609	607.08	-0.071
Same gender groups	3.14	***	0.60	56	55.76	-0.053
Higher age popularity	0		0.02	21414	21467.77	0.070
Age differences	-0.07		0.05	691	692.73	0.040
Age ranges	-0.49	*	0.23	117	117.00	0.000078
Higher ISEI popularity	0		0.002	94054.17	94997.4	0.13
ISEI differences	0		0.004	14834.86	14934.66	0.12
ISEI ranges	-0.01		0.01	2462.76	2657.19	0.12
Number meals	0.08	***	0.02	4859	4914.69	0.36
Anterior friendship ties	0.08		0.17	89.5	95.01	0.66
Number meals * ISEI diff	0		0.0007	98473.81	100028.08	0.42
Same bedroom ties	0.66	***	0.10	256	260.88	0.33

***: p<0.01; **: p<0.05; *: p<0.1

Figure 1: Goodness-of-fit for Model 1

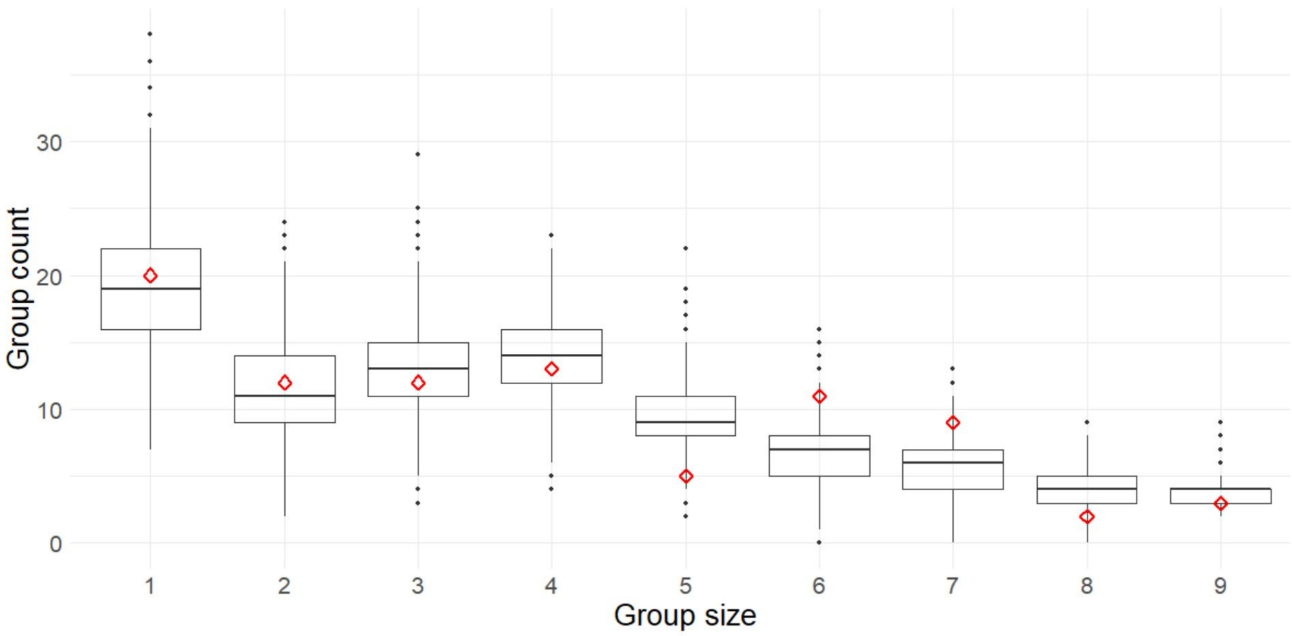


Figure 2: Goodness-of-fit for Model 2

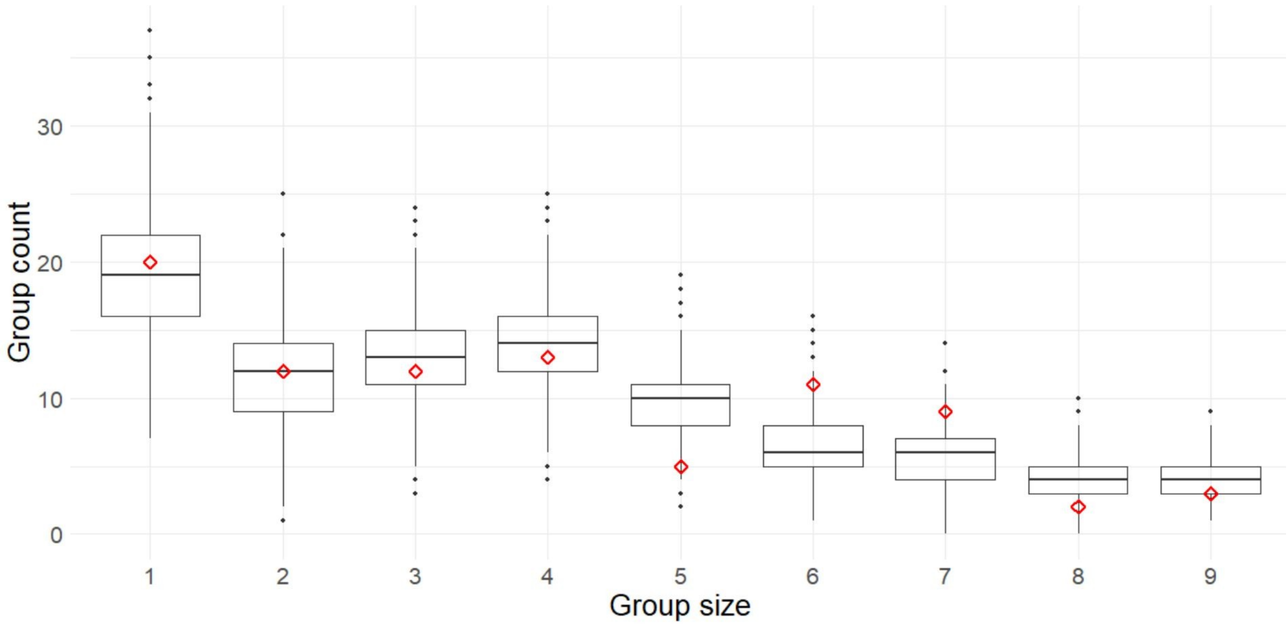
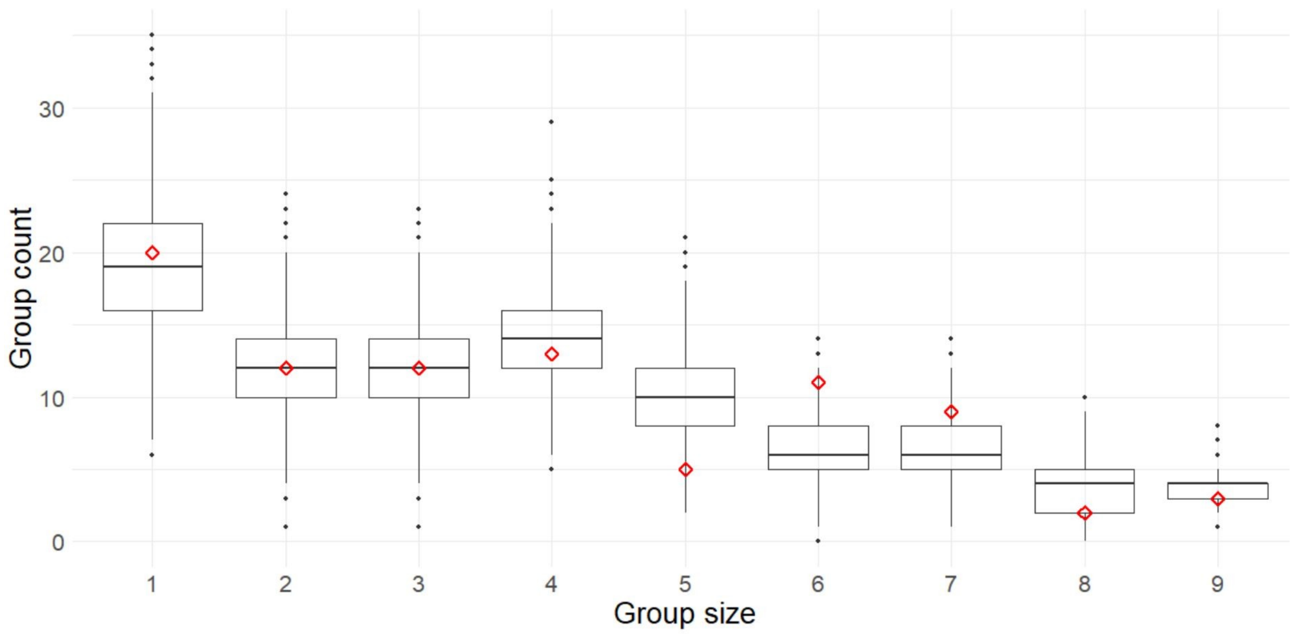


Figure 3: Goodness-of-fit for Model 3



Appendix 8F: Analyses for the 2017 Dataset

NB: there is no friendship network during camp for the 2017 data (it was not collected). Available data are the meal partitions, as well as the network of anterior friendships (recollection by attendees during the camp).

Table 1: Descriptives

Gender		Age				Average ISEI			
Nb girls	Nb boys	Min	Mean	Max	Sd	Min	Mean	Max	Sd
27	33	10	12.98	15	1.32	14.64	54.32	85.13	20.18

Table 2: Permutation Tests for Previous Friendships

	Anterior Friendships		
	Obs.	Av. Perm.	P-val.
Average ISEI	22.67	24.07	0.57
Highest ISEI	22.27	23.21	0.63

Table 3: Permutation Tests for Meal Partitions – All Meals Together

	Average ISEI Differences			Average ISEI Ranges		
	Obs.	Av. Perm.	P-val.	Obs.	Av. Perm.	P-val.
Average ISEI	22.01	23.24	0.11	47.56	47.49	0.97
Highest ISEI	23.93	24.11	0.73	50.07	49.85	0.96

Table 4: Permutation Tests for Meal Partitions – Permutations of the 34 Meals Separately

	Observed			Average in Permutations			Nb of Obs. < av. perm.	Mean p-value	Nb of p-val. < 0.05
	min	mean	max	min	mean	max			
Average ISEI Differences in Co-Sitting Ties									
Average ISEI	18.66	21.98	24.66	22.13	23.49	25.63	29	0.39	3
Highest ISEI	21.28	23.89	26.33	22.14	24.33	26.78	23	0.52	0
Average ISEI Ranges in Groups									
Average ISEI	37.66	48.12	55.82	40.17	47.47	53.14	18	0.54	1
Highest ISEI	42.09	50.73	57.85	43.25	49.96	56.36	15	0.55	0

Note: Reported numbers are the observed values (min, mean, max), the average of values in permuted partitions (min, mean, max), the number of observations below the average of permutations, the average of empirical p-values, and the number of p-values below 0.05.

Figure 1: Histograms of the Average ISEI Differences and Average ISEI Ranges in Groups in the Permuted Partitions

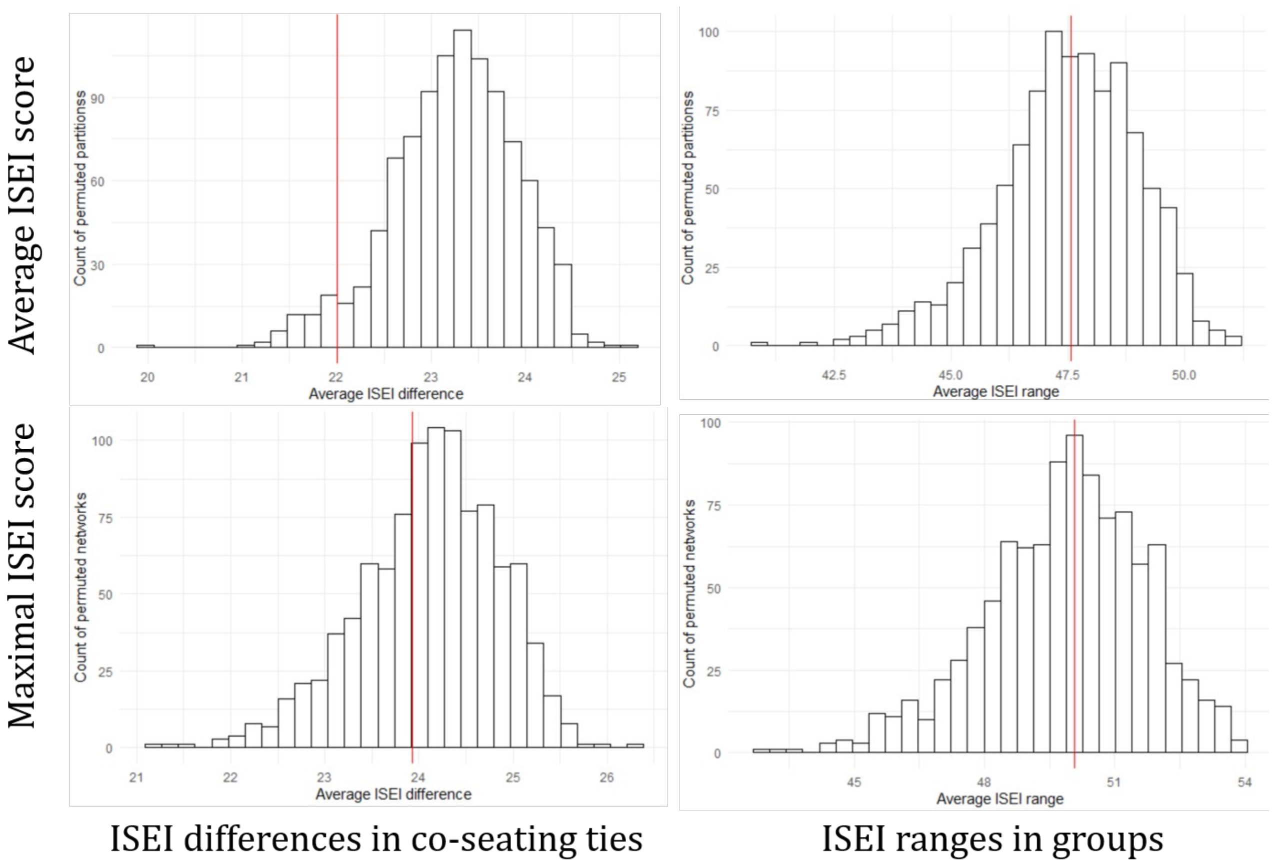


Figure 2: Violin Plots for Permuted Partitions at Every Meal Compared to Observations (Red Square)