DEPARTMENT OF ECONOMICS

Essays in Applied Microeconometrics: Fertility, Nutrition, and Gender Representation

Flavia Cavallini

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Economics
of the European University Institute

Florence, 10 June 2022

European University Institute
Department of Economics

Essays in Applied Microeconometrics: Fertility, Nutrition, and Gender Representation

Flavia Cavallini

Thesis submitted for assessment with a view to
obtaining the degree of Doctor of Economics
of the European University Institute

Examining Board

Prof. Thomas Crossley, EUI, Supervisor
Prof. Alessandro Tarozzi, EUI, Co-Supervisor
Prof. Nadia Campaniello, University of Torino and Collegio Carlo Alberto
Prof. Emilia Del Bono, ISER, University of Essex

Researcher declaration to accompany the submission of written work Department Economics - Doctoral Programme

I Flavia Cavallini certify that I am the author of the work Essays in Applied Microeconometrics: fertility, nutrition, and gender representation 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 acknowledgement is made. This work may not be reproduced without my prior written consent. This authorisation does not, to the best of my knowledge, infringe the rights of any third party.

I declare that this work consists of 16,906 words.

Statement of inclusion of previous work:
I confirm that chapter 3 was jointly co-authored with Alice Dominici and Olivia Masi and I contributed 33% of the work.

Signature and date:

Flavia Cavallini

17/05/2022

Abstract

This thesis is composed of three independent essays in applied microeconomics. The first contributes to the field of labor and health economics and analyzes the effect of local unemployment rates on fertility rates, abortion rates, and the abortions to pregnancies ratio. The second chapter speaks to health and development economics literature, evaluating the impact of agricultural price spikes on farmers' nutrition, considering the case study of quinoa in Peru. The topic of the final chapter lies within the fields of gender and political economics and discusses the effect of gender representation within local governments on expenditure in social services. Even though the three chapters seem separate, all of them share my interest in gender and health economics, as well as causal estimation.

In Chapter 1, I analyze the effect of local unemployment rates on fertility rates, abortion rates, and the abortions to pregnancies ratio, combining population statistics and administrative data on induced abortions performed in Italy between 2004 and 2016. This is the first paper to causally investigate the effect of local economic conditions on abortion choice. Using a shift-share instrument measuring labor demand, I exploit demand-driven shocks to unemployment. A one standard deviation (sd) increase in unemployment induces a 0.9 sd decrease in the fertility rate, a 0.27 sd increase in the abortion rate, and a 0.35 sd increase in the abortion ratio. In percentage terms, these changes mean that a 1 percentage point increase in the unemployment rate brings about a 1.7% decrease in the general fertility rate, a 1.4% increase in the abortion rate, and a 1.8% increase in the abortion ratio. These effects are driven by women above 25 years old, and are particularly large in the 35-49 age group.

In Chapter 2, I consider the impact of food price changes on farmers' particular nutrition, as part of a discussion of the effect of preference shifts in the global North on welfare in the global South. Previous research has yielded contrasting results, while this question is increasingly relevant. The case of quinoa provides an ideal event study, where quinoa prices steeply increased from 2008 onwards, led by increasing

international demand. I study the effect of this price shock on the nutrition of Peruvian households in a difference in differences framework. Results point to a limited effect on nutritional outcomes: in the short- term, neither caloric intake nor diet quality significantly increases in quinoa-farming households and districts.

Chapter 3 investigates the effect of executive female representation on the provision of different social services, in the context of Italy. While Italy is a high-income country, many families still rely on women to take care of children, the elderly, and family members in need of assistance. We exploit a 2014 reform that mandated 40% gender quotas in the executive committees of municipalities with more than 3000 inhabitants. To account for confounding policies introduced at the same cutoff, we employ a difference-in-discontinuities empirical strategy. We find that while the policy was effective in increasing female representation, it did not have an impact on any category of social services expenditures.

To the people that changed my life,
first and foremost me.

Acknowledgements

This is the final stop of the longest and most exciting roller coaster I've ever ridden. These years have changed me forever, professionally and personally, and I am thankful to all that supported me throughout this journey.

First of all, I would like to express gratitude to the team of advisors that followed me in these years, at different points in time. I want to thank Thomas Crossley for being incredibly patient, supportive, and asking the right questions. Though his professional advice was of great value and I keep learning from his lectures on econometrics, what he taught me goes beyond that. Another big thank you goes to Alessandro Tarozzi, who embarked on this supervision journey with enthusiasm and provided many insightful comments and interesting discussions on academic life. I also thank Michele Belot for advising me through the many challenges and for believing in my ideas. I am indebted to Arthur Schram, who provided guidance during the early stages of the thesis, and my coauthors Alice Dominici and Olivia Masi for their collaboration and the lively discussions. I would also like to thank Alicia Adsera and all members of the EUI Microeconomics Working Group for their helpful comments. Furthermore, I thank everyone who made EUI, and especially Villa la Fonte, an incredible place to work in: the students willing to share their many great ideas and failures, the organizers of the working groups and other initiatives, and the Reps. Everyone who listened and did so much more specifically Antonella, Lori, Sarah, Lucia, Alessandro. Our community has a social capital that we should not give for granted, and that takes effort to maintain. Finally, I am indebted to my thesis committee members Nadia Campaniello and Emilia Del Bono, who reviewed my work with extreme care and made invaluable suggestions.

A huge thank you goes to all the people that made my time in Florence exceptional - the many friends I made along the way that made it feel like home. I was never so happy as when I got to cook a good dinner with or for you and share the joys and trials of life. To Giulia, my "wife" in work and daily life, I am so lucky to have finally met you after living on parallel paths for so long. To the boys Mustafa, David, Daniele, I am thankful for the stupid jokes and the memories I will cherish forever. To My, provider of the best hugs in town and great counselor, who taught me to trust my gut once in a while. To Fatima and Johanna, thank you for your friendship and the confidence boosts, from long calls to econometric advice. I am also grateful to Nikolaj, Pietro, Daniela for taking the time to listen and share their thoughts with me.

I thank the friends that supported me from a distance, in Rome and elsewhere. A special mention goes to Federica, whose advice has steered me in the right direction again and again, even when I did not want to hear it, and Sara who is always present and the best company regardless of the physical distance. I thank Nicola, my buddy in Economics from day one; Lorenzo, who always has some sarcasm ready and a fresh perspective on academia; Massimo, Simona, Luca, Flavia, and Elisa for their encouraging words. Additionally, I am indebted to two teachers that were very special to me: Annamaria Avena, who made study fun from the early elementary school days, and my music teacher Damiana Pinti, who pushed me out of my comfort zone musically, personally, and professionally.

Finally, I want to thank my roots, my family, that have provided me with countless resources and the gene of curiosity and perseverance. I am extremely grateful to my mother Luisa, the Literature Ph.D., and my father Roberto, the engineer, who transmitted to me their love for applied problems and the complexity of human behavior - a powerful mix that led me to become a social scientist. I thank the tribe of women with strong opinions in my family, who nourished my soul with kind words, a spoonful of criticism, and their example. In particular, I am grateful to my aunt Giulia for making this journey possible by contributing to my studies. I also thank the many uncles and cousins for their continuous support. Above all, I am thankful to my parents and my sister Chiara for their faith in me.

And to Giulio, who is always on my side. I cannot emphasize enough how important you have been throughout these years.

Table of contents

Chapter 1

Not the right time for children: unemployment, fertility, and abortion

1.1 Introduction

Economic recessions have a far-reaching impact on individuals, as lower income and widespread economic uncertainty influence their decisions in numerous ways.[1] Bleak economic prospects may thus affect fertility decisions such as contraception effort, the timing of childbearing and the voluntary interruption of a pregnancy. Understanding the role of economic factors in fertility choice bears substantial policy relevance for the dependency ratio of a society, public healthcare spending, and labor supply, among others. While several studies have considered the response of contraception effort and timing of childbearing, the same attention has not been dedicated to abortion behavior.

This paper extends our understanding of the importance of local economic conditions in fertility choices. I use administrative, population-level data on abortions and province-level vital statistics to obtain a panel of fertility and abortion rates for Italian provinces between 2004 and 2016. I then exploit the variation in unemployment rates that occurred during the Great Recession and the Sovereign Debt Crisis to investigate the relationship between unemployment, induced abortions, and childbearing.

1. A wide literature has established that economic downturns affect a number of outcomes, including but not limited to physical and mental health (Ruhm 2000; Avdic, Sonja, and Kamhöfer 2021; Black, Jackson, and Johnston 2022), human capital accumulation (Barrow and Davis 2012; Barr and Turner 2015), labor market outcomes (Schwandt and Wachter 2019; Bono and Morando 2021), crime rates (Bell, Bindler, and Machin 2018), marriage and divorce rates (Schaller 2013; González-Val and Marcén 2018).

This is the first paper to causally investigate the effect of local economic conditions on abortion choice. The analysis of abortion patterns cannot be separated from that of fertility, since the number of abortions ultimately depends on the number of pregnancies. As a result, this study adds to the literature by considering the joint response of births and abortions. In addition, I consider the abortion to pregnancies ratio as a dependent variable, which measures the propensity to abortion conditional on pregnancy. While the fertility and the abortion rate measure the incidence of live births and induced abortions among all women, the abortion ratio measures the incidence of abortions among pregnant women. Throughout the analysis, I restrict the sample to women in their childbearing years, between 15 and 49 years old, and of Italian nationality.[2]

Italy provides a setting with homogeneous legislation across regions where induced abortion is covered by the national health insurance, which limits potential confounders and presents a very different scenario compared to countries with substantial legislative heterogeneity such as the United States. In this context, the monetary cost of an abortion is limited to the travel costs to access abortion services, thus allowing to abstract away from the potential effect of (relative) changes in the price of abortions during economic downturns. Nonetheless, women still bear the social and personal cost of an abortion, and face hurdles in accessing abortion services due to a geographically heterogeneous supply. To the extent that they are time invariant, these non-monetary costs are captured by province fixed effects. In addition, results are robust to controlling for the time-varying regional share of gynecologists unavailable for abortion.

I employ unemployment rates as a proxy for both job loss and societal economic uncertainty, for instance through the perceived risk of unemployment. However, unemployment rates may also reflect changes in labor supply induced by fertility choices, perhaps related to unobserved changes in preferences (Hotz, Klerman, and Willis 1997). To isolate fluctuations in the unemployment rate driven by labor demand shocks, I use a shift-share instrument in the spirit of Bartik (1991). In particular, the instrument measures predicted local employment that is unrelated to changes in local labor supply and is based on a weighted average of employment levels across industries. This approach closely relates to recent studies using shift-share instruments to explore the relationship between unemployment rates and fertility (Schaller 2016; Aksoy 2016), or other outcomes such as child maltreatment (Brown, De Cao, et al. 2018). This instrument is a labor demand index that measures predicted local employment that is unrelated to changes in local labor supply. In this context, the identification assumption

2. I exclude both foreign residents and Italian residents born abroad.

is that national employment shocks are independent from relevant province-by-industry unobservables, following Borusyak, Hull, and Jaravel (2022).

I find that a one standard deviation increase in the unemployment rate decreases the fertility rate by 0.9 standard deviations (95% CI [-1.13, -0.78]) and increases the abortion rate by 0.27 standard deviations (95% CI [0.41, 0.12]); the propensity to abort conditional on being pregnant also increases by 0.35 standard deviations (95% CI [0.53, 0.23]). This implies that a standard recession increasing the unemployment rate by 5 percentage points translates into approximately 3 fewer births and 0.5 more abortions per 1000 women; conditioning only on pregnant women, the effect rises to 12 more abortions per 1000 pregnant women.[3] Therefore, the pro-cyclical behavior of fertility rates is the result not only of changes in planned pregnancy but also a higher incidence of abortions, which account for around one sixth of the observed decrease in fertility.

These effects are driven by women above 25 years old, and are particularly large in the 35-49 age group, while young women (15-24) are largely unaffected. Such differences in response are potentially mediated by a number of factors, including heterogeneity in childbearing intentions, parity, marriage rates, labor force participation, and career attainment.

From a theoretical standpoint, standard models of fertility suggest that the effect of a rise in the unemployment rate is ex-ante ambiguous. On the one hand, the negative income effect decreases desired fertility; on the other hand, unemployment lowers the opportunity cost of childbearing, and might thus increase desired fertility through a substitution effect (Becker 1991). The change in childbearing intentions will affect contraception behavior, subject to the budget constraint; pregnant women will then choose between childbirth or abortion.[4] Therefore, the unemployment rate might affect the abortion rate through the behavioral response of both contraception effort and the propensity to abort conditional on being pregnant. Additionally, unemployment spells are related to increases in domestic violence (Pallitto et al. 2013; Anderberg et al. 2016; Bhalotra et al. 2021), which might result in increased abortion demand. Moreover, individual characteristics such as age, socioeconomic status, and job characteristics might determine different opportunity costs of childbearing and different exposure to

3. These changes stand at around an 8% change from the mean following a 5 p.p. increase in the unemployment rate, for all three dependent variables.

4. While standard economic models often assume that agents have perfect control over their fertility, here I account for the stochastic nature of fertility, i.e. the possibility of unintended pregnancies. Moreover, I leave open the possibility that contraception effort may be sub-optimal or not perfectly effective. Therefore, women are faced with three subsequent choices: their childbearing intentions, contraception effort, and abortion or childbearing choice.

unemployment shocks. Therefore, aggregate fertility responses may mask heterogeneous behavior across groups.

A wealth of research has investigated the relationship between fertility choices and economic fluctuations, both at the aggregate and individual levels. Sobotka, Skirbekk, and Philipov (2018) review the literature investigating the effect of economic recessions on fertility in high-income countries. Although the evidence is not uniform, the majority of studies support a pro-cyclical relationship of births to economic fluctuations, though with relatively small and short-lived effects. Among a few, Goldstein et al. (2013), Hofmann and Hohmeyer (2013a), Schneider (2015), Comolli (2017), and Matysiak, Sobotka, and Vignoli (2021) show that fertility decreases with higher unemployment rates. Del Bono, Weber, and Winter-Ebmer (2012) estimate that female job loss due to plant closure reduces average fertility by 5–10% in the short and medium-term. By contrast, Schaller (2016) finds that a deterioration of women's labor market conditions is associated with increases in fertility while a deterioration of men's labor market conditions has the opposite effect, suggesting that female unemployment generates positive opportunity cost effects that offset the negative income effects. This discrepancy in results can be attributed to the different context, where plant closures lead to a potentially permanent decrease in income affecting a specific sub-population (Schaller 2016). Finally, research suggests that short-term parenthood intentions are negatively associated with societal economic uncertainty (Fahlén and Oláh 2018). Modena and Sabatini (2012) and Modena, Rondinelli, and Sabatini (2014) estimate that having a temporary job contract in Italy reduces childbearing intentions by about 15 percentage points for childless women and 10 percentage points for mothers, from an average 25%, and being unemployed has a similar effect. The present study adds to this literature by considering the joint response of births and abortion behavior, and brings additional evidence of the pro-cyclical behavior of childbearing.

In addition, this study relates to the literature investigating the determinants of abortion demand. Although access to abortion services is a widely discussed topic worldwide, research on the determinants of abortion remains scarce and we lack an understanding of the role of economic factors in this decision. Studies of abortion demand have mostly concentrated on changes in the cost of abortion, particularly in the United States, for instance due to legislative restrictions (Haas-Wilson 1996; Bitler and Zavodny 2001; Medoff 2007; Myers and Ladd 2020), abortion clinic closures (Fischer, Royer, and White 2018; Lindo et al. 2020), insurance coverage (Levine, Trainor, and Zimmerman 1996), or the diffusion of oral contraception methods (Ananat and

Hungerman 2012). Some research has explored the role of social welfare policies, such as income support for low-income women (Snarr and Edwards 2009) or child support enforcement (Crowley, Jagannathan, and Falchettore 2012). González and Trommlerová (2021) explore the effect of a universal child benefit in Spain, finding that it increased the birth rate both through an increase in conceptions and a decrease in abortions. The role of unemployment has been largely overlooked; it is included as one of many potential determinants of abortion in the classic models of Medoff (1997) and Blank, George, and London (1996), without accounting for its potential endogeneity with respect to fertility choice. Lima et al. (2016) find that the abortion ratios in 2010-2012 exceeded the predicted trend across several European countries, indicating the economic recession and austerity policies as potential determinants. However, this study only relies on time variation, and the authors cannot make a causal statement nor address the potential heterogeneous response across countries that experienced the recession with different intensity and timing. In this paper, I overcome these issues by taking a causal approach and focusing on a single country, Italy, considering local economic conditions within Italy and using a shift-share instrument that allows reaching causal statements.

The remainder of the paper is organized as follows: Section 1.2 describes the institutional framework of abortions in Italy; Section 1.3 presents the data; Section 1.4 illustrates the empirical strategy. Section 1.5 presents the main results, while Section 1.6 covers robustness checks. Section 1.7 concludes.

1.2 Institutional framework

Abortion in Italy is regulated by Law 194 of 1978 and since then the Istituto Superiore di Sanità (Italian National Institute of Health) has maintained a surveillance system for legally induced abortions, based on quarterly reporting by the regional health authorities.

According to Law 194/1978, all women are eligible to request the voluntary interruption of a pregnancy during the first 90 days of gestation. Beyond this 90 days limit, only therapeutic abortions are permitted, i.e. abortions motivated by medical concerns. A woman seeking an abortion must first obtain a certificate attesting to the pregnancy from either her general practitioner, a private physician, or a public family clinic; parental or judge's consent is required for minors. With the exception of urgent cases, there is a mandatory seven-day period of reflection after the certificate. Induced abortions can be performed either in public hospitals, free of charge, or in authorized private clinics; more than 90% take place in public hospitals (Ministero della Salute 2016).

Article 9 of Law 194/78 regulates the practice of conscientious objection, granting the healthcare personnel the right to refuse to partake in procedures aimed at the termination of a pregnancy, except when these are deemed as life-saving. In 2016, 71% of gynecologists were objectors[5], over 8% of all abortions were sought by women out of their region of residence, and around 13% out of their province of residence (Ministero della Salute 2016). Autorino, Mattioli, and Mencarini 2020 show that a higher prevalence of objecting professionals is associated with a higher share of women having an abortion outside the region and longer waiting times. To account for this interregional mobility, I count abortions based on the province of residence of the woman, rather than the province where they occurred.

Moreover, differences in conscientious objection lead to considerable geographical heterogeneity in the supply of abortion services. Coverage is highest in Umbria, Liguria, and Toscana, where more than 90% of structures with a gynecology ward offer induced abortion in 2016, and lowest in Campania, Bolzano, and Molise, where less than 35% do (Ministero della Salute 2016). In this study, systematic time-invariant differences in the supply of abortion services across provinces are captured by province dummies, and in the robustness analysis I account for potential changes in the preferences of doctors over time.

Over the years, several regulatory changes have occurred: the legalization of emergency contraception pills in 2000 and 2012[6]; the introduction of medication abortion in 2009; the availability of emergency contraception pills without a medical prescription from 2015. The role of these regulatory changes as potential confounders is addressed in Section 1.6.

1.3 Data and descriptive statistics

In this section, I describe the main sources of data used in this paper and how I select my sample. Moreover, I define the dependent variables and discuss their evolution over time. Further details on the construction of variables are provided in Appendix 1.D.

5. Additionally, in 2016 49% of anaesthesiologists and 44% of non-medical staff were objectors.

6. Emergency contraception pills were legalized in Italy in 2000 if based on Levonorgestrel, and 2012 if based on Ulipristal acetate.

1.3.1 Data and construction of the sample

I employ yearly data on Italian provinces between 2004 and 2016 provided by the Italian National Institute of Statistics (ISTAT). The geographical unit of the analysis is a province (NUTS III division), with boundaries fixed to 2004 to keep geography constant.[7] The final sample includes 1339 observations, corresponding to a balanced panel of 103 provinces across the years 2004-2016.

To construct province-level rates, I combine two sources of data: population statistics at the province level and administrative data on induced abortions. Population statistics include data on population and live births from the General Register Office, disaggregated by the province of residence and age of the mother. Data on abortions pertain to administrative data on voluntary interruptions of pregnancies collected by ISTAT.[8] This covers all legal induced abortions performed in Italy, both in public and private facilities. For each procedure, the medical staff compiles a standardized form (module D.12) with details on the procedure and socio-demographic characteristics of the woman. In particular, this form records the age, citizenship, area of residence, marital status, and reproductive history of the woman. Appendix 1.A.3 discusses measurement error in the abortion data and presents several robustness checks.

I focus my analysis on women of Italian nationality and born in Italy between 15 and 49 years old, i.e. in their childbearing years. Women of foreign nationality are excluded from the final sample because they might react differently to economic factors, due to differences in socio-demographic composition, cultural values, access to health services (Spinelli et al. 2006), and labor market exposure.[9] In fact, the abortion rate of foreign citizens is three times higher than that of Italian women (Ministero della Salute 2016). In addition, I consider only interruptions of pregnancies taking place before

7. Seven new provinces were established during the observation window: 4 in 2006 and another 3 in 2010. I absorb each of these new provinces into their parent province; the correspondence is one-to-one except for Olbia Tempio, which I assign to Sassari since it was composed of 24 municipalities from Sassari and only 2 from Nuoro. Therefore, from an initial sample of 110 provinces, I remain with 103 provinces. In 2009, seven municipalities moved from the Pesaro to the Rimini province; results are robust to dropping these provinces, as reported in Table A1.

8. Data analysis was conducted at the Laboratory for Elementary Data Analysis (Laboratorio per l'Analisi dei Dati Elementari) of ISTAT, in compliance with legislation concerning the protection of statistical secrecy and personal data.

9. I also exclude foreign-born women that have acquired Italian citizenship, which constitute around 6% of the sample of Italian women in the abortion data.

the statutory 90 days limit, since abortions performed after this date should respond primarily to medical concerns.[10]

1.3.2 Dependent variables

This study considers three dependent variables: the general fertility rate, the abortion rate, and the abortion ratio. The general fertility rate (henceforth GFR) measures the average number of births in a year for every 1000 women who are in their childbearing years, i.e. between 15 and 49 years old. Similarly, the abortion rate indicates the incidence of abortions in the population of reproductive-aged women, i.e. the number of abortions per 1000 women in their childbearing years by the province of residence. Focusing on the province of residence rather than the province of abortion allows establishing a closer connection between local economic conditions and the abortion choice, on top of accounting for cross-province migration to access abortion services (see Section 1.2). A rise in the abortion rate can be the result of both a higher pregnancy rate (as a result of reduced contraception, for instance) or a larger share of unwanted pregnancies, keeping the pregnancy rate constant. For this reason, I also consider the abortion ratio, which measures the propensity to abort conditional on being pregnant, thus capturing also changes in the pregnancy rate. This ratio is computed as the share of abortions over pregnancies, where the number of pregnancies is proxied by the sum of live births and abortions.[11] As a result, the abortion ratio is increasing in the abortion rate and decreasing in the general fertility rate, as follows:

$$Ab.ratio = \frac{Abortions}{Pregnancies_{000}} \approx \frac{N.Ab}{N.Ab + N.Livebirths} = \frac{Ab.rate}{Ab.rate + GFR} \tag{1.1}$$

These aggregate measures are affected by the age structure of the population; I therefore replicate the analysis using age-specific rates. For example, I construct group-specific

10. Abortion regulation provides for the interruption of a pregnancy after 90 days from conception only in case of serious risk to the woman's life or severe fetal malformation. If abortions performed after the statutory 90 days limit reflected only medical concerns, they would provide an interesting placebo group. However, it is also possible that part of these abortions are related to economic conditions, as previous literature has established an increase of miscarriages (Bruckner, Mortensen, and Catalano 2016) and better infant health outcomes (Dehejia and Lleras-Muney 2004) in times of high unemployment rates. What effect economic fluctuations might have on fetal malformations and mothers' medical state remains an open question for future research, although out of the scope of this analysis.

11. I define the abortion ratio as the number of abortions per 1,000 pregnancies, in line with the definition by the Guttmacher Institute, while the CDC defines it as the number of abortions per 1,000 live births.

fertility rates by dividing the number of births by the appropriate at-risk population, i.e. women aged 15-49 in the relevant demographic group.

Finally, I use several indicators to proxy for local economic conditions. My main variable of interest is the aggregate unemployment rate, but I also consider age-specific unemployment rates. To address concerns of endogeneity of unemployment rates, I construct an instrument based on supply-side employment data by industry from the regional accounts, as described in Section 1.4.

1.3.3 Descriptive statistics

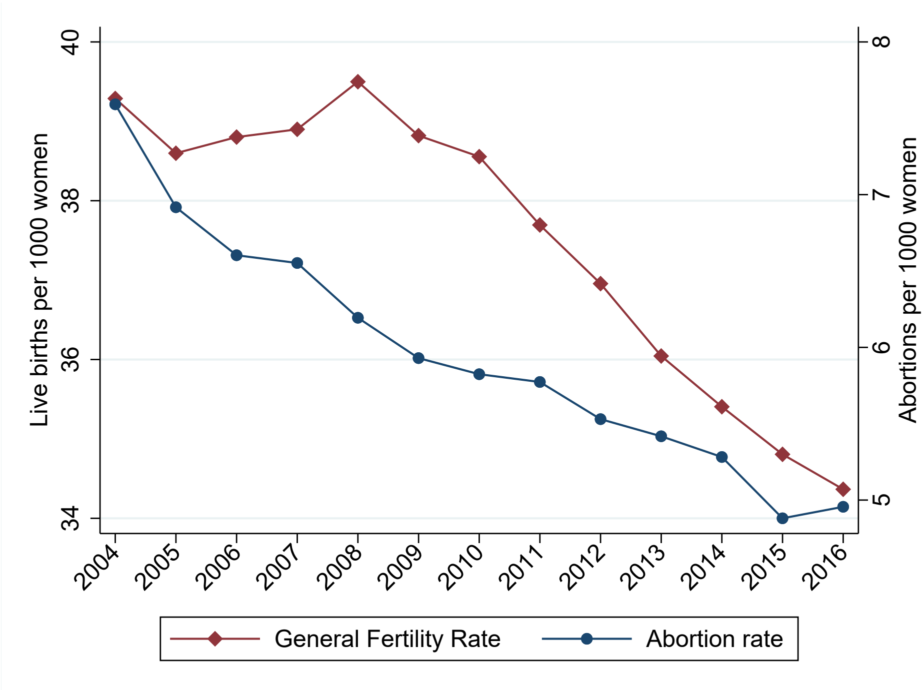
The evolution of the main variables of interest at the national level is presented in Figure 1.1. The general fertility rate, reported in Panel 1.1a, is relatively stable in the first half of the sample and reaches its peak in 2008, at around 39 live births per 1000 women of childbearing age. It then starts to rapidly decrease, down to 34 live births per 1000 women of childbearing age in 2016. The abortion rate is decreasing throughout the observation period, but at a slower rate from 2009 onwards; by 2016, it stands at around 5 abortions per 1000 women in their childbearing years. Panel 1.1b reports the evolution of the abortion to pregnancies ratio and the unemployment rate. The abortion ratio initially decreases steadily, but it almost flattens out from 2008 as a result of the sudden decrease in childbearing and the slower decrease in abortions. Finally, the national unemployment rate shows substantial variation over time, confirming the years between 2008 and 2014 to be a period of prevailing economic instability and underlining the double-dip nature of the recession. Given the different nature of the two crises, local labor markets were affected heterogeneously depending on their industrial composition.[12]

Table 1.1 reports the summary statistics for the province-level data. On average, the fertility rate is highest for women between 25 and 35 years old, but this age class also corresponds to the highest abortion rate. On the other hand, once accounted for the number of pregnancies, the propensity to abort is much higher for the younger (15-24) and older (35+) women. This means that childbirth and abortion are relatively less frequent in the younger and older female population because these age classes have fewer pregnancies but, in the event of a pregnancy, they are more likely to abort. The mean unemployment rate throughout the sample stands at around 9%, with a standard

12. For an overview of changes in employment and the geographical concentration of industries, see Section 1.B.1 in the Appendix. Section 1.C of the Appendix further elaborates on the evolution of the crisis, and its different impact across geographical areas.

Figure 1.1: National trends

(a) Fertility and abortion rates



(b) Propensity to abort and unemployment

Table 1.1: **Summary statistics - province data**

	Mean	Median	SD
Fertility rate by age group, per 1000 women			
GFR	36.21	35.96	3.35
GFR 15-24	12.83	11.19	5.16
GFR 25-34	76.82	76.32	7.82
GFR 35-49	22.55	22.43	2.81
Abortion rate by age group, per 1000 women			
Ab.rate	5.43	5.28	1.57
Ab.rate 15-24	7.74	6.42	2.13
Ab.rate 25-34	8.35	7.60	2.28
Ab.rate 35-49	3.67	3.32	1.03
Abortion ratio by age group, per 1000 pregnancies			
Ab.ratio	129.83	127.51	33.14
Ab.ratio 15-24	350.98	356.16	94.88
Ab.ratio 25-34	91.84	90.56	26.08
Ab.ratio 35-49	132.66	126.09	40.78
Unemployment rate by gender			
Unempl	9.37	8.12	5.26
Unempl_f	11.39	9.65	5.96
Unempl_m	8.04	6.77	5.18
Unemployment rate by age group			
Unempl 15-24	29.20	28.05	14.25
Unempl 25-34	13.35	11.08	8.60
Unempl 35-64	6.09	5.39	3.43
Additional economic indicators			
Empl. rate	57.88	62.30	9.79
Irregular workers, %	12.68	10.77	4.57
Real GDP per capita (000 €)*	22.13	22.36	5.84
Real Value Added per capita (000 €)*	19.89	20.06	5.17
Referendum on abortion law (1981)			
Yes votes, %	29.19	29.54	5.78
Observations		1,339	
Provinces		103	

Notes: The table provides within cell means for the 103 provinces used in the baseline specification. The share of irregular workers is measured at the regional level.

*GDP and Value Added are CPI adjusted to 2004 Euros.

deviation of 5 percentage points, and it is particularly high for females and younger workers. Moreover, the average employment rate is 57% and share of irregular workers is 12% on average, with a maximum of 25%. For summary statistics of the province-level data and the underlying abortion micro dataset, see Table 1.1 in the Appendix.

1.4 Empirical strategy

The goal of the analysis is to study the relationship between childbearing, abortion, and local economic conditions, proxied by the unemployment rate. To analyze different aspects of fertility choice, I consider multiple dependent variables: the general fertility rate, the abortion rate, and the abortion ratio. When studying fertility rates I lag the unemployment rate by one year to consider the unemployment rate in the year of conception.[13] All regressions control for province fixed effects and a linear time trend, which capture the confounding effect of unobserved time-invariant local characteristics and linear trends over time that are common to all provinces.[14] For instance, province dummies capture social attitudes towards abortion, i.e. the social cost of induced abortion, and time-invariant supply-side constraints in abortion services.[15] Therefore, the baseline specification is the following:

$$Fert_{p,t} = \alpha_0 + \beta_0 Unempl_{p,t-1} + \gamma_{0,p} + \delta_0 year + \epsilon_{0,p,t} \tag{1.2}$$

$$Ab_{p,t} = \alpha_1 + \beta_1 Unempl_{p,t} + \gamma_{1,p} + \delta_1 year + \epsilon_{1,p,t} \tag{1.3}$$

$$Ab.ratio_{p,t} = \alpha_2 + \beta_2 Unempl_{p,t} + \gamma_{2,p} + \delta_2 year + \epsilon_{2,p,t} \tag{1.4}$$

where $Fert_{p,t}$, $Ab_{p,t}$ and $Ab.ratio_{p,t}$ are the outcomes for province p in year t, $Unempl_{p,t}$ is the local unemployment rate, and γ_p are province dummies. Standard errors are clustered at the province level, and province boundaries are fixed to 2004.

The general fertility and abortion rates are affected by the underlying age composition of the population, as childbearing decisions significantly vary over a woman's lifecycle

13. By definition, three-quarters of the conceptions in year t will be realized in year $t+1$ due to the 9 months gestation period. Since monthly data on live births is not available, I approximate the time of conception with the year preceding the birth. Results are robust to considering the contemporaneous unemployment rate.

14. A sensitivity analysis of the results to alternative time specifications is reported in the Appendix 1.A.3, including location-specific time trends or a quadratic time polynomial.

15. Moreover, province fixed effects also account for time-invariant measurement error in both the dependent and independent variables. This applies to mismatches between the actual and official place of residence, particularly for students, where province dummies account for the average student outflows from provinces with fewer universities; the average cross-province employment mobility; and misreporting of births and induced abortions.

(Hotz and Miller 1988; Del Boca and Sauer 2009). Thus, if the female population of a province is predominantly young, we would expect the fertility rate to be higher than in a province with a predominantly old female population. Moreover, changes in the labor market heterogeneously impact different age groups; for instance, younger women might be more inclined than older women to postpone childbearing when facing adverse job prospects. To address this issue, I replicate the analysis focusing on age-specific fertility and abortion rates as dependent variables.

Using local unemployment rates as a proxy of local economic conditions allows capturing the effect of both individual job displacement and economic uncertainty, as the rate of unemployment correlates with the perceived risk of unemployment (Anderberg et al. 2016), the job separation rate, and it is negatively related to the job-finding rate. Although unemployment rates can understate the magnitude of a recession by not accounting for discouraged workers, they are the best available proxy to capture changes in the labor market conditions at this level of disaggregation. In addition, the unemployment rate is useful in exploring fertility behavior because it is less likely to be endogenous to childbearing or abortion choices than other variables such as own wages.

However, the results of the OLS estimation cannot be interpreted as causal for a number of reasons. First, there is an issue of reverse causality: an increase in fertility (abortions) can induce more women to drop out (stay in) the labor force (Bloom et al. 2009; Kalist 2004). Unemployment rates are therefore correlated with fertility-induced changes in the labor supply. Moreover, recent literature shows that fertility indicators lead economic recessions by several quarters (Buckles, Hungerman, and Lugauer 2021). Second, there are omitted variables that affect both fertility and unemployment, such as the age composition of the labor force or unobserved preferences (Hotz, Klerman, and Willis 1997). A simultaneous increase in childbearing intentions and a decrease in female labor force participation will increase the fertility rate while decreasing the denominator of the unemployment rate, thus leading to upward biased OLS estimates. Viceversa, a change in preferences that induces both a decrease in abortions and labor force participation will lead to downward biased OLS estimates.

To address these concerns of endogeneity, I construct an instrument for the local unemployment rate following the approach developed by Bartik 1991 and employed by Bound and Holzer (2000), Autor and Duggan (2003), Moretti (2013), and Schaller (2016). The instrument averages national employment across industries, using local industry employment shares as weights, to produce a measure of predicted local employment that is unrelated to changes in local labor supply. More specifically, this variable interacts predetermined differences in the industrial composition within a local labor market with national industry employment trends over time.

Formally, the instrument is constructed as

$$B_{p,t} = \sum_{k=1}^{K} \chi_{p,k,t_0} E_{-p,k,t} \tag{1.5}$$

where χ_{p,k,t_0} is the employment share of sector k in province p and base period t_0, and $E_{-p,k,t}$ is number of people employed in sector k, period t in Italy, excluding province p. Figure B1 in the Appendix maps the initial shares χ_{p,k,t_0}, which represent the importance of each sector for local employment. To construct age-specific instruments, I adjust the employment shifts by the contemporaneous employment share of each age group at the national level, respectively.Thus, the age-specific instrument will be [16]

This instrument identifies demand-driven changes in the unemployment rate: by keeping industry composition constant over time and exploiting temporal variation originating only from the national employment level, I abstract from changes in local labor supply and sorting into industries. Then, the instrument captures variation in unemployment that is driven by changes in the national economy, but varies across counties because of historical differences in the distribution of nationwide industry employment by location.

Recent methodological literature has underlined that identying variation in shift-share designs can stem from either the national shifts or the local shares (Jaeger, Ruist, and Stuhler 2018; Adao, Kolesár, and Morales 2019; Goldsmith-Pinkham, Sorkin, and Swift 2020; Borusyak, Hull, and Jaravel 2022). In the present setting, local industrial composition is likely to be endogenous, because local industrial composition determines opportunities for female employment and thus correlates with the unobserved labor supply shocks.[17] In addition, other unobserved shocks, such as the expansion of female-dominated industries, are likely to affect outcomes through the same mixture of exposure shares. I thus follow the interpretation suggested by Borusyak, Hull, and Jaravel (2022), where identification relies on the exogenous assignment of national industry employment shocks. This translates into assuming that national employment shocks are independent from relevant province-by-industry unobservables.

16. A similar group-specific weight is implemented by Schaller (2016); formally, the age-specific instrument corresponds to the following: $B_{a,p,t} = \frac{E_{a,IT,t}}{E_{IT,t}} \sum_{k=1}^{K} \chi_{p,k,t_0} E_{-p,k,t}$

17. These issues are particularly salient when exposure shares are recent: since industrial composition is persistent over time and the initial shares are only 4 years before the sample window (2000), future changes in fertility are reasonably related to the initial employment shares.

The validity of the instrument then hinges upon the atomic structure of the national industrial composition, and the assumption that single provinces are small enough within each industry not to affect its national employment level. This assumption however is likely to be violated by large provinces, and in fact the four most populous Italian provinces make up for more than 5% of national employment in a variety of sectors.[18] To address the finite sample bias coming from the use of own-observation information, I compute the national employment of each industry excluding own province employment. The first stage relationship is discussed in Section 1.B.2 of the Appendix, where figure B3 shows a strong negative correlation between the observed unemployment rate and the Bartik instrument. The relevance of the instrument is confirmed by the statistical significance of the point estimates of the first stage regression and the magnitude of the Kleinebergen-Paap F-statistic.

The instrument is computed using supply-side employment data on the number of workers participating to the production process in each industry province, thus not accounting for residents that work outside of the province but including non-residents that work in the province. This measurement error does not pose a threat to the analysis as long as the unobserved cross-province employment migration patterns are not correlated with childbearing or abortion.[19]

As a robustness check, I construct alternative versions of the instrument. First, I employ different sets of weights: the local employment share χ_{p,k,t_0} can be the number of employed in a sector and province over either total province employment or the working-age population of the province. Second, I manipulate the base year t_0: in the main estimation I use the first year out of sample (2003); alternatively, I compute it using the year 2000 as a base year. Finally, I consider industry value-added as an alternative shift; however, this measure is only available at the national level so the resulting instrument is not of the leave-one-out type. Figure B4 illustrates the differences between these alternative instruments, while details on their construction are reported in the Data appendix.

18. Rome, Milan, Naples, and Turin populate the right tail of the distribution of province weights in national industry employment. In particular, Milan and Rome represent more than 5% of all industries except agriculture and manufacturing, respectively. The industry with the highest weight is information and communications, where Rome and Milan make up for around 17% of national employment each.

19. To the extent that this measurement error is related to time-invariant factors such as population size, it is going to be captured by province fixed effects. Moreover, results are similar when measuring employment by industry and province of residence with data from the Labor Force Survey.

1.5 Results

Table 1.2 reports the main results of the paper for the three dependent variables of interest.

Columns (1) and (4) analyze the general fertility rate, reporting respectively the OLS and IV estimates. The OLS results indicate that the unemployment rate has a negative and statistically significant effect on fertility rates, where a one standard deviation increase in the unemployment rate reduces the GFR by 0.21 standard deviations. The IV estimation confirms the direction of the effect and yields larger estimates in magnitude: increasing the unemployment rate by one standard deviation translates into a reduction of the fertility rate by 0.95 standard deviations (95% CI [-1.13, -0.78]). The reported Kleibergen-Paap LM under-identification test and the Kleibergen-Paap F statistic strongly reject the hypothesis of a weak instrument; additionally, Table B2 of the Appendix reports the first stage estimates attesting to the relevance of the instrument.

Columns (2) and (5) report the results for the abortion rate. The abortion rate increases with the unemployment rate, and again the IV estimates are almost twice the OLS ones: a one standard deviation increase in the unemployment rate brings about a 0.25 standard deviation change in the abortion rate (95% CI [0.39, 0.12]).

Finally, columns (3) and (6) present the main results for the abortion to pregnancies ratio. The OLS coefficients indicate that a one standard deviation increase in unemployment is associated with a 0.18 standard deviations change in the propensity to abort conditional on being pregnant. This effect however more than doubles when we move to the IV estimation: a one standard deviation increase in unemployment is now associated with a 0.38 standard deviations change in the abortion ratio (95% CI [0.52, 0.23]). Overall, the propensity to abort conditional on being pregnant increases when unemployment rises, as a result of both the reduced number of births and the increased number of abortions.

Coefficient estimates from the IV analysis are larger, in absolute value, than the OLS results, consistently with the expected reverse-causality bias discussed in Section 1.4: lower fertility induces higher female labor supply, and thus lower unemployment rates, generating attenuation bias in OLS. Moreover, measurement error in unemployment rates could also be causing OLS coefficients to be biased downward (in magnitude).

Table 1.2: Main specification - standardized variables

	OLS			IV		
	GFR	Ab.rate	Ab.ratio	GFR	Ab.rate	Ab.ratio
	(1)	(2)	(3)	(4)	(5)	(6)
$Unempl_{t-1}$	-0.212***			-0.956***		
	(0.052)			(0.089)		
$Unempl_t$		0.144 ***	0.187***		0.253***	0.375***
		(0.032)	(0.038)		(0.069)	(0.074)
Observations	1236	1339	1339	1236	1339	1339
R^2	0.581	0.561	0.353			
KP LM p value				0.000	0.000	0.000
KP F-stat				360.5	285.5	285.5

Standard errors in parentheses. All regressions include province fixed effects and a linear time trend.
* $p < .05$, ** $p < .01$, *** $p < .001$

1.5.1 Age heterogeneity

Different age groups might experience different substitution effects because labor force participation, wage levels, career expectations, and the probability of having a stable partner change with age. Moreover, focusing on age groups allows disentangling the reaction of women at different points of their childbearing cycle, that have different childbearing intentions, contraceptive use patterns, and number of previous children. In this section I explore the heterogeneity of response by age, dividing women of childbearing age into three groups: from 15 to 24; from 25 to 34; and from 35 to 49 years old.

Previous literature on fertility has painted a varied picture: Ananat and Hungerman (2012), Goldstein et al. (2013), Schneider (2015), Comolli (2017) find childbearing of younger groups to be the most responsive to changes in the unemployment rate, an effect driven by first births. Schaller (2016) and Del Bono, Weber, and Winter-Ebmer (2012) instead find older age groups to be more responsive; Comolli (2017) finds women in their late thirties (35-39) to be the second most hit group. On the one hand, young women might be the most responsive, as youth unemployment was most affected by the recession (see Table 1.1) and fertility plans can be revised more easily at younger ages (Goldstein and Cassidy 2014); on the other hand, they are more likely to use contraception[20] Older women instead are more likely to be in a stable relationship[21]; however, economic recessions might affect both divorce and marriage rates (Schaller 2013; González-Val and Marcén 2017). Women between 25 and 34 are in the prime of their childbearing years and professional career, and therefore most likely to postpone childbearing. Changes in the behavior of this group will be particularly relevant because the incidence of births is highest in this age class (see Table 1.1). Women above 35 years old are closest to the end of their reproductive life, but they are also more likely to already have children and face a stronger trade-off between quality and quantity of children.[22] Finally, the recession might have brought about a reduction in births through lower recourse to assisted reproductive treatments such as IVF, as they are only partially covered by national insurance. Since these treatments are used predominantly by women older than 35, such changes are going to affect disproportionately the group of women above 35 years old[23].

Table 1.3 replicates the analysis relating age-specific dependent variables to the corresponding age-specific unemployment rates; Figure 1.2 presents graphically the estimates from Table A4. Age-specific unemployment rates reflect the conditions that women in each age group face in the labor market, particularly if the labor market is segmented.[24] For the IV estimation, I employ age-specific instruments that measure the predicted employment level for each age group, as described in Section 1.4.

Columns (1-3) report the results for age-specific fertility rates. The OLS estimates indicate that youth unemployment does not affect the fertility rate of women aged 15-24, as the estimated coefficient is close to zero and not statistically significant, while the response of older women's childbearing to changes of age-specific unemployment rates is negative and statistically significant at the 0.1% confidence level. The IV estimation uncovers an interesting relation: higher youth unemployment slightly increases the fertility rates of younger women, while higher unemployment for the middle-aged groups

20. In fact, survey data suggests that, conditional on being sexually active, the share of women using contraception is higher among women younger than 25 and to be out of the labor force.(Loghi and Crialesi 2017). Moreover, survey data suggests that women between 18 and 25 years old are the main users of emergency contraception (Bastianelli, Farris, and Benagiano 2005; Bastianelli et al. 2016).

21. Women in a stable couple are less likely to use contraception, even when not intentionally seeking a pregnancy (Loghi and Crialesi 2017).

22. In 2016, the average age of the mother at first birth was 32 in Italy(Istat 2017). Hofmann and Hohmeyer (2013b) show that couples with children respond significantly to economic concerns by reducing fertility, while childless couples do not.

23. Assisted reproductive treatment in Italy was first regulated in 2004 (L. 40/2004), at the beginning of my sample. In 2016, the average age of women resorting to assisted reproduction was 37, and these treatments accounted for 3% of live births in Italy (Ministero della Salute 2018)

24. Relating age-specific dependent variables to the overall unemployment rate yields similar results, reported in section 1.A.1 of the Appendix.

largely decreases their fertility rates.[25] Specifically, a one standard deviation increase in youth unemployment brings about a 0.15 standard deviation increase in the fertility rate of women aged 15-24, and this effect is statistically significant at the 5% confidence level. For women in the 24-34 age group, a one standard deviation increase in their unemployment rate decreases the general fertility rate by 0.52 standard deviations, and this effect is statistically different from zero at the 0.1% confidence level. The effect is even larger for women between 35 and 49 years old, where the GFR decreases by 1.5 standard deviations per one standard deviation increase in unemployment.

Columns (4-6) consider the relation between age-specific abortion rates and age-specific unemployment rates. The OLS analysis points to a null effect of unemployment rates on the abortion rate of young women, although not precisely estimated, and a limited positive effect for women older than 25 years old, which is statistically different from zero. The IV analysis confirms that unemployment rates do not have a statistically significant impact on the abortion choice of women in the youngest age group, while they have a positive and statistically significant effect on the abortion rates of older women. The IV estimates are larger in magnitude than the OLS ones and indicate that a one standard deviation increase of the respective age-specific unemployment rates brings about a 0.3 standard deviation increase in the abortion rate of women in the 24-34 age group, which rises to a 0.4 standard deviation increase for women older than 35.

Finally, columns (7-9) report the results for the age-specific abortion to pregnancies ratios. Again, the OLS estimation indicates that the behavior of younger women does not correlate with changes in the unemployment rate, while for older women there is a positive and statistically significant correlation between the age-specific unemployment rate and their propensity to abort conditional on being pregnant. The results of the IV estimation confirm the small effect for women in the 15-24 class, and the positive effect for older women, which follows naturally from the decrease in fertility rates and increase in abortion rates attested by the previous columns.

25. For estimates to be comparable across age groups they should refer to similar changes in unemployment, while here each refers to a one standard deviation change of the age-specific unemployment rate. However, here the standard deviation of the age-specific unemployment rates is decreasing across the age profile (see Table 1.1), while estimates are increasing over the age profile. Thus, translating the coefficients into a comparable unit change in unemployment would only magnify the result that the response of fertility and abortion increases across age categories. The fact that differences in response are not driven by differences in unemployment units of measurement is confirmed by the estimates relating the age-specific dependent variables to the total unemployment rate, presented in Section 1.A.1 of the Appendix.

Table 1.3: Age specific fertility rates - standardized

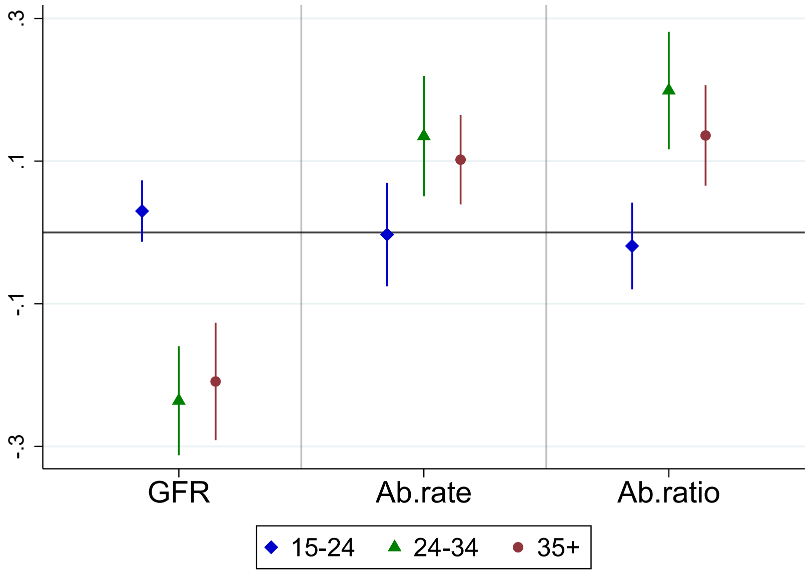
	GFR			Ab.rate			Ab.ratio		
	15-24	25-34	35-49	15-24	25-34	35-49	15-24	25-34	35-49
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
OLS									
$Unempl_{t-1}$	0.030	-0.236***	-0.209***						
	(0.022)	(0.039)	(0.042)						
$Unempl_t$				-0.003	0.135**	0.102**	-0.019	0.199***	0.136***
				(0.037)	(0.043)	(0.032)	(0.031)	(0.042)	(0.036)
IV									
$Unempl_{t-1}$	0.155*	-0.525***	-1.511***						
	(0.066)	(0.123)	(0.132)						
$Unempl_t$				0.084	0.291*	0.416***	-0.060	0.210	0.948***
				(0.081)	(0.139)	(0.097)	(0.083)	(0.132)	(0.118)
Observations	1236	1236	1236	1339	1339	1339	1339	1339	1339
R^2	0.368	0.279	0.066	0.477	0.286	0.469	0.121	0.193	0.474
KP LM p value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
KP F-stat	97.744	82.949	213.311	113.627	54.553	175.792	113.627	54.553	175.792

Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. The local unemployment rates are instrumented for using a leave-one-out Bartik instrument based on the number of employed individuals in each sector, using 2003 weights. The R^2 refers to the OLS estimation.

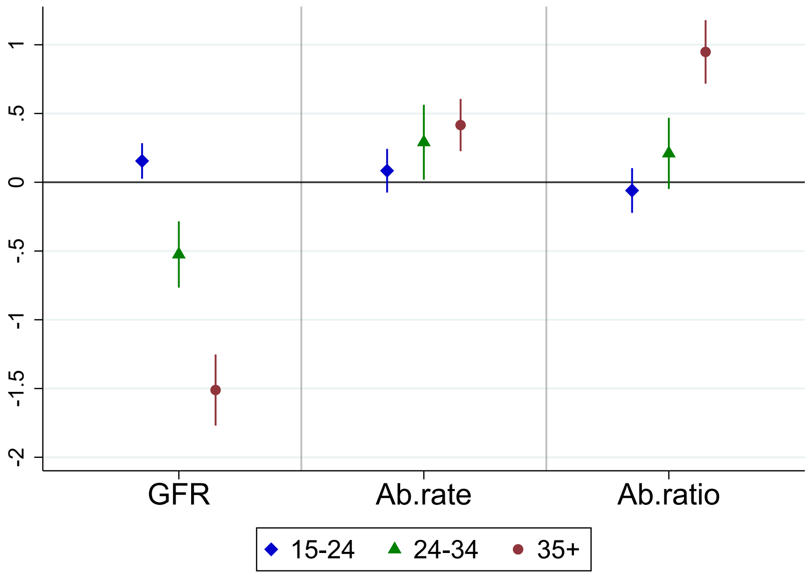
* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 1.2: Regression estimates - age heterogeneity

(a) OLS estimates



(b) IV estimates



Notes: 95% CI reported.

A one standard deviation increase in the respective unemployment rate increases the abortion ratio of women in the 25-34 age group by 0.2 standard deviations, as a consequence of the rise in pregnancies induced by the increase in births, but the estimate is not statistically different from zero. For women above 35 years old, their propensity to abort significantly increases by 0.95 standard deviations following a 1 standard deviation increase in their unemployment rate.

Overall, a common thread emerges from this analysis: unemployment rates have a statistically significant effect on childbearing and abortion, and their relevance for fertility choice is increasing over the life cycle. Notably, unemployment rates have a limited positive effect on young women's fertility and no statistically significant impact on the abortion rates of women in the youngest age group, while they decrease fertility rates and increase abortion rates of older women. The fact that the response of childbearing to changes in unemployments is increasing in magnitude across the age profile is consistent with Schaller (2016) and Del Bono, Weber, and Winter-Ebmer (2012). Since the average age at first birth is 32 in Italy (Istat 2017), the reduction of births from women aged between 25 and 35 is likely to translate into a mere postponement of births. The large response of women between 35 and 49 years old instead suggests a potential permanent reduction in fertility, since these women are closer to the end of their reproductive cycle. The change in births for the oldest age group also captures changes in assisted reproductive treatment; however, given the limited diffusion of IVF procedures[23], any changes in the use of these treatments are likely only a contributing factor to the observed decrease in births.

1.6 Robustness checks

This section discusses several robustness checks, presented graphically in Figures 1.3 and 1.4; the corresponding regression estimates are reported in Table A1 of the Appendix. Additional checks are reported in the Appendix, in particular addressing measurement error in the abortion data (see Section 1.A.3).

Figure 1.3 plots regression estimates for each of the dependent variables for a number of robustness checks, using respectively OLS and IV estimation. The graph first presents the baseline coefficients of Table 1.2 and then compares them to alternative specifications. The second specification focuses on a restricted sample excluding the provinces of Pesaro, Rimini, and the region of Puglia. This allows to more precisely keep geographical boundaries constant over time, since in 2009 seven municipalities belonging to these provinces changed province. In addition, I exclude the region of Puglia because in 2008 it introduced a policy of free hormonal contraception for specific groups, which might have affected abortion behavior. The third specification replicates the analysis of the abortion dependent variables controlling for the regional share of objecting gynecologists. While systematic time-invariant differences in the supply of abortion services across provinces are captured by province dummies, changes in the preferences of doctors over time are a potential source of variation in abortion supply. Indeed, the percentage of objectors in Italy has increased over time, from below 60% in 2002 to over 70% in recent years (Autorino, Mattioli, and Mencarini 2020). The robustness of results to these alternative specifications suggests that the results are not driven by changes in geography nor in the the availability of abortion services.

Figure 1.4 explores the sensitivity of results to alternative specifications of time. The linear time trend in the baseline specification removes the monotonic relationship between the outcome variables and time, thus leaving potential non-linear behavior in the residual term. The second specification considers province-specific linear time trends; since health services are administered at the regional level, provinces in different regions might have been following different trends.[26] The third specification considers a quadratic time trend, and last specification includes time dummies for the introduction of potentially confounding national policies: the availability of medication abortion from 2009, a baby bonus policy from 2013, the labor market reform, baby bonus policy, and availability of emergency contraception without prescription from 2015.[27]

Due to data limitations there is not sufficient within-period variation to sustain the analysis with yearly time dummies, particularly in the 2SLS estimation.[28] Notably, the point estimates are rather stable across specifications, though in some cases with a loss of precision.

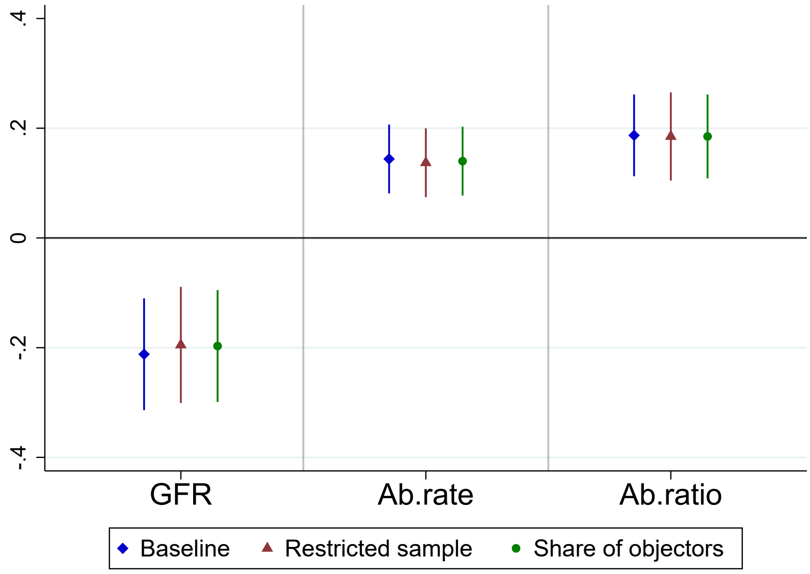
26. Similar results are obtained if accounting for region-specific time trends.

27. For a detailed description of these policies, see Section 1.A.3.

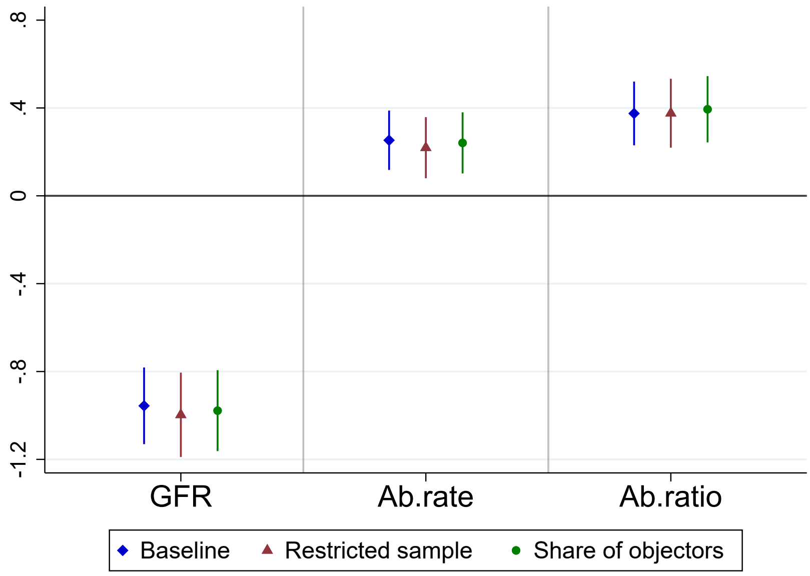
28. Furthermore, recent literature has underlined how the use of TWFE fundamentally rests on modeling assumptions, rather than being a nonparametric estimation strategy (Imai and Kim 2021). Though this research has focused on cases of binary treatment, the use of TWFE is likely to be just as problematic with continuous treatment such as the one under consideration here.

Figure 1.3: Robustness checks

(a) OLS estimates



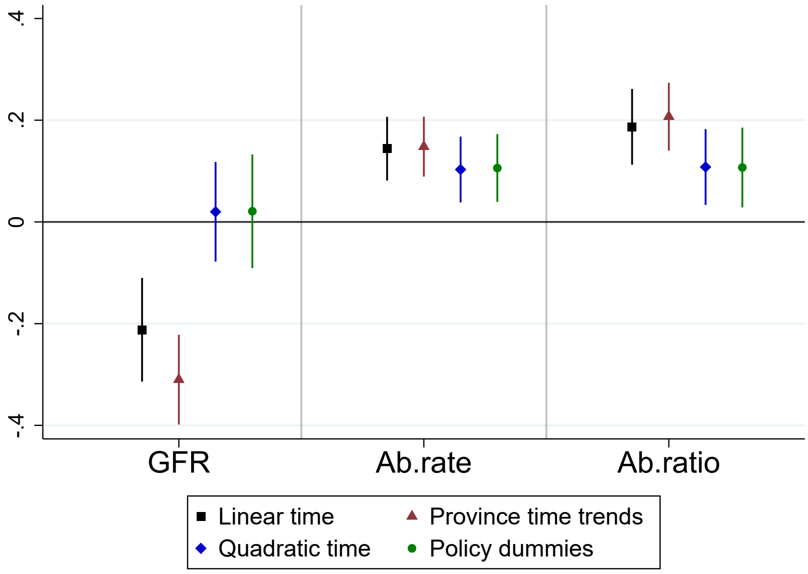
(b) IV estimates



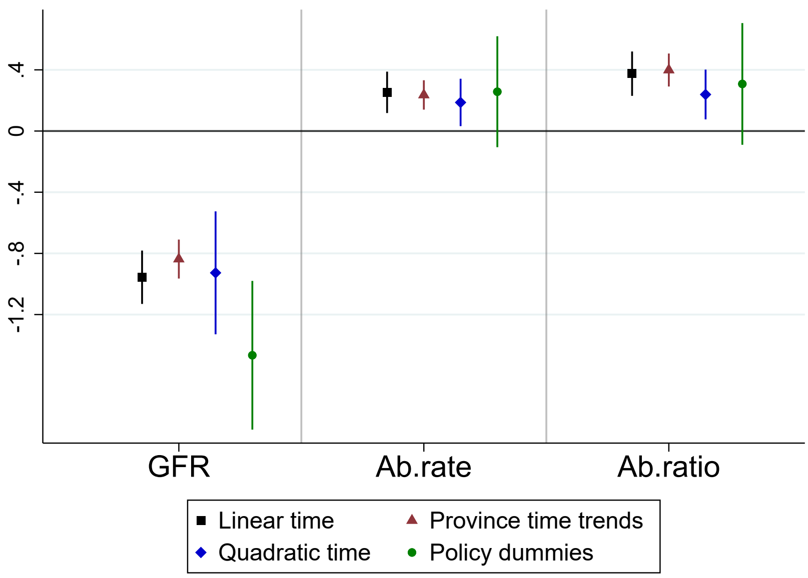
Notes: 95% CI reported.

Figure 1.4: Time specification

(a) OLS estimates



(b) IV estimates



Notes: 95% CI reported.

1.7 Conclusion

This paper explores the relationship between local unemployment rates and fertility choice, focusing on both childbearing and abortion behavior. The literature has been aware of the role of economic conditions in childbearing, suggesting that fertility behaves procyclically. However, the same attention has not been dedicated to understanding the response of abortion behavior, particularly using causal methods.

I investigate the response of births and abortions to changes in the unemployment rate in the context of the Great Recession and Sovereign Debt Crisis in Italy. The empirical results suggest that as unemployment conditions worsen the general fertility rate decreases, while the abortion rate and the propensity to abort conditional on pregnancy increase. Thus, both childbearing and abortion behave procyclically. This evidence also points to the fact that the pro-cyclical behavior of fertility rates is the result not only of changes in planned pregnancy but also a higher incidence of abortions. The estimated effect on childbearing stands at around 1.4% change in fertility rate following a 1 p.p. change in the unemployment rate and is consistent with previous literature, such as Schaller (2016). This is however relatively limited compared to the estimated effect of job displacement (Del Bono, Weber, and Winter-Ebmer 2012), suggesting that local conditions have a smaller impact on childbearing than individual employment status.

Further analysis indicates that the reaction is increasing over the age profile, where younger women do not adjust their behavior (their childbearing even increases, though marginally) but older women do. Dynamic models of fertility predict that transitory fluctuations in wages might affect the timing of births, inducing a postponement of fertility, but they will not impact expected total fertility in the presence of perfect capital markets and certainty (Happel, Hill, and Low 1984; Hotz, Klerman, and Willis 1997). The reaction of women in the 24-34 age group can be reasonably interpreted as a fertility postponement. However, the reaction of women aged between 34 and 49 is even stronger. These women are the most likely to already have children but are also closer to the end of their reproductive life. A strong reaction from this group might therefore signal changes in their demand for additional children or having children at all. Future research can delve into this question and explore whether fertility and abortion responses vary by parity, i.e. whether the changes affect the intensive or extensive margin or fertility, and whether the total realized fertility of women exposed to recessions in their last reproductive years is affected.

This study brings supportive evidence to the theory that fertility behaves procyclically, an important factor to be taken into account by policy makers. Public policies may play a key role in fertility decisions during economic downturns, for instance through child subsidies or the provision of abortion and family planning services. Learning about the impact of the business cycle on fertility and abortion rates has therefore important implications in terms of allocation of resources, both financial and human, in particular in the presence of budget cuts to the health sector. Additional policy implications include the design of policies to mitigate the adverse effects of labor market swings, such as free or subsidized contraception during unemployment spells. Similar policies were introduced in Puglia in 2008 and since then other regions have followed; the present study thus leaves scope for future research towards evaluating the impact of these policies.

References

Adao, R., M. Kolesár, and E. Morales. 2019. "Shift-share designs: Theory and inference." *The Quarterly Journal of Economics* 134 (4): 1949–2010.

Aksoy, C. G. 2016. "The effects of unemployment on fertility: evidence from England." *The BE Journal of Economic Analysis & Policy* 16 (2): 1123–1146.

Ananat, E. O., and D. M. Hungerman. 2012. "The power of the pill for the next generation: oral contraception's effects on fertility, abortion, and maternal and child characteristics." *The Review of Economics and Statistics* 94 (1): 37–51.

Anderberg, D., H. Rainer, J. Wadsworth, and T. Wilson. 2016. "Unemployment and domestic violence: Theory and evidence." *The Economic Journal* 126 (597): 1947–1979.

Autor, D. H., and M. G. Duggan. 2003. "The rise in the disability rolls and the decline in unemployment." *The Quarterly Journal of Economics* 118 (1): 157–206.

Autorino, T., F. Mattioli, and L. Mencarini. 2020. "The impact of gynecologists' conscientious objection on abortion access." *Social science research* 87:102403.

Avdic, D., C. Sonja, and D. A. Kamhöfer. 2021. "Economic downturns and mental health in Germany." *European Economic Review* 140:103915.

Barr, A., and S. Turner. 2015. "Out of work and into school: Labor market policies and college enrollment during the Great Recession." *Journal of Public Economics* 124:63–73.

Barrow, L., and J. Davis. 2012. "The upside of down: Postsecondary enrollment in the great recession." *Economic Perspectives* 36 (4).

Bartik, T. J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research.

Bartiloro, L., L. Carpinelli, P. F. Russo, and S. Pastorelli. 2012. "L'accesso al credito in tempo di crisi: le misure di sostegno a imprese e famiglie." *Questioni di economia e finanza,* no. 111.

Bastianelli, C., M. Farris, and G. Benagiano. 2005. "Reasons for requesting emergency contraception: a survey of 506 Italian women." *The European Journal of Contraception & Reproductive Health Care* 10 (3): 157–163.

Bastianelli, C., E. Rosato, M. Farris, and G. Benagiano. 2016. "Emergency contraception: A survey of 1773 women." *The European Journal of Contraception & Reproductive Health Care* 21 (6): 455–461.

Becker, G. S. 1991. *A Treatise on the Family.* Harvard University Press.

Bell, B., A. Bindler, and S. Machin. 2018. "Crime scars: recessions and the making of career criminals." *Review of Economics and Statistics* 100 (3): 392–404.

Bhalotra, S., D. Britto, P. Pinotti, and B. Sampaio. 2021. "Job displacement, unemployment benefits and domestic violence." *IZA Discussion Paper* No. 14543.

Bitler, M., and M. Zavodny. 2001. "The effect of abortion restrictions on the timing of abortions." *Journal of health economics* 20 (6): 1011–1032.

Black, N., A. Jackson, and D. W. Johnston. 2022. "Whose mental health declines during economic downturns?" *Health economics* 31 (1): 250–257.

Blank, R. M., C. C. George, and R. A. London. 1996. "State abortion rates the impact of policies, providers, politics, demographics, and economic environment." *Journal of health economics* 15 (5): 513–553.

Bloom, D. E., D. Canning, G. Fink, and J. E. Finlay. 2009. "Fertility, female labor force participation, and the demographic dividend." *Journal of Economic Growth* 14 (2): 79–101.

Bono, E. D., and G. Morando. 2021. "For some, luck matters more: the impact of the great recession on the early careers of graduates from different socio-economic backgrounds."

Borusyak, K., P. Hull, and X. Jaravel. 2022. "Quasi-experimental shift-share research designs." *The Review of Economic Studies* 89 (1): 181–213.

Bound, J., and H. J. Holzer. 2000. "Demand shifts, population adjustments, and labor market outcomes during the 1980s." *Journal of labor Economics* 18 (1): 20–54.

Brown, D., E. De Cao, et al. 2018. *The impact of unemployment on child maltreatment in the United States.* Technical report. Institute for Social and Economic Research.

Bruckner, T. A., L. H. Mortensen, and R. A. Catalano. 2016. "Spontaneous pregnancy loss in Denmark following economic downturns." *American journal of epidemiology* 183 (8): 701–708.

Buckles, K., D. Hungerman, and S. Lugauer. 2021. "Is fertility a leading economic indicator?" *The Economic Journal* 131 (634): 541–565.

Comolli, C. L. 2017. "The fertility response to the great recession in europe and the United States: Structural economic conditions and perceived economic uncertainty." *Demographic Research* 36 (1): 1549–1600.

Consiglio Nazionale dell'Economia e del Lavoro. 2009. "Rapporto sul mercato del lavoro 2008-2009."

———. 2011. "Rapporto sul mercato del lavoro 2010-2011."

———. 2014. "Rapporto sul mercato del lavoro 2013-2014."

Crowley, J. E., R. Jagannathan, and G. Falchettore. 2012. "The Effect of Child Support Enforcement on Abortion in the United States." *Social science quarterly* 93 (1): 152–172.

De Philippis, M., A. Locatelli, G. Papini, and R. Torrini. 2021. *La crescita dell'economia italiana e il divario Nord-Sud: trend storici e prospettive alla luce dei recenti scenari demografici, forthcoming.*

Dehejia, R., and A. Lleras-Muney. 2004. "Booms, busts, and babies' health." *The Quarterly journal of economics* 119 (3): 1091–1130.

Del Boca, D., and R. M. Sauer. 2009. "Life cycle employment and fertility across institutional environments." *European Economic Review* 53 (3): 274–292.

Del Bono, E., A. Weber, and R. Winter-Ebmer. 2012. "Clash of career and family: Fertility decisions after job displacement." *Journal of the European Economic Association* 10 (4): 659–683.

Drago, R., K. Sawyer, K. M. Shreffler, D. Warren, and M. Wooden. 2011. "Did Australia's baby bonus increase fertility intentions and births?" *Population research and policy review* 30 (3): 381–397.

Fahlén, S., and L. S. Oláh. 2018. "Economic uncertainty and first-birth intentions in Europe." *Demographic Research* 39 (October): 795–834.

Fischer, S., H. Royer, and C. White. 2018. "The impacts of reduced access to abortion and family planning services on abortions, births, and contraceptive purchases." *Journal of Public Economics* 167:43–68.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift. 2020. "Bartik instruments: What, when, why, and how." *American Economic Review* 110 (8): 2586–2624.

Goldstein, J., D. D. Karaman Örsal, M. Kreyenfeld, and A. Jasilioniene. 2013. "Fertility Reactions to the "Great Recession" in Europe." *Demographic Research* 29 (July): 85–104.

Goldstein, J. R., and T. Cassidy. 2014. "A cohort model of fertility postponement." *Demography* 51 (5): 1797–1819.

González, L., and S. K. Trommlerová. 2021. "Cash transfers and fertility: How the introduction and cancellation of a child benefit affected births and abortions." *Journal of Human Resources,* 0220–10725R2.

González-Val, R., and M. Marcén. 2017. "Divorce and the business cycle: a cross-country analysis." *Review of Economics of the Household* 15 (3): 879–904.

———. 2018. "Unemployment, marriage and divorce." *Applied Economics* 50 (13): 1495–1508.

Haas-Wilson, D. 1996. "The impact of state abortion restrictions on minors' demand for abortions." *Journal of Human Resources,* 140–158.

Happel, S. K., J. K. Hill, and S. A. Low. 1984. "An economic analysis of the timing of childbirth." *Population studies* 38 (2): 299–311.

Hofmann, B., and K. Hohmeyer. 2013a. "Perceived economic uncertainty and fertility: Evidence from a labor market reform." *Journal of Marriage and Family* 75 (2): 503–521.

Hofmann, B., and K. Hohmeyer. 2013b. “Perceived economic uncertainty and fertility: Evidence from a labor market reform.” *Journal of Marriage and Family* 75 (2): 503–521.

Hotz, V. J., J. A. Klerman, and R. J. Willis. 1997. “The economics of fertility in developed countries.” *Handbook of population and family economics* 1:275–347.

Hotz, V. J., and R. A. Miller. 1988. “An empirical analysis of life cycle fertility and female labor supply.” *Econometrica: Journal of the Econometric Society,* 91–118.

Imai, K., and I. S. Kim. 2021. “On the use of two-way fixed effects regression models for causal inference with panel data.” *Political Analysis* 29 (3): 405–415.

Istat. 2017. “Natalità e fecondità della popolazione residente.”

Jaeger, D. A., J. Ruist, and J. Stuhler. 2018. *Shift-share instruments and the impact of immigration.* Technical report. National Bureau of Economic Research.

Kalist, D. E. 2004. “Abortion and female labor force participation: Evidence prior to Roe v. Wade.” *Journal of Labor Research* 25 (3): 503–514.

Levine, P. B., A. B. Trainor, and D. J. Zimmerman. 1996. “The effect of Medicaid abortion funding restrictions on abortions, pregnancies and births.” *Journal of Health Economics* 15 (5): 555–578.

Lima, J. M., A. Reeves, F. Billari, M. McKee, and D. Stuckler. 2016. “Austerity and Abortion in the European Union.” *European Journal of Public Health* 26 (3): 518–519.

Lindo, J. M., C. K. Myers, A. Schlosser, and S. Cunningham. 2020. “How far is too far? New evidence on abortion clinic closures, access, and abortions.” *Journal of Human resources* 55 (4): 1137–1160.

Loghi, M., and R. Crialesi. 2017. “La salute riproduttiva della donna.” *Letture statistiche - Temi.*

Malak, N., M. M. Rahman, and T. A. Yip. 2019. “Baby bonus, anyone? Examining heterogeneous responses to a pro-natalist policy.” *Journal of Population Economics* 32 (4): 1205–1246.

Matysiak, A., T. Sobotka, and D. Vignoli. 2021. “The Great Recession and fertility in Europe: A sub-national analysis.” *European Journal of Population* 37 (1): 29–64.

Medoff, M. H. 1997. “A pooled time-series analysis of abortion demand.” *Population Research and Policy Review* 16 (6): 597–605.

———. 2007. “Price, restrictions and abortion demand.” *Journal of Family and Economic Issues* 28 (4): 583–599.

Ministero della Salute. 2016. “Relazione del Ministro della Salute sulla attuazione della legge contenente norme per la tutela sociale della maternità per l‘interruzione volontaria di gravidanza (Legge 194/78).”

———. 2018. “Relazione del Ministro della Salute al Parlamento sullo Stato di Attuazione della Legge contenente norme in materia di procreazione medicalmente assistita (legge 19 febgraio 2004, n. 40, articolo 15)–anno 2016, Roma, 28 giugno 2018.”

Modena, F., C. Rondinelli, and F. Sabatini. 2014. “Economic insecurity and fertility intentions: The case of Italy.” *Review of Income and Wealth* 60 (S1): 233–255.

Modena, F., and F. Sabatini. 2012. “I would if I could: Precarious employment and childbearing intentions in Italy.” *Review of Economics of the Household* 10 (1): 77–97.

Moretti, E. 2013. “Real wage inequality.” *American Economic Journal: Applied Economics* 5 (1): 65–103.

Myers, C., and D. Ladd. 2020. “Did parental involvement laws grow teeth? The effects of state restrictions on minors’ access to abortion.” *Journal of health economics* 71:102302.

Pallitto, C. C., C. Garcıa-Moreno, H. A. Jansen, L. Heise, M. Ellsberg, C. Watts, et al. 2013. “Intimate partner violence, abortion, and unintended pregnancy: results from the WHO Multi-country Study on Women’s Health and Domestic Violence.” *International Journal of Gynecology & Obstetrics* 120 (1): 3–9.

Ruhm, C. J. 2000. “Are recessions good for your health?” *The Quarterly journal of economics* 115 (2): 617–650.

Schaller, J. 2013. "For richer, if not for poorer? Marriage and divorce over the business cycle." *Journal of Population Economics* 26 (3): 1007–1033.

———. 2016. "Booms, busts, and fertility testing the Becker model using gender-specific labor demand." *Journal of Human Resources* 51 (1): 1–29.

Schneider, D. 2015. "The great recession, fertility, and uncertainty: Evidence from the United States." *Journal of Marriage and Family* 77 (5): 1144–1156.

Schwandt, H., and T. von Wachter. 2019. "Unlucky Cohorts : Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets." *Journal of Labor Economics* 37.

Snarr, H. W., and J. A. Edwards. 2009. "Does income support increase abortions?" *Social Choice and Welfare* 33 (4): 575–599.

Sobotka, T., V. Skirbekk, and D. Philipov. 2018. "Economic Recession and Fertility in the Developed World." *Population and Development Review* 37 (2): 267–306.

Spinelli, A., E. Forcella, S. Di Rollo, and M. Grandolfo. 2006. *L'interruzione volontaria di gravidanza tra le donne straniere in Italia.* Rapporto ISTISAN 06/17. Istituto Superiore di Sanità.

APPENDIX A

Appendix 1.A Additional results

Subsection 1.A.1 Age heterogeneity - response to the aggregate unemployment rate

Table A4 replicates the analysis relating age-specific dependent variables to the aggregate unemployment rate. Results are remarkably consistent with what was reported in Table 1.3, especially concerning the IV coefficients.

Age-specific unemployment rates reflect the market conditions that women in each age group are facing, particularly if the labor market is segmented. The response of different age groups to changes in the aggregate unemployment rate instead captures the response to overall market uncertainty. We might expect aggregate unemployment to play an important role in particular for young women, who are less likely to be in the labor force and rely on parental employment status.

Columns (1-3) report the results for age-specific fertility rates. Both the OLS and the IV estimates indicate that aggregate unemployment does not have a statistically significant effect on the fertility rate of women aged 15-24, while its effect on older women is negative and statistically significant at the 99.9% confidence level. For women in the 24-34 age group, a one standard deviation increase in the unemployment rate decreases the general fertility rate by 0.57 standard deviations and is statistically different from zero at the 99.9% confidence level. The effect is even larger for women between 35 and 49 years old, where the GFR decreases by 1.4 standard deviations per one standard deviation increase in unemployment.

Columns (4-6) consider the relation between age-specific abortion rates and the aggregate unemployment rates. The OLS analysis points to a positive effect of unemployment rates on the abortion rate for all age groups, though of limited economic significance. The IV analysis instead finds that unemployment rates do not have a statistically significant impact on the abortion choice of women in the youngest age group, while they have a positive and statistically significant effect on the abortion rates of older women. The IV estimates are larger in magnitude and indicate that a one standard deviation increase of the respective age-specific unemployment rate brings about a 0.24 standard deviation increase in the abortion rate of women in the 24-34 age group, which rises to a 0.39 standard deviation increase for women older than 34.

Table A4: Age specific fertility rates - standardized - total unemployment

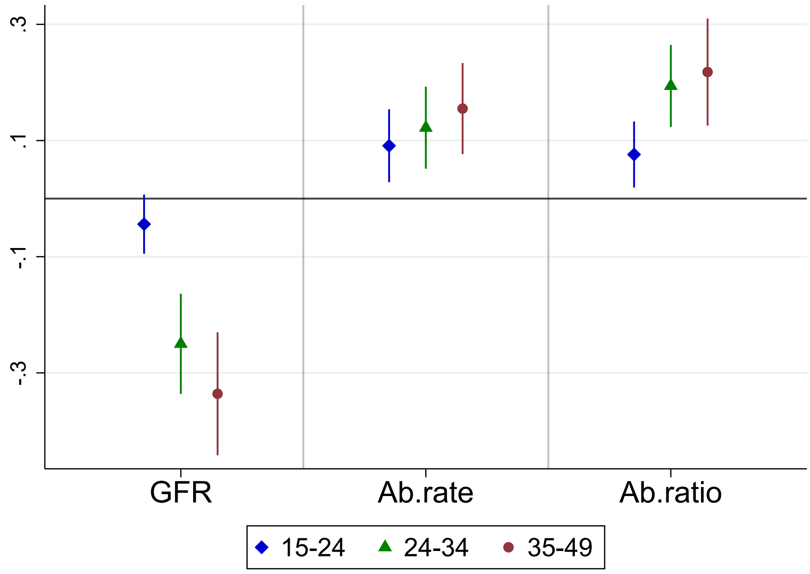
	GFR			Ab.rate			Ab.ratio		
	15-24	25-34	35-49	15-24	25-34	35-49	15-24	25-34	35-49
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
OLS									
$Unempl_{t-1}$	-0.044	-0.250***	-0.366***						
	(0.026)	(0.044)	(0.054)						
$Unempl_t$				0.091**	0.122**	0.155***	0.076*	0.194***	0.218***
				(0.032)	(0.036)	(0.040)	(0.029)	(0.036)	(0.047)
IV									
$Unempl_{t-1}$	0.024	-0.569***	-1.413***						
	(0.051)	(0.077)	(0.106)						
$Unempl_t$				-0.039	0.204*	0.415***	-0.1816*	0.299***	0.696***
				(0.084)	(0.082)	(0.090)	(0.084)	(0.080)	(0.097)
Observations	1236	1236	1236	1339	1339	1339	1339	1339	1339
R^2	0.369	0.279	0.094	0.480	0.283	0.472	0.125	0.189	0.481
KP LM p value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
KP F-stat	360.531	360.531	360.531	285.542	285.542	285.542	285.542	285.542	285.542

Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. The local unemployment rates are instrumented for using a leave-one-out Bartik instrument based on the number of employed individuals in each sector, using 2003 weights. The R^2 refers to the OLS estimation.

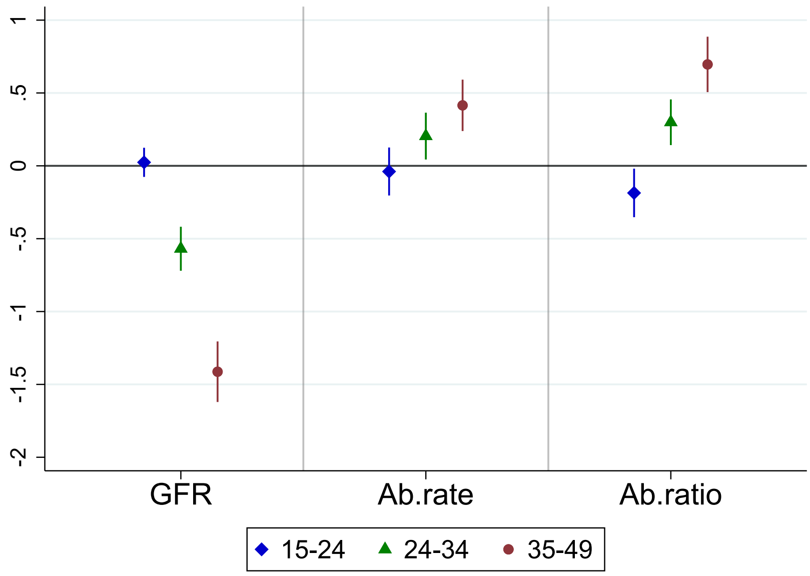
* $p < .05$, ** $p < .01$, *** $p < .001$

Figure A5: Regression estimates - age heterogeneity - aggr. unemployment

(a) OLS estimates



(b) IV estimates



Notes: 95% CI reported.

Finally, columns (7-9) analyze the response of age-specific abortion to pregnancies ratios. Again, the OLS estimation indicates that there is a positive and statistically significant correlation between the age-specific unemployment rate and their propensity to abort conditional on being pregnant for all age groups. The results of the IV estimation uncover an interesting relation: higher unemployment slightly decreases the propensity to abort of younger women, while it increases that of older women. These estimates are statistically significantly different from zero, respectively at the 95% and 99.9% confidence levels. In particular, a one standard deviation increase in the unemployment rate decreases the abortion to pregnancies ratio of women in the 15-24 age group by a 0.18 standard deviation, and increases that of women in the 25-34 age group by 0.33 standard deviations. For women older than 34, their propensity to abort significantly increases by 0.67 standard deviations following a 1 standard deviation increase in the unemployment rate. The increase in the abortions to pregnancies ratio of older women follows naturally from the decrease in fertility rates and increase in abortion rates attested by the previous columns.

Subsection 1.A.2 Geographic heterogeneity

This section considers a different margin of heterogeneity in the response to change in unemployment rates, distinguishing by geographical area. Because of the differences in labor market characteristics, female empowerment (including labor force participation), and access to abortions, we can expect the reaction of fertility and abortion to show substantial variation across geographic areas.

Figure A6 presents the standardized effects disaggregating by geographical area, and Table A5 reports the corresponding estimates. Dependent and independent variables are standardized at the national level, to preserve the comparability of the estimates.

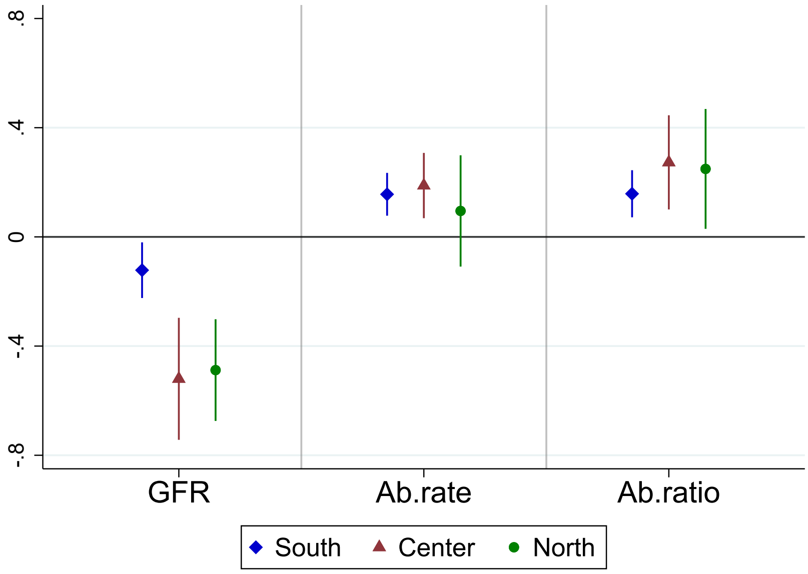
The direction of the estimated coefficients is homogeneous across geographical areas, while there is a noticeable pattern in the magnitude of the effects. Both fertility and abortion outcomes are less responsive to changes in the unemployment rate in Southern provinces, and the effect on the abortion rate and ratio is not significantly different from zero in the IV estimation. According to the IV estimation results, the general fertility rate responds similarly to a 1 standard deviation increase in unemployment in the Central and Northern provinces, while the reaction of the abortion rate is driven by Central Italy.

I explore two potential mechanisms behind this heterogeneity, which in particular might explain the low response from Southern areas. First, in more religious or abortion averse provinces the social cost of abortion might be higher, resulting in lower responsiveness to economic factors. Second, higher informality in the labor market might translate into less salience of unemployment rates for agents' choices. In Table A6, I proxy for these two channels using respectively the vote share against Law 194/78 in the 1981 referendum[29] and the share of irregular workers. The coefficients for the response to the unemployment rate are stable across specifications and close to the baseline estimate, particularly in the IV estimation., indicating that Columns (1), (4), (7) suggest that views against abortion do not have a statistically significant effect on fertility or abortion choice; though the OLS coefficient for the fertility rate is marginally significant and positive, this is not confirmed by the IV estimation. The share of irregular workers instead has a statistically significant effect on all dependent variables in the IV estimation. The coefficients on the interaction term between unemployment and the share of irregular workers suggest that a context of labor informality makes unemployment rates less relevant for fertility choice, as expected. In fact, the estimates for the interaction consistently go in the opposite direction of the main coefficient of interest, therefore reducing its magnitude in absolute terms. Labor informality might therefore a contributing factor to the zero effect found in Southern provinces, though the limited size of the estimates indicates that this channel alone does not explain this result entirely.

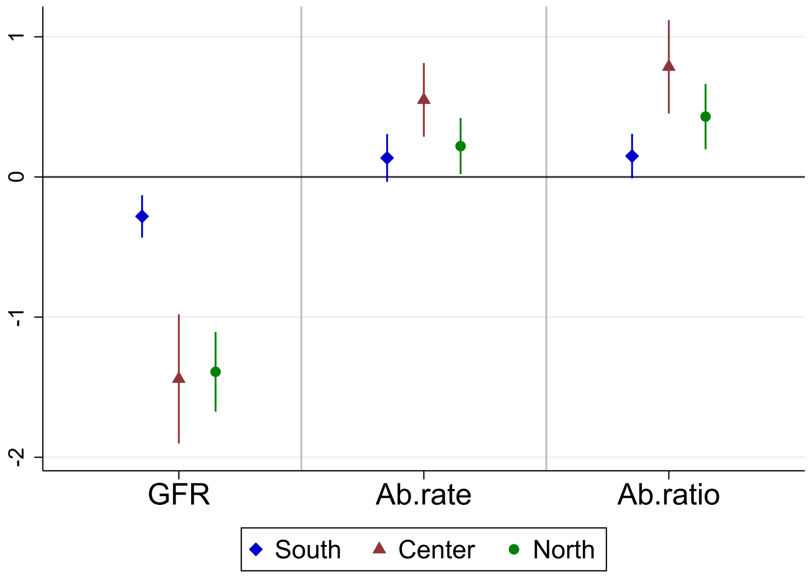
29. More specifically, I consider the share of favorable votes to the fifth referendum question, opposing the legalization of induced abortion.

Figure A6: Regression estimates - geographical heterogeneity

(a) OLS estimates



(b) IV estimates



Notes: 95% CI reported.

Table A5: Geographic heterogeneity - standardized

	GFR			Ab.rate			Ab.ratio		
	South	Center	North	South	Center	North	South	Center	North
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
OLS									
$Unempl_{t-1}$	-0.122*	-0.520***	-0.488***						
	(0.052)	(0.111)	(0.093)						
$Unempl_t$				0.156***	0.188**	0.095	0.158***	0.273**	0.249 *
				(0.040)	(0.061)	(0.104)	(0.044)	(0.088)	(0.112)
IV									
$Unempl_{t-1}$	-0.281***	-1.411***	-1.390***						
	(0.077)	(0.235)	(0.145)						
$Unempl_t$				0.136	0.559***	0.220	0.149	0.786***	0.431 **
				(0.087)	(0.134)	(0.102)	(0.081)	(0.170)	(0.119)
Observations	432	252	552	468	273	598	468	273	598
R^2	0.539	0.523	0.673	0.477	0.742	0.574	0.284	0.530	0.350
KP LM p value									
KP F-stat	192.9	83.88	274.6	145.8	67.35	259.0	145.8	67.35	259.0

Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. The local unemployment rates are instrumented for using a leave-one-out Bartik instrument based on the number of employed individuals in each sector, using 2003 weights. The R^2 refers to the OLS estimation.

* $p < .05$, ** $p < .01$, *** $p < .001$

Table A6: Heterogeneity - attitudes towards abortion and informal labor markets

		GFR			Ab. rate			Ab. ratio	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					OLS				
Unempl	-0.215***	-0.354***	-0.343***	0.143***	0.150**	0.150**	0.189***	0.210***	0.189***
	(0.048)	(0.050)	(0.050)	(0.032)	(0.050)	(0.048)	(0.036)	(0.050)	(0.048)
Ref._{1981}*Unempl	0.128*		0.085	0.017		0.030	-0.066		-0.057
	(0.053)		(0.050)	(0.043)		(0.044)	(0.044)		(0.046)
%Irreg			0.154			0.127			0.115
			(0.98)			(0.080)			(0.081)
%Irreg *Unempl		0.124***	0.108**		-0.005	-0.009		-0.020	-0.004
		(0.035)	(0.037)		(0.031)	(0.032)		(0.0323)	(0.035)
					IV				
Unempl	-0.957***	-0.917***	-0.916***	0.250***	0.225***	0.216***	0.380***	0.335***	0.342***
	(0.089)	(0.079)	(0.078)	(0.069)	(0.063)	(0.062)	(0.073)	(0.067)	(0.066)
Ref._{1981}*Unempl	0.018		-0.005	0.039		0.054	-0.054		-0.038
	(0.049)		(0.046)	(0.039)		(0.038)	(0.045)		(0.046)
%Irreg		0.216***	0.214**		0.107	0.122*		0.108	0.097
		(0.065)	(0.067)		(0.063)	(0.062)		(0.062)	(0.062)
%Irreg *Unempl		0.267***	0.267***		-0.045	-0.054		-0.080**	-0.074*
		(0.029)	(0.030)		(0.029)	(0.029)		(0.031)	(0.031)
Observations	1236	1236	1236	1339	1339	1339	1339	1339	1339
R^2	0.589	0.597	0.602	0.562	0.562	0.564	0.356	0.353	0.358
KP LM p value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
KP F-stat	176.918	254.518	169.645	139.050	197.915	132.467	139.050	197.915	132.467

Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. With GFR as dependent variable, all time varying controls are lagged by one year. The local unemployment rates are instrumented for using a leave-one-out Bartik instrument based on the number of employed individuals in each sector, using 2003 weights. The R^2 refers to the OLS estimation.

* $p < .05$, ** $p < .01$, *** $p < .001$

Subsection 1.A.3 Robustness checks

This section reviews various robustness checks, first reporting the estimates of alternative specifications considered in the main text, and second discussing potential measurement error in the abortion dependent variables.

Sensitivity to time specifications

The linear time trend removes the monotonic relationship between the outcome variables and time, thus leaving potential non-linear behavior in the residual term. Figure A1 reports the estimates from alternative time specifications such as location-specific time trends or a quadratic time polynomial, showing that estimates are not sensitive to these alternative time specifications.

Confounding policies and others

Table A1 reports the estimates for several robustness checks for each dependent variable; these are the same presented graphically in Figure 1.3. Results are fundamentally unchanged, as coefficients remain close to the benchmark in Table 1.2.

Columns (1-2) include time dummies for the introduction of potentially confounding national policies: the availability of medication abortion from 2009, a baby bonus policy from 2013, the labor market reform, baby bonus policy, and availability of emergency contraception without prescription from 2015. These policies are described in detail in the next subsection. Columns (3-4) add regional linear time trends to the baseline specification of Equation 1.2. Since health services are administered at the regional level, different regions might have been following different trends. Columns (5-6) present the results when excluding from the sample the provinces of Pesaro, Rimini, and the region of Puglia. Finally, columns (7-8) control for the share of objecting gynecologists for the abortion dependent variables, accounting for potential changes in the preferences of doctors over time.

Confounding national policies

Several important policy changes were implemented between 2004 and 2016, which potentially had an impact on fertility behavior. Birth allowances and childcare vouchers have been implemented in different forms and magnitudes over the years, as they are approved and fine-tuned yearly in the national budget plan. In 2009, a "Fund for loans to families with newborns" was launched to incentivize access to credit for

Table A1: Robustness

(a) GFR

	Policy dummies		Regional time trends		Restricted sample	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Unempl_{t-1}	0.021	-1.146***	-0.254***	-0.868***	-0.190***	-0.970***
	(0.057)	(0.248)	(0.042)	(0.065)	(0.053)	(0.096)
Observations	1236	1236	1236	1236	1164	1164
R^2	0.630		0.647		0.570	
KP F		108.564		499.754		328.6
KP LM pval		0.000		0.000		0.000

(b) Abortion rate

	Policy dummies		Regional time trends		Restricted sample		% objectors	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Unempl_{t}	0.093**	0.307*	0.119***	0.262***	0.127***	0.253**	0.120***	0.260***
	(0.035)	(0.134)	(0.031)	(0.051)	(0.034)	(0.082)	(0.033)	(0.077)
Obs	1339	1339	1339	1339	1261	1261	1301	1301
R^2	0.561		0.616		0.546		0.536	
KP F		101.869		481.319		254.420		255.389
KP LM pval		0.000		0.000		0.000		0.000

Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. * $p < .05$, ** $p < .01$, *** $p < .001$
The restricted sample in columns (1-2) excludes the provinces of Pesaro-Urbino, Rimini, and the region of Puglia from the sample.
The policy dummies used in columns (3-4) control for the introduction of medication abortion from 2009, a baby bonus policy from 2013, a labor market reform in 2015, and emergency contraception without prescription from 2015.
The local unemployment rates are instrumented for using a leave-one-out Bartik instrument based on the number of employed individuals in each sector, using 2003 weights.

Robustness (cont.)

(a) Abortion ratio

	Policy dummies		Regional time trends		Restricted sample		% objectors	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Unempl_t	0.093*	0.349*	0.162***	0.409***	0.168***	0.394***	0.162***	0.397***
	(0.040)	(0.142)	(0.034)	(0.057)	(0.041)	(0.086)	(0.038)	(0.081)
Obs	1339	1339	1339	1339	1261	1261	1301	1301
R^2	0.368		0.401		0.328		0.328	
KP F		101.869		481.319		254.520		255.389
KP LM pval		0.000		0.000		0.000		0.000

Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. * $p < .05$, ** $p < .01$, *** $p < .001$
The restricted sample in columns (1-2) excludes the provinces of Pesaro-Urbino, Rimini, and the region of Puglia from the sample.
The policy dummies used in columns (3-4) control for the introduction of medication abortion from 2009, a baby bonus policy from 2013, a labor market reform in 2015, and emergency contraception without prescription from 2015.
The local unemployment rates are instrumented for using a leave-one-out Bartik instrument based on the number of employed individuals in each sector, using 2003 weights.

households with newborns in 2009-2014, by offering loans at a subsidized rate with a State guarantee amounting to 50% of the amount in case of insolvency [30]. Since 2012, the government introduced a monthly voucher for babysitting and childcare services targeted at full-time working mothers with babies born between 2013 and 2016[31]. Moreover, in 2014 a means-tested benefit for medium-low income family households with newborns was approved, consisting of a minimum 80E/month allowance paid up to the age of three, with larger amounts for lower-income households[32]. Although demographic experts have not reached a consensus as to whether such short-term measures are effective to redress the population imbalance, they might have provided temporary support for mothers in uncertain economic times (Drago et al. 2011; Malak, Rahman, and Yip 2019).

In addition, the 2014 Jobs Act (Law 78/2014) implemented a wide-ranging reform of the Italian labor market, reducing firing costs and eliminating restrictions on the use of temporary contracts with the aim of reducing unemployment and labor market dualism. The changes induced by the reform might have affected both the unemployment rate and job security and consequently fertility and abortion through unobserved factors.

Finally, some regulatory changes interested the abortion procedure directly. Medication abortion was introduced in 2009, but its use remains rather limited and there is no evidence that it has induced an increase in abortion incidence. The share of medication abortions increased over time, from 5% in 2010 to 16% in 2016, particularly in the North and Center regions (Ministero della Salute 2016). Moreover, from 2015 emergency contraception can be purchased in pharmacies without a medical prescription, except in the case of minors[33].

To control for such policies, in Table A1 I replicate the analysis including an indicator variable for years from 2009 onwards to account for the availability of medical abortion; a time dummy from 2013 onwards to account for the kindergarten voucher; and a time

30. By end 2011, the participation rate of banks stood at 82% and 70% of the allocated funds had been employed; however, only 1% of households eligible in 2009 participated (Bartiloro et al. 2012). The remaining funds were used to extend the initiative, which was originally planned only for births between 2009 and 2011.

31. The value of the kindergarten voucher was initially set to 300 €/month and increased to 600 €/month from 2015. The voucher can be used as an alternative to optional parental leave for a maximum of 6 months, reduced to 3 months for self-employed mothers.

32. A similar lump-sum baby bonus of 1,000 Euros was in place between 2004 and 2006 but limited to second and higher parity births in 2004 and 2006. Consequently, this measure received only a limited number of requests. For its limitations, this measure is not considered in the present study; nevertheless, note that in 2004 it is captured by the constant, in 2005 it should have no effect on fertility choice because it was approved retroactively.

33. This regulatory change initially applied to the Ellaone, Stromalidan, and Escapelle emergency contraception drugs; the same provision was extended to Norlevo in 2016.

dummy for years from 2015 onwards to capture the effect of the baby bonus policy, the labor market reform and emergency contraception access.

Measurement error in the abortion data

Although the filing of Modello D12 is compulsory for all induced abortions performed in public or private hospitals, there are various sources of missing data in the number of abortions per province of residence. Table A3 presents estimates for the abortion outcomes taking into account different sources of measurement error.

First, this data will only reflect the number of legally performed abortions. According to the Ministry of Health, covert abortions accounted for 20% of total abortions by Italian women in 2016 (Ministero della Salute 2016), which means that the data used in this analysis captures the vast majority of the phenomenon. Moreover, covert abortions are partially measured by miscarriages, accounted for in Table A3.

Second, women can choose not to share their personal data, in which case the abortion is recorded in the data but all information regarding the province of residence and birth is omitted. This appears to be a minor concern since only 1% of observations do not report the province of residence.

Third, there are inconsistencies in the transmission of data from regions to the National Statistics Agency. Underreporting the number of abortions leads to a downward measurement error in the abortion rate and the abortion ratio. Comparing the number of reported abortions to official regional estimates, which integrate the incomplete data with hospital discharge data, I find that incomplete reporting affects 17% of region-year cells and that incomplete observations have on average 5% of total abortions missing. Incomplete data affects almost all regions at least once in the observed period, but only Campania and Sicilia report incomplete data for most years in the sample[34]. Columns (1-4) of Table A3 report the estimation results when excluding region-year cells with incomplete data from the analysis. The estimates maintain their statistical significance and are larger in magnitude than in the full sample: a one standard deviation increase in the unemployment rate brings about a 0.38 standard deviation increase in the abortion ratio and a 0.54 standard deviation increase in the abortion ratio.

34. The region-years affected by incomplete data transmission, according to ISTAT, are: Abruzzo (2009, 2012), Basilicata (2009, 2014), Calabria (2008), Campania (2002-2003, 2005-2014), Friuli-Venezia Giulia (2005, 2006), Liguria (2013), Lombardia (2014), Marche (2014), Molise (2005), Puglia (2012,2013), Sardegna (2008,2009, 2013-2015), Sicilia (2004-2012, 2014, 2015), Umbria (2010-2012), Veneto (2015,2016).

Table A3: Measurement error in abortion data

	Incomplete region-years				Miscarriages			
					Ita		Tot	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ab.rate	0.130**	0.394***						
	(0.041)	(0.076)						
Ab.ratio			0.169***	0.563***	0.175***	0.380***	0.170***	0.405***
			(0.048)	(0.081)	(0.037)	(0.073)	(0.037)	(0.072)
Obs	1094	1094	1094	1094	1339	1339	1339	1339
$\bar{Y}$	5.54	5.54	132.84	132.84	116.80	116.80	114.45	114.45
R^2	0.639		0.439		0.355		0.381	
KP LM		0.000		0.000		0.000		0.000
KP F-stat		238.972		238.972		285.542		285.542

Columns (5-6) measure the estimated number of pregnancies as the sum of live births and induced abortions for Italian women by province of residence, and miscarriages of Italian women by province of abortion; columns (7-8) consider the number of miscarriages by province of residence, but independent of citizenship.
The local unemployment rates are instrumented for using a leave-one-out Bartik instrument based on the number of employed individuals in each sector, using 2003 weights.
Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. * $p < .05$, ** $p < .01$, *** $p < .001$

Finally, the benchmark abortions to pregnancies ratio measure suffers from measurement error because it proxies the number of pregnancies only with live births and abortions. The estimated number of pregnancies is downward biased since it does not account for stillbirths and miscarriages, leading to an upward biased estimate of the abortion ratio. This measurement error is non-random insofar as the number of miscarriages is related to economic instability, for instance through maternal stress; Bruckner, Mortensen, and Catalano (2016) show that unexpected increases in the unemployment rate correlate with a rise in the number of spontaneous abortions one month later. Therefore, failing to account for spontaneous abortions might lead to an upward biased estimate of the effect of unemployment rates on abortion ratios. Columns (5-8) of Table A3 report the estimation results when including the number of spontaneous abortions in the calculation of the abortion ratio. Data on miscarriages is extracted from the yearly reports from the ISTAT data archive[35]. Columns (5-6) measure the estimated number of pregnancies as the sum of live births and induced abortions for Italian women by the province of residence, and miscarriages of Italian women by the province of abortion; columns (7-8) consider the number of miscarriages by the province of residence, but

35. Indagine sulle dimesse dagli istituti di cura per aborto spontaneo, ISTAT.

independent of citizenship. Since both corrections remain subject to some measurement error, the preferred proxy for the number of pregnancies remains the sum of live births and induced abortions. Coefficients remain statistically significant at the 0.1% confidence level and suggest that the effect of a 1 standard deviation increase in the unemployment rate ranges between 0.38 and 0.41 abortion ratio standard deviations.

Appendix 1.B Bartik instrument

Subsection 1.B.1 Industry sectors

Table B1 reports the list of sectors included in the analysis; in total, these are 10 sectors from the ATECO 2007 classification, which corresponds to the NACE Rev.2. Some industries are only available as a group, for example the manufacturing and extraction industries.

Figure B1 presents the initial geographical variation of the local employment shares for different industries, i.e. the shares used to compute the Bartik instrument. There is substantial variation across provinces in the importance of each sector for the local labor market, with Northern provinces concentrating on industrial production and services, and Southern areas concentrating on public administration and agriculture. Figure B2 shows the year-to-year growth of industry employment across the main sectors. The shaded area represents the years of crisis. The contraction during the crisis appears to encompass the majority of industries, especially in 2013, and manufacturing and construction stand out as the most affected sectors.

Table B1: Sectors used for the Bartik instrument

NACE code, Rev. 2	Sector
A	Agriculture, forestry and fishing
B-E	Mining and quarrying; manufacturing; electricity, gas, steam and air-conditioning supply; water supply, sewerage, waste management and remediation
F	Construction
G-I	Wholesale and retail trade, repair of motor vehicles and motorcycles; Transportation and storage; Accommodation and food service activities
J	Publishing, audiovisual and broadcasting activities; Telecommunications; IT and other information services
K	Financial and insurance activities
L	Real estate activities
M-N	Professional, scientific and technical activities; Administrative and support service activities
O-Q	Public administration and defence, compulsory social security; Education; Human health services; Residential care and social work activities
R-U	Arts, entertainment and recreation; Other services

Notes: Industries are aggregated to match the industry data available at the province level.

Figure B1: Industry shares in 2003

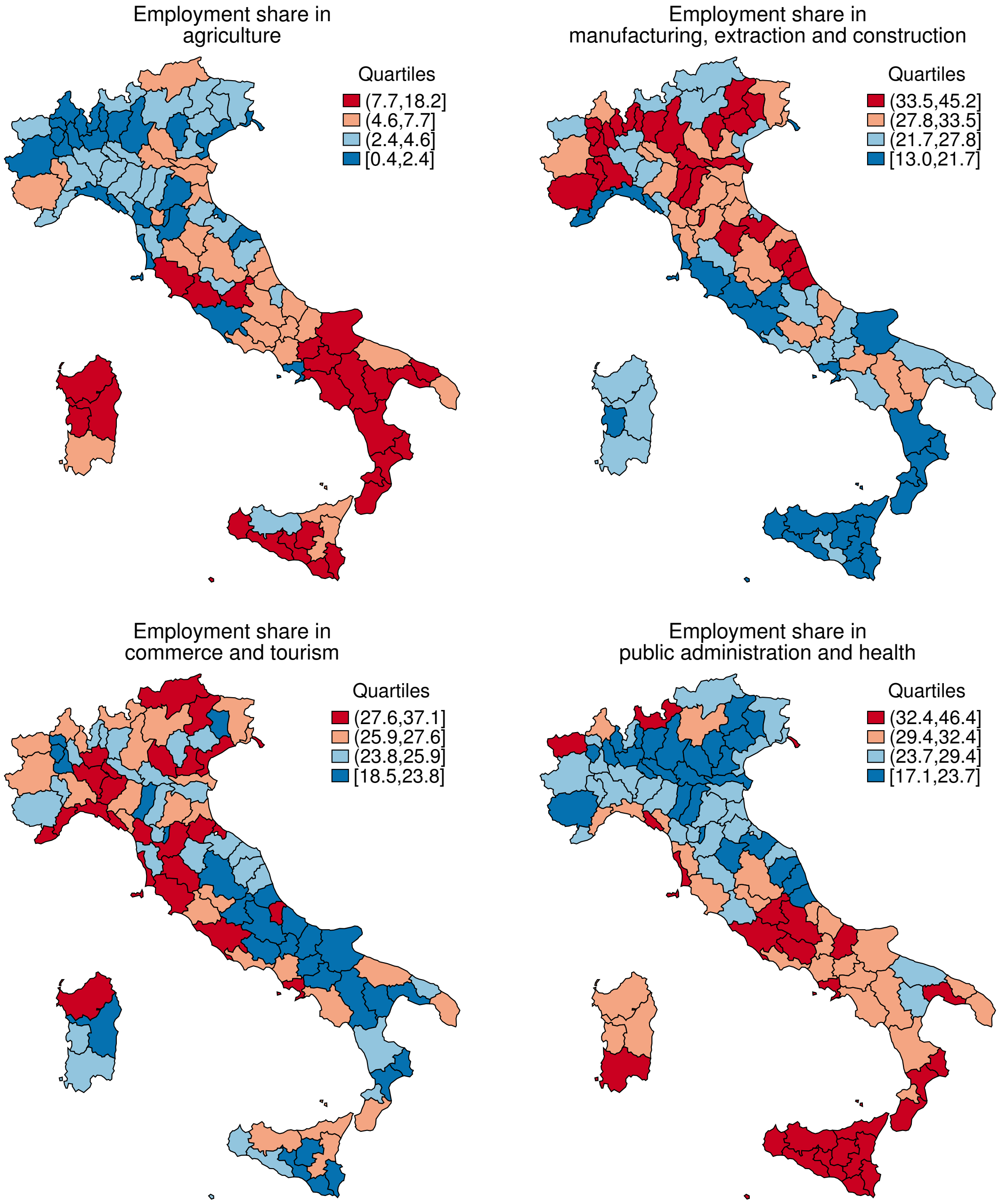
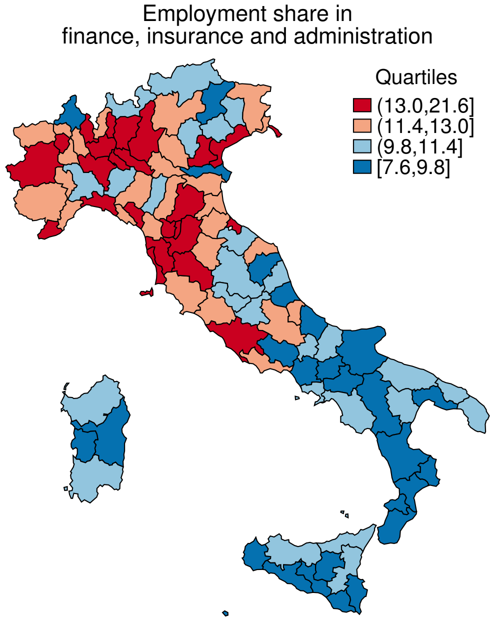
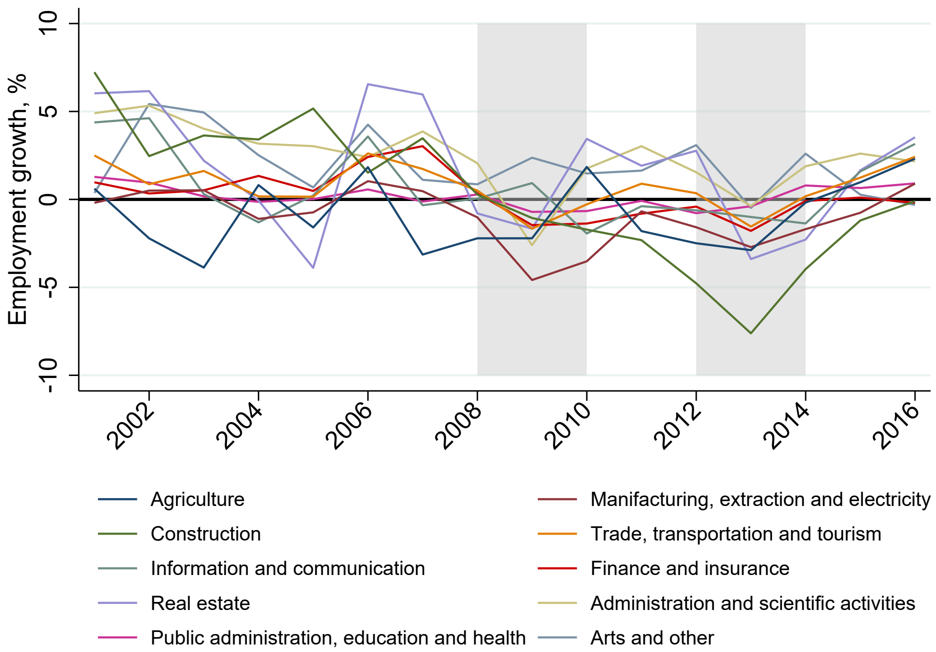


Figure B2: Employment growth

Industry shares in 2003 (cont.)





Subsection 1.B.2 First stage relationship

The first stage estimates the relationship between the province unemployment rate and the constructed shift-share instrument, which measures the predicted province employment level. If the labor force were constant, the number of employed individuals would perfectly determine the unemployment rate, so this index of employment and the unemployment rate would be almost inversely proportional. In the presence of changes in labor force participation, the level of employment will be informative of the unemployment rate insofar as changes in the labor force are not fully absorbed by only one of these two variables.

Figure B3 presents the first stage relationship graphically. Intuitively, the unemployment rates are negatively correlated with the employment level predicted by the Bartik instrument.

Table B2 reports the first stage estimates of the IV analysis, corresponding to the following regressions:

$$Unempl_{p,t-1} = \zeta_0 + \eta_0 Bartik_{p,t-1} + \theta_{0,p} + \chi_0 year + \mu_{0,p,t} \tag{1.6}$$

$$Unempl_{p,t} = \zeta_1 + \eta_1 Bartik_{p,t} + \theta_{1,p} + \chi_1 year + \mu_{1,p,t} \tag{1.7}$$

where the coefficients of interest are η_0, η_1. The large Kleibergen-Paap F-statistic informs us that the instrument is relevant, and the Kleibergen-Paap LM test rejects the null hypothesis of under-identification.

Figure B3: First stage relationship

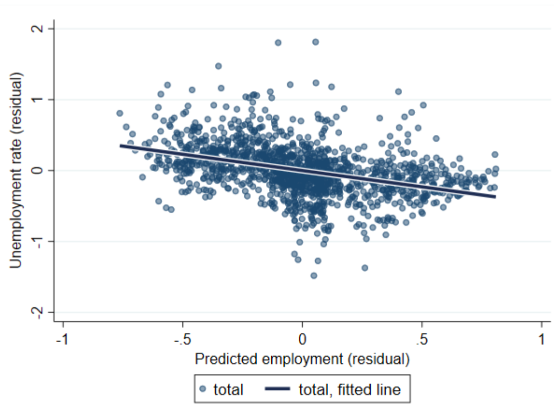


Table B2: First stage estimates - standardized

(a) Main specification

	GFR	Ab.rate & Ab.ratio
	(1)	(2)
	$Unempl_{t-1}$	$Unempl_t$
$Bartik_{t-1}$	-0.511***	
	(0.027)	
$Bartik_t$		-0.460***
		(0.027)
Observations	1236	1339
KP LM p-value	0.000	0.000
KP F-stat	360	285

(b) Age specific rates

	GFR			Ab.rate & Ab.ratio		
	15-24	25-34	35-49	15-24	25-34	35-49
	(1)	(2)	(3)	(4)	(5)	(6)
$Bartik_{t-1}$	-0.889***	-0.659***	-1.008***			
	(0.090)	(0.072)	(0.069)			
$Bartik_t$				-0.740***	-0.479***	-0.871***
				(0.069)	(0.065)	(0.066)
Observations	1236	1236	1236	1339	1339	1339
KP LM p-value	0.000	0.000	0.000	0.000	0.000	0.000
KP F-stat	97	82	213	114	54	176

Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. The local unemployment rates are instrumented for using a leave-one-out Bartik instrument based on the number of employed individuals in each sector, using 2003 weights.
* $p < .05$, ** $p < .01$, *** $p < .001$

Subsection 1.B.3 Alternative Bartik instruments

This section presents alternative versions of the shift-share instrument and discusses their performance in the main specification 1.2). For details on the construction of each measure, see the Data appendix below.

Figure B4 presents the evolution over time of different versions of the Bartik instrument, taking the (population-weighted) average instrument and changing the initial shares, the contemporaneous shifts, or adjusting the shifts to be gender or age-specific. The benchmark measure used in the main analysis considers the number of employed in each industry, weighted by the industry employment share in 2003. Panel B4a illustrates what happens when using different shifts, specifically comparing the index based on employment level versus value-added. Notably, the instrument based on value-added rebounds quickly after the crisis, while employment remains depressed. A quick comparison with Figure 1.1b suggests that the value-added-based index is less reflective of the extent of the crisis in the labor market. Moreover, value-added data is only available at the national level, so the resulting instrument cannot feature a leave-one-out correction. Panel B4b compares alternative instruments by changing the initial provincial shares, both in terms of content and base year, which results in a simple rescaling of the instrument. I consider as benchmark share the province-industry employment share, as in Schaller (2016), and alternatively the ratio of local industry employment to the working-age population, as in Brown, De Cao, et al. (2018)[36]. Panel B4c presents the gender-adjusted instruments, that reflect the different evolution over time of female and male employment by accounting for national demographic composition within industry (following Schaller, 2016). Male employment appears to be more affected by the double-dip recession, which is consistent with manufacturing and construction being the hardest hit industries (see Figure B2). Finally, Panel B4d illustrates the age-adjusted instruments, which notably show a different evolution over time across age groups. While the predicted level of employment is increasing over time for the oldest age group (35-65) while it is decreasing for younger workers. Since the age-specific instrument is constructed as the benchmark instrument rescaled by the time-varying employment share of each age group, these differences reflect the evolution in the composition of the national workforce over time, where the share of older workers

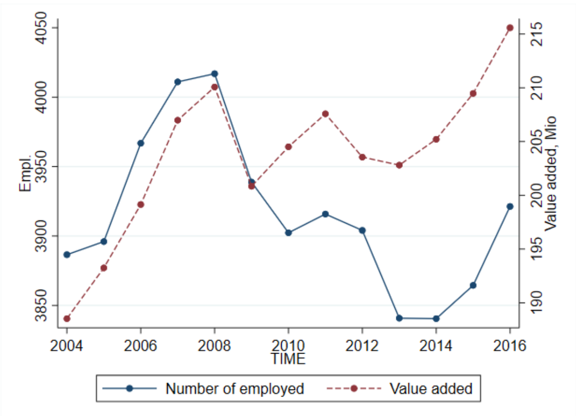
36. Unfortunately, data on the working-age population is not available at the province level for the year 2000. Moreover, the ratio of local industry employment to the working-age population by definition does not sum up to 1 within a province. This motivates the choice of the employment share as the benchmark in the main analysis.

has kept increasing throughout the observation period. Namely, the share of employed individuals between 35 and 65 years old increased from 65% in 2004 to 77% in 2016, while the share of younger workers decreases correspondingly. However, note that for the purpose of this analysis only the deviations from the linear time trend of each instrument are relevant, which instead evolve quite similarly across groups.

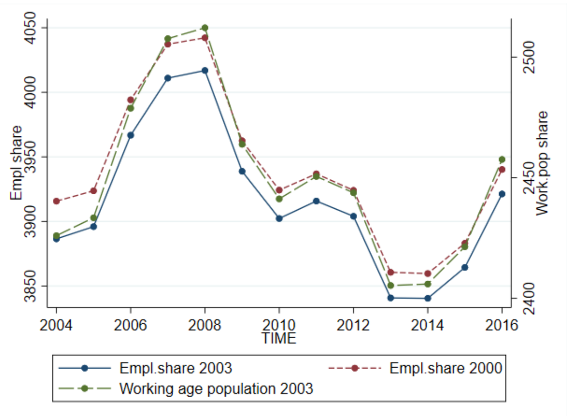
Finally, Table B3 reports on how estimates vary when employing different versions of the Bartik instrument.

Figure B4: Bartik instrument alternatives - average province

(a) Changing shifts

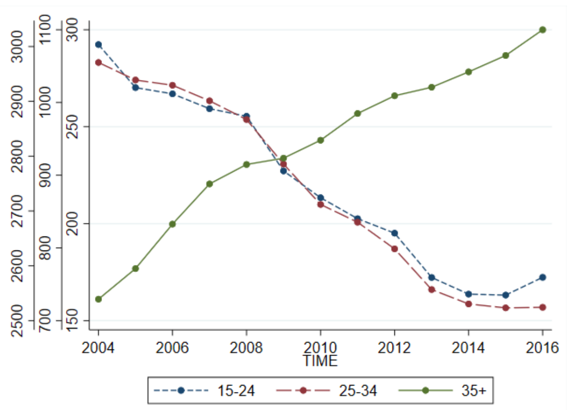


(b) Changing shares



(c) By gender

(d) By age



Notes: the figure plots the predicted value of different versions of the instrument for the median province (in terms of the distribution of the benchmark instrument), Cuneo. Panel B4a shows the predicted number of employed individuals and the predicted value added, while all other panels show the predicted number of employed individuals using different weights.

Table B3: Alternative Bartik estimators - standardized

(a) Second stage

	(1) E_$empl_{03}$	(2) E_$empl_{00}$	(3) E_wp_{03}	(4) VA_$empl_{03}$	(5) VA_$empl_{00}$	(6) VA_wp_{03}
	Dependent variable: GFR					
$Unempl_{t-1}$	-0.956***	-0.974***	-1.124***	-0.777***	-0.784***	-0.804***
	(0.085)	(0.091)	(0.102)	(0.061)	(0.061)	(0.060)
Observations	1236	1236	1236	1236	1236	1236
	Dependent variable: Abortion rate					
$Unempl_t$	0.267***	0.268***	0.305***	0.186***	0.181***	0.170***
	(0.070)	(0.075)	(0.077)	(0.050)	(0.050)	(0.048)
Observations	1339	1339	1339	1339	1339	1339
	Dependent variable: Abortion ratio					
$Unempl_t$	0.382***	0.372***	0.440***	0.324***	0.319***	0.328***
	(0.074)	(0.079)	(0.083)	(0.053)	(0.053)	(0.052)
Observations	1339	1339	1339	1339	1339	1339

(b) First stage

	(1) E_$empl_{03}$	(2) E_$empl_{00}$	(3) E_wp_{03}	(4) VA_$empl_{03}$	(5) VA_$empl_{00}$	(6) VA_wp_{03}
	Dependent variable: GFR					
$Bartik_{t-1}$	-0.511***	-0.524***	-1.642***	-0.575***	-0.606***	-1.472***
	(0.027)	(0.030)	(0.101)	(0.025)	(0.026)	(0.066)
Observations	1236	1236	1236	1236	1236	1236
KP LM pvalue	0.000	0.000	0.000	0.000	0.000	0.000
KP F-stat	360.531	310.931	264.509	537.680	525.489	504.050
	Dependent variable: Abortion rate and ratio					
$Unempl_t$	-0.460 ***	-0.465***	-1.500***	-0.567***	-0.595***	-1.467***
	(0.027)	(0.030)	(0.098)	(0.025)	(0.026)	(0.063)
Observations	1339	1339	1339	1339	1339	1339
KP LM pvalue	0.000	0.000	0.000	0.000	0.000	0.000
KP F-stat	285.542	242.509	233.017	526.748	513.950	533.836

Columns (1)-(3) refer to the leave-one-out Bartik instrument measured using the employment level, weighting industries by their employment share in 2003, 2000 and the working population in 2003, respectively. Column (1) corresponds to the preferred Bartik instrument used in the main text. Columns (4)-(6) refer to the leave-one-out Bartik instrument measured using industry real value added, weighting industries by their employment share in 2003, 2000 and the working population in 2003, respectively. Standard errors in parentheses. All regressions include province fixed effects and a linear time trend. * $p < .10$, ** $p < .05$, *** $p < .01$

Appendix 1.C Descriptive analysis of the recession

This section describes the recession in Italy, pointing out its different impacts on different age classes and geographical areas.

Subsection 1.C.1 The recession in Italy

Figure 1.1 in the main text illustrates the evolution of the unemployment rate over time, showing evident spikes between 2008 and 2014. The first increase in the unemployment rate corresponds to the Great Recession, which mostly affected employment in manufacturing firms (CNEL, 2009). The second increase relates to the Sovereign debt crisis, from the second half of 2011 to 2013, which caused a credit crunch and drop in internal demand, affecting mostly the construction and mechanic industry (CNEL, 2014).

The effect of recessions on the unemployment rate is partially mediated by changes in labor force participation. Figure C5 illustrates the evolution of the labor force over time, underlying the relationship between employed and unemployed. While labor force participation decreased slightly in 2009, it increased significantly in 2012 as inactive workers re-entered the labor market; as a consequence, the observed increase in the unemployment rate in 2009 is relatively small compared to the one in 2013. The Great Recession in 2009 thus induced change both along the extensive and intensive margin: the labor force participation rate decreased slightly, and the number of unemployed increased. In 2012 instead, participation in the labor market increased, but only to the benefit of unemployment, which further increased in 2013.

Subsection 1.C.2 North and South

Figures C6 and C7 give us a picture of the geographic differences in terms of unemployment, while Figure C8 portrays mean differences in the outcome variables across regions.

Figure C6 illustrates the evolution over time of unemployment rates in the average (population-weighted) provinces of Southern, Central, and Northern Italy. Two patterns emerge from this graph: first, the female unemployment rate is consistently higher than that of males regardless of geographic location; second, the average unemployment rate becomes smaller as we move towards the North. In particular, the labor market of Southern provinces is characterized by a higher unemployment rate, partially due to a higher prevalence of inactivity, irregular contracts, and low female labor force participation (De Philippis et al. 2021). Moreover, unemployment rates in the Southern provinces seem to respond less strongly to the Great Recession, while in the Center and North they increase in correspondence to both crises. This can be partially traced back to the different industrial composition of these macro-areas, as the export-oriented manufacturing industry that was greatly impacted by the Great Recession concentrates in Northern provinces, as illustrated by Figure B1. Moreover, the response of labor supply to the recession was also geographically heterogeneous: in Southern Italy, the increase in the unemployment rate during the first recession was partially offset by a large increase in inactive workers, while in the Central and Northern regions most of the reduction in employment translated directly into a rise of unemployment (CNEL, 2011).

Figure C7 maps the before-after change in unemployment rates across provinces. Although there is a slight prevalence of increase in unemployment rates in Northern provinces, a glance at the map does not reveal a clear North-South divide in terms of the labor market effects of the crises.

Finally, Figure C8 presents the distribution of the three dependent variables across regions. At a glance, there is no clear relation between fertility and abortion: high abortion rates are found in regions both in the top and bottom quartile of the fertility distribution. At the regional level, significant differences in the rates of fertility and abortion can be observed, but a clear pattern cannot be discerned except for the fact that abortion rates are (naturally) lower in areas where abortion services are harder to access, i.e. with higher prevalence of conscientious objection by medical staff.

Figure C5: Employment, unemployment and labor force participation

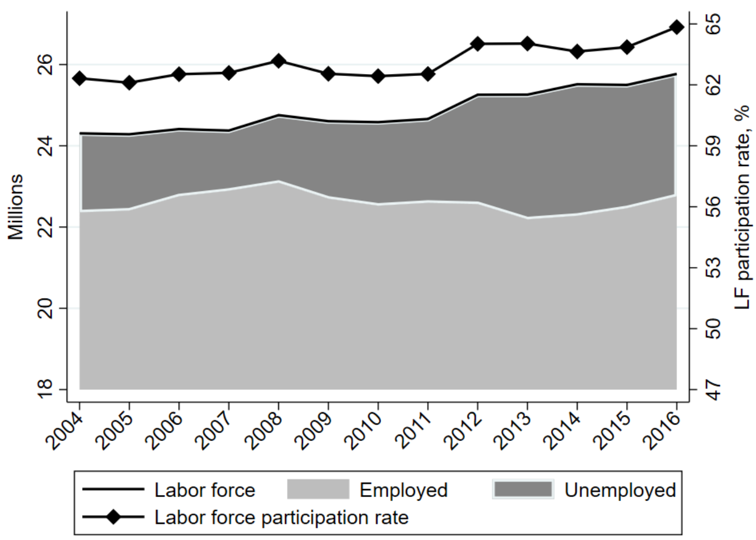
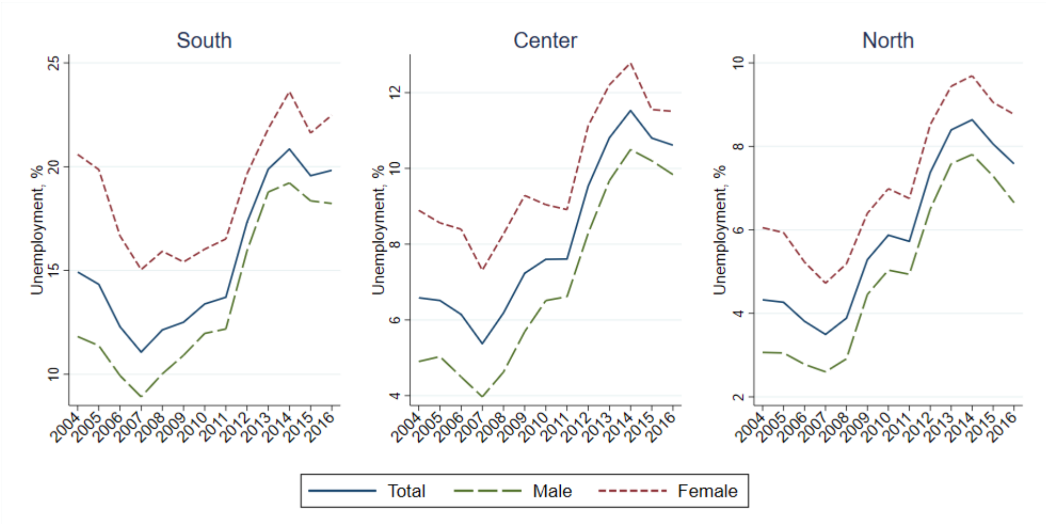
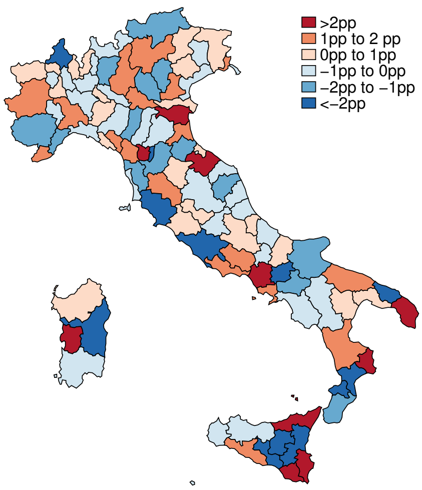


Figure C6: Unemployment by macro-areas and gender



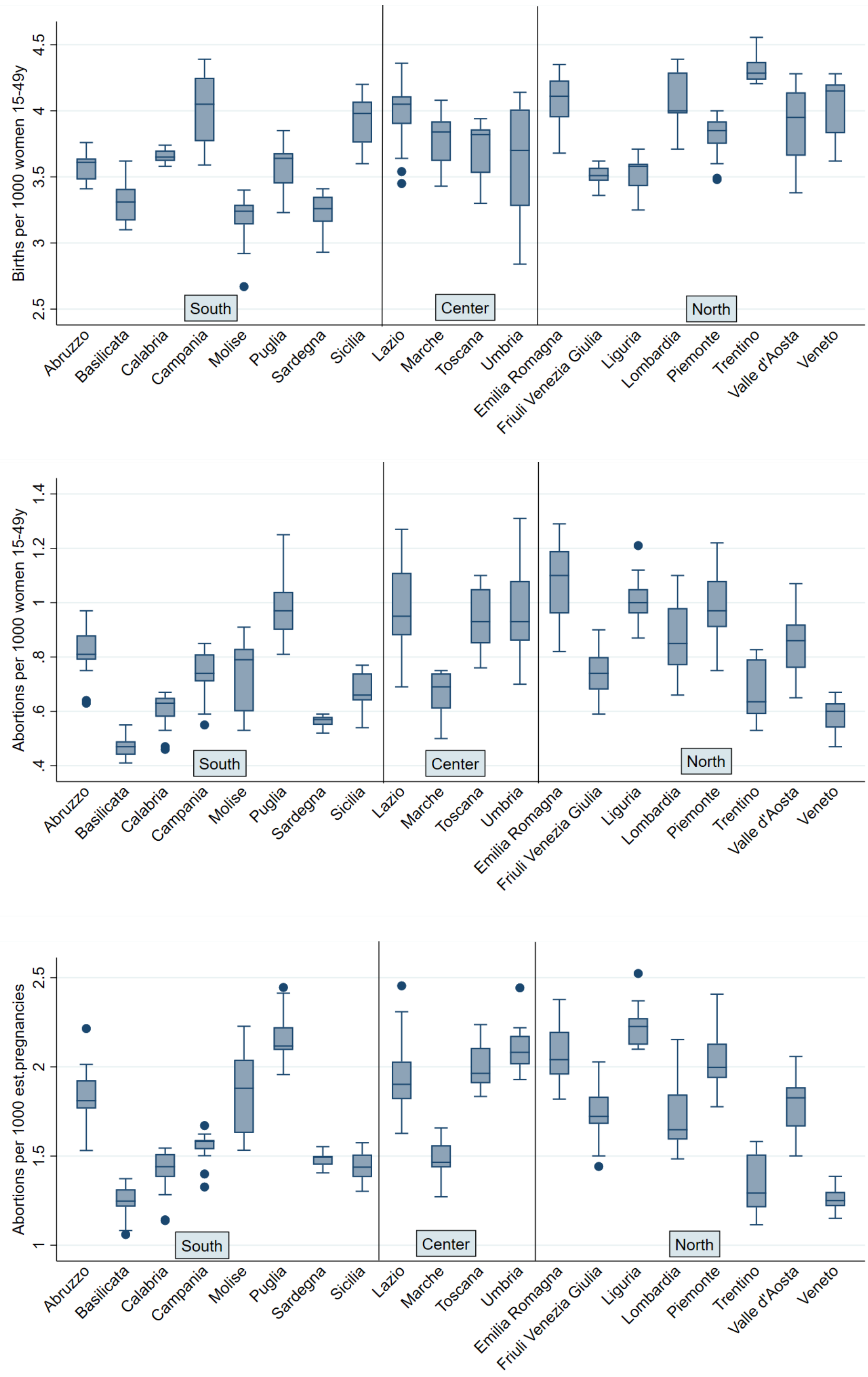
Notes: The figure plots the unemployment rate for the average (population weighted) province in the Center-North and South macroareas.

Figure C7: Changes in unemployment



Notes: This map presents the change in the unemployment rate before and after the recession. I regress the unemployment rate on region-specific time trends and calculate the average residuals for the pre-recession period 2004-2007 and for the post-recession period 2015-2016. I then subtract the pre-recession average from the post-recession average.

Figure C8: Distribution across regions



Notes: Abortion rates and ratio refer to the region of abortion, extracted from Ministero della Salute (2016).

Appendix 1.D Data appendix

Table D4: Data appendix

Variable	**Source**	**Method**
General fertility rate; population and live births	ISTAT, General Register Office	The GFR measures the number of live births for 1000 women in their childbearing age (15-49), considering only women of Italian nationality.
Abortion rate	ISTAT , Rilevazione sulle Interruzioni Volontarie di Gravidanza (Laboratorio ADELE)	The abortion rate measures the number of voluntary abortions for 1000 women in their childbearing age (15-49), considering only women of Italian nationality.
Abortion ratio	ISTAT, Rilevazione sulle Interruzioni Volontarie di Gravidanza (Laboratorio ADELE)	The abortion ratio measures the share of voluntary abortions over the number of estimated pregnancies in childbearing age, i.e. the sum of voluntary abortions and live births, considering only women of Italian nationality.
Miscarriages	ISTAT, Indagine sulle dimesse dagli istituti di cura per aborto spontaneo	Number of spontaneous abortions by province of abortion, 2010-2016.
Unemployment rate	ISTAT, Rilevazione sulle Forze di Lavoro	Computed using employment and unemployment counts from the LFS
Share of irregular workers	ISTAT, European System of Accounts	Share of employed individuals with irregular work per region, 2004-2016.
Referendum 1981	Ministry of Internal Affairs, Historical Archive of Elections	Share of yes votes to question 5 of the 17/05/1981 Referendum.

Table D5: Bartik instrument, data and formulas

Data and source	**Method**	**Formula**
Employment Bartik Employment by industry and province of work - ISTAT, Regional accounts	I take a weighted average of the national-level number of employed individuals added in each of the 10 sectors considered (see Table B1), where the weights are the local employment shares of each industry in 2003. The employment data measures the number of workers by place of work, i.e. it measures the workers that participate in the production process in each province: it does not account for residents that work outside of the province, while it includes non-residents that work in a firm located in the province. An alternative measure uses the number of employed by industry weighted by the share of working age population employed in each industry and province in 2003.	$B_{p,t} = \sum_{k=1}^{K} \chi_{p,k,t_0} E_{-p,k,t}$ $\chi_{p,k,t_0} = \frac{E_{p,k,t_0}}{E_{p,t_0}}$ Alternative weight: $\chi_{wp,p,k,t_0} = \frac{E_{p,k,t_0}}{Pop_{15-65,p,t_0}}$
Value Added Bartik VA by industry, Italy - ISTAT	I take a weighted average of the national-level real value added in each of the 10 sectors considered (see Table B1), where the weights are the local employment shares of each industry in 2003. Real value added by industry is measured in thousands of 2004 Euros.	$B_{AV,p,t} = \sum_{k=1}^{K} \chi_{p,k,t_0} VA_{IT,k,t}$

37. When the shift measure is national employment in an industry, adjusting for industry gender concentration translates into using directly the gender-specific national employment in an industry as the shift.

Gender-specific Bartik	For the gender-specific Bartik instrument I weight each industry by its gender concentration, i.e. the share of female (male) employees in the industry at the national level[37].	$$B_{g,p,t} = \sum_{k=1}^{K} \chi_{p,k,t_0} \xi_{g,IT,k,t} E_{-p,k,t}$$ $$\xi_{g,IT,k,t} = \frac{E_{g,IT,k,t}}{E_{IT,k,t}}$$
Age-specific Bartik	For the age-specific Bartik instrument, I rescale the standard Bartik instrument by the share of employed of each age class at the national level over the total number of employed.	$$B_{a,p,t} = \psi_{a,IT,t} \sum_{k=1}^{K} \chi_{p,k,t_0} E_{-p,k,t}$$ $$\psi_{a,IT,t} = \frac{E_{a,IT,t}}{E_{IT,t}}$$

Chapter 2

Do food price shocks affect farmers' nutrition? A study on rising quinoa prices in Peru

2.1 Introduction

> "The resulting quandary - local farmers earn more, but fewer Bolivians reap quinoa's nutritional rewards - has nutritionists and public officials grasping for solutions."
>
> — S. Romero and S. Shahriari [1]

The past decade has seen Western consumers devote increasing attention to their dietary choices, following a sequence of fashionable diets and nutrition trends. The recent superfood, gluten-free and plant-based nutrition trends have brought to the culinary forefront ancient grains, often of exotic origin, marketed as re-discovered healthy foods (Dreibus (2014); Loyer (2016); Harvard T.H. Chan School of Public Health, 2019). In the age of globalization, these sudden changes in Western consumers' behavior can have far-reaching consequences in the producing countries, starting from their effect on agricultural prices: Mišečka et al. (2019) show that attention driven behavior has short- and long-run impacts on the prices of agricultural commodities. A perfect collision of these nutrition trends, quinoa is a prime example of the potential side effects in the global South of changing dietary habits in the global North. Lessons learned from

1. Romero, S. & Shahriari, S. *Quinoa's Global Success Creates Quandary at Home.* The New York Times, 2011, March 13, pp. A6..

quinoa could shed a light on similar phenomena affecting food commodities such as açaí berries in Brazil, avocados and chia seeds in Mexico, or goji berries in China.

Quinoa consumption dates back to the Inca civilization, but was only recently re-discovered by the international community: in less than a decade, it went from being a largely unknown commodity to an extremely popular one among Western consumers. Quinoa came in high demand from mid-2008, especially in the United States, after Oprah Winfrey included the grain in her 21 days cleanse diet.[2] This sudden increase in demand could not be immediately satisfied by the limited supply, with Bolivia and Peru providing 92% of the global quinoa supply in 2008. Between 2008 and 2011 soaring international demand caused quinoa prices to almost triple, with the producer price in Peru increasing by 103% in 2009 alone (Figure 2.1). The price of quinoa further increased in 2013, declared by the United Nations Food and Agriculture Organization (FAO) the *International Year of Quinoa* with the aim of drawing attention to its potential role in providing food security, nutrition, and poverty eradication. The building up of international and national competition in production finally led to a drop in prices in 2015.

The boom of international demand for quinoa has left behind questions concerning its unintended consequences: the rise in quinoa prices might have increased the income of local farmers, but also made it less affordable and easily substituted with nutritionally poorer alternatives. This concern stems from quinoa production being highly geographically concentrated in the Andean Highlands regions of Bolivia and Peru, where quinoa is a traditional staple food for marginalized rural communities. The controversy was picked up by international news outlets, expressing concern that small farmers had effectively been damaged by the quinoa fad (Romero and Shahriari 2011; Friedman-Rudovsky 2012; Blythman 2013).

While the increase in quinoa prices opened a window of opportunity for its traditional farmers, leading to potentially large increases in income and improved food security, it could also have had unforeseen negative implications. Standard consumption theory would suggest that increasing demand for quinoa should increase the income of quinoa farmers through higher prices and increased sales volume, provided that the price change is transmitted to primary sellers and not completely absorbed by intermediaries. This would increase the welfare of local producers, and indirectly contribute to increased access to food through market purchases. However, the increased price of quinoa could

2. "The 21 day cleanse", Oprah's blog, June 2008. Available at http://www.oprah.com/food/The-21-Day-Cleanse-Oprahs-Blog-3

create a shift in local farm-household diets towards cheaper and less nutritious food. Ofstehage (2012) shows that the high prices of quinoa encouraged Bolivian farmers to sell the better quality quinoa and retain the lower quality one for local markets and self-consumption, and in some cases turn to cheaper and less nutritious alternatives. In addition, the high prices also reduced the accessibility of this grain for local consumers, potentially increasing food insecurity. This concern is partially attenuated by the fact that national consumption of quinoa increased in the medium term, sustained by government initiatives aimed at promoting internal demand.

Existing research on the welfare implications of increasing staple food prices has come to diverse conclusions. Some authors find a significant negative effect of higher staple food prices on food consumption (Alem and Söderbom 2012; Attanasio et al. 2013) and nutritional intake (Harttgen, Klasen, and Rischke 2016). Others find a null or even positive effect of a price increase on staple food consumption (Hasan 2016), household welfare and poverty (Deaton 1989; Mghenyi, Myers, and Jayne 2011). These studies have focused on price increases of staple foods in the context of a general rise in food prices such as the 2007-2008 food price crisis, whereas the case of quinoa stands out as a single commodity boom, where other crops were not affected, as shown by Figure 2.1.

Research focusing on the quinoa fad has yielded reassuring but unconvincing results. Stevens (2017) investigates the role of traditional tastes for quinoa in determining nutritional intakes, as households might have reduced consumption of other nutrient-dense foods to continue consuming quinoa. Using detailed household data from the Peruvian National Household Survey (Encuesta Nacional de Hogares, henceforth ENAHO) collected by the Instituto Nacional de Estadistica e Informatica (INEI), he compares nutrition patterns in 2004 and 2012. The author concludes that although regional tastes for quinoa do exist, they do not have a statistically significant impact on nutrition outcomes, while there is a significant income effect. However, this analysis does not distinguish between the consumption patterns of producers and consumers of quinoa. Bellemare, Fajardo-Gonzalez, and Gitter (2018) instead study the welfare effects of the quinoa price rise, using a pseudo-panel dataset from 2004 to 2014 obtained from aggregating the ENAHO data at different geographic levels (district, province, and region). Measuring household welfare by the total value of household consumption, they compare the consumption patterns of quinoa producers and consumers and find different effects depending on the level of aggregation. For smaller geographic units, higher concentrations of quinoa consumption or production are associated with a small and statistically significant increase in household welfare, while at the regional level higher concentrations of quinoa consumption or production are associated with small declines in welfare of less than 1% of total household consumption. In conclusion, they rule out that the rise in international demand for quinoa had large negative consequences on Peruvian household welfare. However, the authors do not discuss the mechanisms behind these results, and how the effect of the quinoa price increase on household expenditure might interplay with the effect on income. Finally, Gamboa et al. 2017 estimate a structural model of quinoa production and consumption using original farm-household survey data collected in 2015 in the region of Junín, a traditional quinoa-producing region in Peru. They conclude that the positive effect of the quinoa price increase on farm profits and income is such that the net effect on quinoa consumption is positive.

This paper aims at a better understanding of the causal effect of income changes on nutrition, by analyzing nutritional outcomes more closely and differentiating across geographical areas based on quinoa production. Using Peruvian ENAHO data from 2004 to 2012, I study the changes in Peruvian households' nutritional choices before and after the quinoa price shock of 2008. This analysis thus focuses on the first, unexpected, increase in the price of quinoa which can be reasonably conceived as an exogenous shock for local farmers. In particular, I restrict the sample to agricultural households residing in the Peruvian Highlands, the mountainous region of Peru where quinoa is traditionally cultivated, to exclude coastal areas where quinoa production spread as a response to the price increase. I first compare quinoa-farming households to other agricultural households using a difference in differences framework. I then move on to a comparison of households across districts historically engaged in quinoa farming or not, to address concerns of endogeneity of quinoa production and spillover across households within the same district. The failure to satisfy the parallel pre-trends assumption for some outcome variables puts some limitations on the causal interpretation of estimates. Results of both specifications point to a limited effect on nutritional outcomes in terms of caloric intake and overall diet quality, as measured by the Diet Quality Index developed by Kim et al. (2003).

The paper is organized as follows: Section 2 delineates the characteristics and production process of quinoa; Section 3 describes the dataset, and Section 4 outlines the empirical strategy and sample selection. Section 5 presents the results from the main specification, and Section 6 provides a variety of specification checks. Section 7 concludes.

2.2 Quinoa: history, characteristics, production

Quinoa, or *Chenopodium Quinoa*, is an ancient grain that played a substantial part in the Incan diet. With the arrival of the Spaniards, it was relegated from the table of the settlers and stigmatized for centuries as peasant food. Its consumption remained alive only in the traditionally quinoa-producing Andean highlands, feeding large, marginal populations alongside other native grains such as corn or amaranth (Mujica 2015). As such, in the early 2000s quinoa was considered a Neglected and Underutilized Species (NUS) (Padulosi et al. 2014). Its nutritional content, genetic variety, and climate adaptability have earned quinoa the title of both *super food* and *super crop*, bringing it to the attention of the international scientific community since the 1990s (McDonell 2015).

From a nutritional point of view, quinoa has a relatively higher content of proteins and minerals and lower fat compared to other cereals (Repo-Carrasco, Espinoza, and Jacobsen 2003). The protein content for 100 gr of quinoa ranges between 9 and 16 gr, while the same figure is around 10 for maize and 14 for wheat (Nowak, Du, and Charrondière 2016). Moreover, quinoa is a complete protein food, because it contains an adequate proportion of each of the nine essential amino acids necessary in the human diet (James 2009).[3] For its nutritional characteristics, quinoa is of particular interest for vegetarian and vegan consumers, but also for policy makers as a potential tool to eradicate malnutrition.

In addition, quinoa has an extraordinary biological variety and versatility: its different varieties are suitable for various climates, withstanding temperatures from -8°C to 38°C, and are highly resistant to drought and saline soils. Because of its high adaptability, marginal agricultural soils are frequently used to grow quinoa: these soils have poor or excessive drainage, low natural fertility, or very acidic to alkaline conditions. For its qualities as a crop, quinoa has been portrayed as a climate change resistant crop and researched in an effort to identify or develop varieties most suitable for adverse conditions.

The quinoa price shock and quinoa production

Historically, Bolivia and Peru were the main global producers of quinoa, accounting for 92% of world production in 2008. In 2014 Peru became the leading producer and

3. Other examples of single-source complete proteins are red meat, eggs, soybeans, buckwheat, and hempseed.

exporter, producing 62% of the global quinoa supply (Montero and Romero 2017). Exports of Peruvian quinoa began in 2005, the United States being the main destination market (Carimentrand et al. 2015).

Figure 2.1 plots the producer price for quinoa against the price of other crops cultivated in Peru, where the quinoa price stands out with noticeable spikes in 2009 and 2013.[4] The first price increase corresponds to the initial phase of quinoa becoming world-renowned for its nutritional properties. In June 2008, Oprah Winfrey inserted quinoa in her 21 days cleanse diet[5], contributing to quinoa's popularity among American consumers.

Prices of other crops considered in Figure 2.1 also started rising in 2007 and 2008, but by no means as much as quinoa. Nonetheless, the progressive increase of quinoa prices in 2008 could have been interpreted as relatively normal in the context of the global commodity price crisis, where the international prices of cereals such as wheat, soybeans, maize, and rice surged upward. In 2009, the quinoa price increased by 103% compared to the previous year, and from then onwards its evolution is quite separate from that of other crops. This confirms that the boom was specific to quinoa, as the prices for other grains did not follow the same pattern. From 2009 onwards prices kept steadily increasing, and between 2008 and 2013 the price jumped from 3.78 to 7.56 PEN/kg.[6]

The second price spike, in 2013, corresponds to the *International year of quinoa* declared by the UN Agency FAO in a campaign aimed at further promoting quinoa. Following the demand boom, other countries have started producing in large quantities, although Peru and Bolivia remain the main producers. The building up of international competition and the rise of internal production finally led to a sudden drop in prices in 2015, which however remain far higher than initially.[7]

Based on Figure 2.1, I consider the year 2008 as the start of the quinoa boom; the inspection of monthly producer prices reported in Figure 2.6 in the Appendix further pins it down to May 2008.

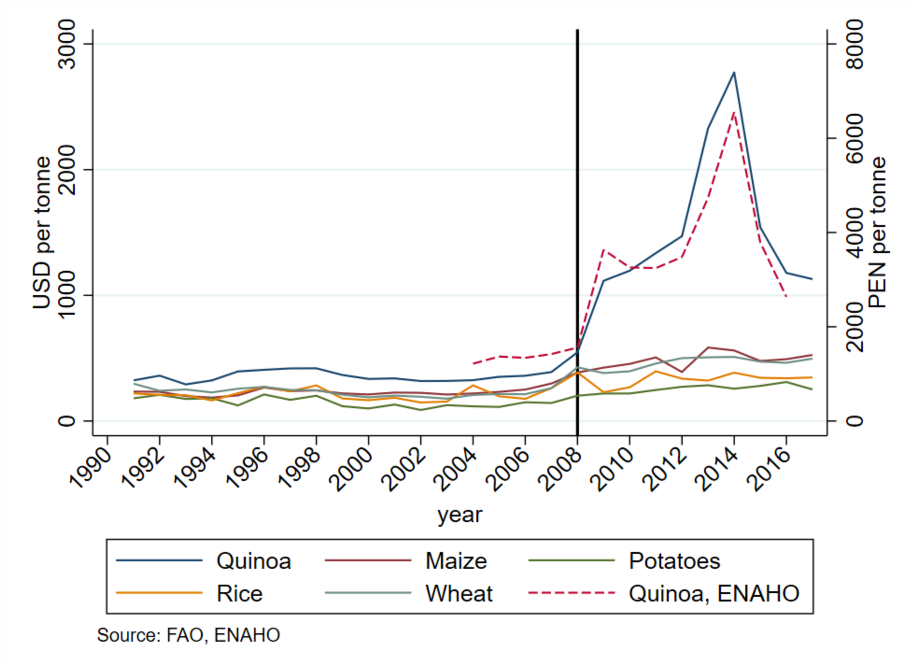
4. For the monthly evolution of local quinoa prices, see Figure 2.6 in the Appendix

5. "The 21 day cleanse", Oprah's blog, June 2008. Available at http://www.oprah.com/food/The-21-Day-Cleanse-Oprahs-Blog-3

6. Corresponding to a jump from 1.26 to 2.96 US dollars per kg, converted at historical exchange rates.

7. In 2017, the producer price for quinoa in Peru was 3.68 US dollars per kg (FAO 1997).

Figure 2.1: Producer prices for quinoa and staple crops, Peru



Quinoa production in Peru has traditionally concentrated in the mountainous area of the Sierra, mostly in the regions of Puno and Junín, which until 2011 represented 99% of national production (FAO 2015). Here, quinoa was a traditional staple and subsistence crop among Andean households, cultivated mostly by smallholders and often exchanged locally through bartering in local markets (Carimentrand et al. 2015; Gamboa et al. 2017).

For a niche, small-scale industry such as this, the sudden spotlight meant increasing prices and opportunities for upward income mobility, as supply could not immediately match demand. The sudden rise in international demand for quinoa might have thus a change of paradigm, transforming quinoa from a common staple and subsistence crop into a high-value cash crop (Repo-Carrasco, Espinoza, and Jacobsen 2003). In fact, the increase in selling price encouraged farmers to increase production both along the intensive and extensive margin. Bedoya-Perales et al. (2018) investigate expansion patterns of quinoa production in traditional quinoa regions in the Peruvian Andean region between 1997 and 2014. They find that overall quinoa acreage expanded by 7% between 2008 and 2012, compared to 51% between 2013 and 2014; crop yield (kg/ha) increased by 0.09% and 45% respectively.

Over time, the boom in prices altered both trading and production systems. Local government and international organizations supported the development of farmer associations, certification processes for organic quality, productivity investments such as irrigation systems, and higher-yield varieties (Carimentrand et al. 2015).

Market pressure also led to the expansion of quinoa production out of its traditional cultivation area, reaching the coastal regions. This process was not immediate, as it required adapting to very different climatic conditions, the Peruvian coast being a warm, almost desertic environment. Large-scale production on the coast substantially picked up only in 2014, when it accounted for 40% of national production (Romero 2015). Production of quinoa in non-traditional regions is characterized by being large-scale, technology-intensive, and allowing for harvesting twice a year, due to the warmer weather.

Finally, institutional efforts aimed at rural development and improving nutritional quality intensified after the quinoa boom. Both local governments and international organizations started several initiatives, e.g. providing agricultural training and information campaigns on the nutritional value of quinoa; such initiatives are discussed in more detail in Appendix 2.E.

To limit the confounding effect generated by the market entry of large-scale producers and governmental and international initiatives targeting quinoa farmers, the analysis considers only the years following the first quinoa price spike, until 2013.

2.3 Data and sample selection

The analysis draws on three sources of information. The main one is a repeated cross-section of households from the National Household Survey (Encuesta Nacional de Hogares, or ENAHO) collected by the Peruvian National Institute of Statistics and Information Technology (INEI), between 2004 and 2012. This information is then augmented with nutritional food-nutrient conversion tables published by the Peruvian Center for Food and Nutrition and the United States Department of Agriculture (USDA) National Nutrient Database.

The Peruvian National Household Survey has been carried out since 1995 to produce a wide range of indicators including data on the evolution of poverty, labor market indicators, income, and living conditions of Peruvian households. The survey covers approximately 22,000 households each year, selected through a three-step stratified cluster sampling procedure and roughly evenly split between urban and rural areas. Within each of the 25 administrative regions of Peru, population centers are divided into 8 strata according to population size. The selection of primary sampling units occurs within region-specific strata, where the sampling probability of each population center is proportional to the number of households in them. The secondary sampling units are census conglomerates, each comprising an average of 120 private dwellings. Conglomerates are drawn with a probability proportional to the number of resident households, and with implicit stratification based on socio-economic variables. Finally, households are randomly sampled from the selected conglomerates. In rural areas, 8 households per conglomerate are drawn, while in urban areas 6 households per conglomerate are selected.

The ENAHO is a comprehensive survey, spanning a wide range of topics. The head of the household provides information on household composition, expenses, and food consumption.[8] Information on schooling and employment is provided directly by each member of the household. The survey on agricultural production is directly administered to all individuals that have engaged in agricultural production in the previous week. It is mandatory for respondents to provide the information, and the national non-response rate was around 7% in 2013-2014.[9] Pertinent to this analysis, the ENAHO survey reports detailed information both on agricultural production and food expenditure at the household level. Food consumption is reported via a two-week recall, while agricultural production is based on a one-year recall. The INEI then produces annualized estimates of consumption and expenditure for each good, assuming constant consumption throughout the year.[10]

To obtain a measure of nutritional quality, I couple the data on household food expenditure from the ENAHO survey with nutritional data, using the food-nutrient conversion table published by the Peruvian Center for Food and Nutrition (Reyes García, Gómez-Sánchez Prieto, and Espinoza Barrientos 2017), as done by Stevens (2017). For foods that have no corresponding entry in the Peruvian nutrition data, primarily hard liquor and sweets, I use information from the USDA National Nutrient Database for Standard Reference (US Department of Agriculture 2016). In total, I match a list of 661 unique food items. Moreover, data on fat content is obtained from the Argentinean

8. If the head of the household is not present, another member of the household is interviewed. If the main respondent does not know some information, the interviewer is instructed to ask other members of the household. In the sample, the head of the household answers to the interview in the majority of cases (46%), which is a male in 67% of cases; if the head of the household is not present his spouse (26%) or oldest son (26%) answer.

9. Non-response rates show large variations between urban areas (8.8%) and rural areas (2.3%) and among socioeconomic categories, with higher non-response rate for higher-income households. Survey weights correct for non-response.

10. If a household did not purchase a specific commodity in the preceding two weeks period, the annualized extrapolation is reported as zero.

Dietary Guidelines (Ministerio de Salud de la Nación 2016), allowing to distinguish between saturated, monounsaturated, and polyunsaturated fatty acids. Finally, I use Recommended Dietary Allowances by age and gender (Trumbo et al. 2002; Appel et al. 2005) to compute the recommended nutritional intake for each household. This procedure allows obtaining household-level data on the consumption of calories, protein, fat, carbohydrates, and other nutrients. This approach to measuring food avalability closely follows Anríquez et al. (2008) and Ecker, Weinberger, and Qaim (2010). These estimates on the one hand overestimate food consumption, by not accounting for food waste, but also underestimate it by assuming that the household only consumes the items reported in the survey. Overall, the net effect of this measurement error is ambiguous.

To have a comprehensive measure of diet quality, I then compute the Diet Quality Index - International (henceforth DQI-I) developed by Kim et al. (2003). This index comprises a variety of aspects of diet quality, measuring food variety, adequacy of nutrient intake, moderation of unhealthy nutrients, and overall diet balance.[11] Overall the DQI-I ranges from 0 to 100, with higher values indicating a higher quality diet. Further details on the nutritional variables and the construction of the DQI-I are reported in Section 2.D of the Appendix.

In the main analysis, I consider daily per capita consumption of calories and other nutrients as an outcome variable, where each household member has equal weight.[12] Per capita caloric intake only provides a rough indicator of food availability and a household's nutritional status, since caloric needs vary by age and sex. However, this measure is positively related to diet adequacy, as measured by the DQI-I index, which accounts for household size, age, and sex composition to compute household-specific nutritional recommended intakes Figure 2.10 in the Appendix further illustrates the relation between daily per capita caloric intake and the four DQI-I subindices. Finally, in the estimation, I use the logarithm of caloric consumption, as common in the nutrition literature.

Peru is divided into 25 administrative regions and the autonomous Lima Province, which are further subdivided into 196 provinces and 1,869 districts[13]. In the present study, I restrict the sample geographically to areas in the Peruvian Highlands region,

11. Details on the construction of the index are reported in the section 2.D.2 of the Appendix.

12. Because caloric consumption is measured from food purchases at the household level, I do not observe intra-household food allocation.

13. In a comparative perspective, Peruvian districts would correspond to the administrative level of U.S. counties. A district is an area with at least 3,5000 residents at the time of its foundation (this threshold increases to 4,000 in the Andean highlands).

where quinoa traditionally occurs. In addition, I only consider districts that are sampled both before and after the price shock and that are engaged in agricultural production, i.e. that have a positive share of households declaring to be agricultural producers. By doing so, I exclude purely urban areas to obtain a more comparable control sample and exclude the coastal quinoa production, which developed mostly as a response to the quinoa boom. Moreover, I restrict the time window of the sample by using survey waves until 2012, in order to limit the potential issues related to the second price spike and the market entry of large-scale producers.

The final dataset comprises 736 districts and 35,699 households; on average, a district has 39 surveyed households per year. Figure 2.2 presents the geographical distribution of different macro-areas and the within-district concentration of quinoa farmers.[14] The Figure illustrates how quinoa-producing districts concentrate in the Peruvian Highlands, particularly in the Sierra.[15] Figure 2.3 reports the share of quinoa-farming households in Peruvian districts before and after the price increase, providing a sense of spatial variation in quinoa production. As expected, the geographical distribution of quinoa-producing areas does not change much between 2004 and 2012, since the expansion of quinoa production to non-traditional areas mostly happened from 2013 onwards. Moreover, districts with the highest concentration of quinoa farmers are located in the Puno and Junin regions.

14. The reported share of quinoa-farming households per district is measured as the within district time mean before 2008.

15. The comparison of households in the Sierra versus other macro-regions is further discussed in section 2.A.2 of the Appendix.

Figure 2.2: Macro areas and quinoa

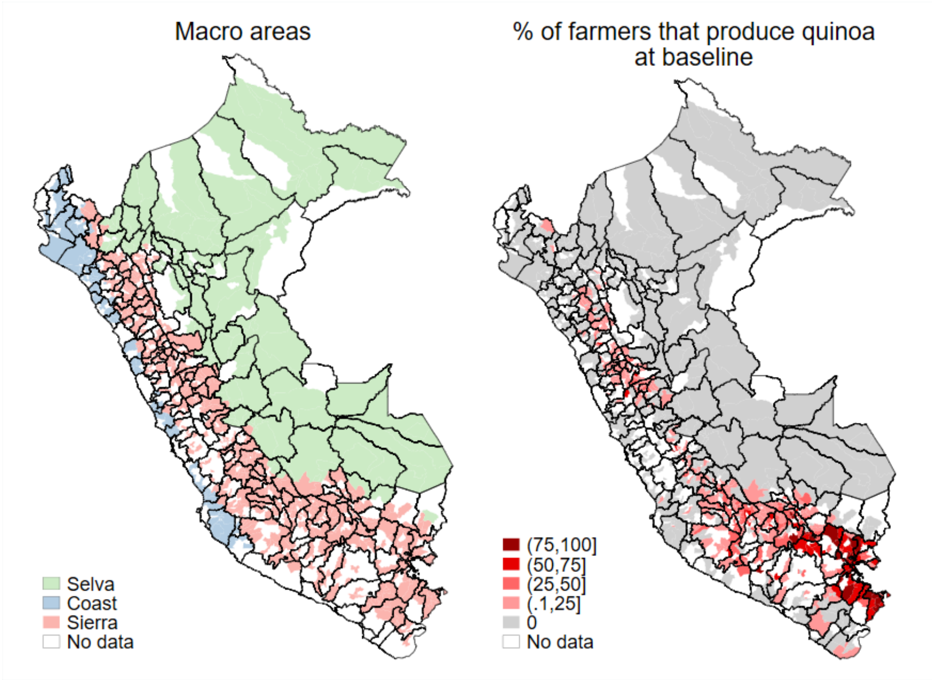
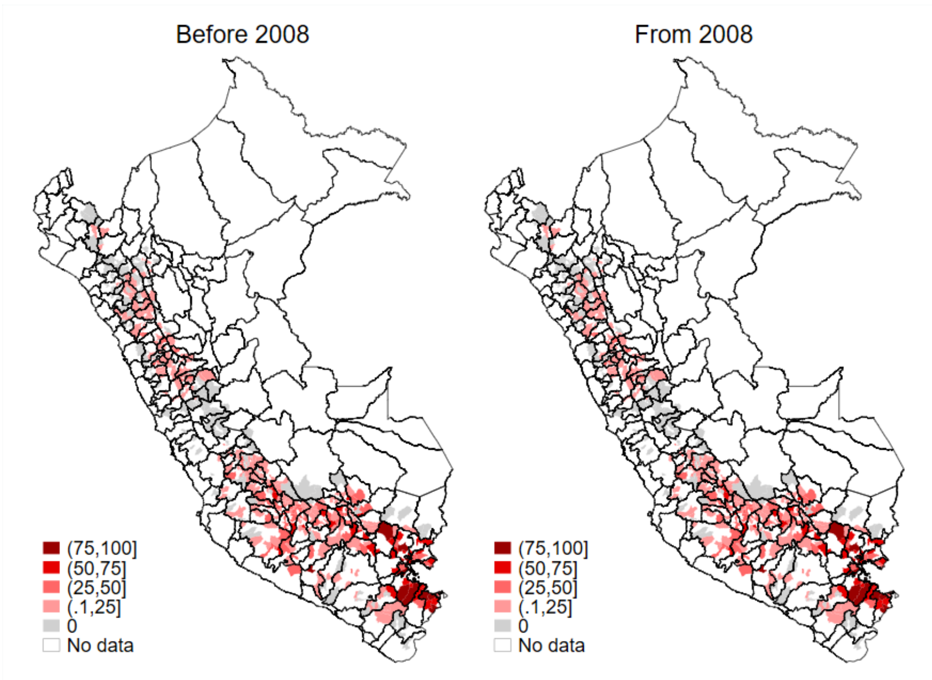


Figure 2.3: Share of quinoa-farming households by district



2.4 Empirical strategy and sample selection

I use a difference in differences (DID) framework to study the effect of the quinoa boom on various nutrition outcomes. I first estimate the impact of the price shock on quinoa farming households, then I move on to considering districts traditionally engaged in quinoa farming.

In the baseline specification, I compare quinoa farming households with households engaged in other agricultural production in the same province. Therefore, treatment assignment is naively defined as being in a quinoa farming household in a specific year. Formally, I estimate the following regression:

$$\begin{aligned} Y_{i,t} = & \alpha + \beta Quinoa_{i,t} + \gamma Quinoa_{i,t} * Post_{08} + \delta X_{i,t} \\ & + \zeta_t + \eta_m + \theta_p + \lambda_s + \epsilon_{i,t} \end{aligned} \tag{2.1}$$

where I observe each outcome variable Y for household i, residing in district d, surveyed in year t and month m. All regressions use sampling weights and cluster the standard errors by secondary sampling unit (a conglomerate of dwellings) to account for the sampling design and spatial correlation between households.

The coefficient of interest is the effect of the price shock on quinoa-farming households, γ. Variable $Post_{08}$ indicates observations after the first price increase, i.e. from May 2008 onwards.[16] I include a vector of household characteristics X_i, namely the household gender composition, and the gender, age, and educational attainment of the household head.[17] In addition, I control for year t, month of survey m, province p and population size s dummies. Location fixed effects capture common confounders within a province, such as agricultural specialization. Population size dummies classify districts by the number of residents and allow comparing areas with similar population sizes, thus further distinguishing between urban and rural areas to increase the precision of the estimates.[18] Survey month dummies account for potential seasonalities in food and agricultural production reporting: since yearly food consumption is extrapolated from a 2-week recall survey, the household nutrition outcome might be affected by the month of interview. For example, if quinoa-producing areas are surveyed in months of the year

16. Figure 2.6 in the Appendix displays the monthly evolution of local quinoa prices.
17. Table 2.1 in the Appendix documents the balance of characteristics across quinoa-farming and other farming households at baseline.
18. Results are robust to the exclusion of these population size dummies.

corresponding to the quinoa harvest, the survey might overestimate the consumption of quinoa.

Defining treatment assignment based on whether a household farms quinoa, however, can hinder the causal interpretation of the estimates. First, this treatment indicator is potentially endogenous, as the decision of being a quinoa producer can be influenced by the profitability of quinoa, hence its market price. However, the share of quinoa farmers within each district does not significantly change across 2008, suggesting that at least initially the price change did not attract new producers.[19] Second, the validity of the DID analysis relies on the absence of spillover effects from treated to control units. If the rise in international demand for quinoa induces local economic growth, it will indirectly affect also households that do not directly farm quinoa.

As an alternative estimation strategy, I redefine treatment at the aggregate level to avoid self-selection into treatment and spillover effects. I compute the share of quinoa-farming households in each district at the beginning of the sample[20], thus measuring the ex-ante traditional taste for quinoa farming and land suitability of each district. From this share, I construct a binary treatment indicator $Quinoa_{d,t_0}$, which indicates districts that produced quinoa at the beginning of the sample. Thus, in this second specification, I consider assigned to treatment households residing in districts that traditionally farmed quinoa, and compare them to households residing in districts that engage in other agricultural production. Formally, I estimate the following regression:

$$\begin{aligned} Y_i =& \alpha + \beta_1 Quinoa_{d,t_0} + \gamma_1 Quinoa_{d,t_0} * Post_{08} + \delta_1 X_i + \\ & + \zeta_{1,t} + \eta_{1,m} + \theta_{1,p} + \lambda_{1,s} + \epsilon_{1,i} \end{aligned} \tag{2.2}$$

where i represents the household, d the district, t the year and m the survey month. Observations are weighted by their sampling weights and I cluster the standard errors by secondary sampling unit. The coefficient of interest is now the effect of the shock on quinoa-producing districts, γ_1. Thus, I compare households across districts with different ex ante quinoa specialization within the same province, stratum, and month.

This alternative specification is preferred because it overcomes the issue of potential spillovers within a district, by comparing households across districts. Moreover, since

19. Stevens (2017) also finds that the proportion of quinoa producers does not change between 2004 and 2012.

20. Since by sampling design the panel of districts is unbalanced, I consider the average share of quinoa-farming households in the district before 2008.

treatment is now predetermined, selection into treatment is time-invariant and does not pose a threat to the identification of causal estimates. However, the use of aggregate treatment risks diluting the treatment effect because in treated districts only a portion of households effectively farms quinoa. In fact, only 27% of households in treated districts engaged in quinoa production.

2.4.1 Pre-treatment characteristics

Although the DID framework does not require covariate balance across treatment groups before treatment, inspecting such differences can provide useful insights. Table 2.1 reports the balance of characteristics across quinoa-farming and other farming households before treatment, i.e. before May 2008, and the respective two-sample mean comparison t-tests. The results highlight that there are no statistically significant differences in household composition, number of children, and age of the household head. However, quinoa-farming households have statistically higher average education and are less likely to be unemployed, but their household heads are younger and more educated. In terms of nutrition, quinoa farming households have consistently higher caloric intake, as their diet is composed of a larger share of proteins and a lower share of fats. This is consistent with the nutritional contribution of (auto consumed) quinoa to their diet, which provides additional proteins. This translated into them having a significantly higher estimated diet quality at baseline. In terms of income, quinoa farmers have on average lower net total income, which is the result of lower monetary income and lower value of their agricultural production. On average, a member of a quinoa-farming household earns 128 pesos less per year (around 33$) than a member of a non-quinoa-farming household. On the other hand, per capita expenditures are similar across the two groups, where quinoa farming households have higher food expenditure but also a larger share of the food budget covered by self-consumption. Finally, in quinoa-farming households farmers are more likely to be female; moreover, plot size and the number of cultivated crops are higher.

Table 2.2 compares characteristics at baseline across treated and control districts. Household members in quinoa-farming districts tend to have fewer male members, less education, and a lower unemployment rate. Similar to the comparison by quinoa-producing status, individuals in quinoa-farming districts have higher caloric intake, diet quality, and lower overall income but the same spending level as individuals in other districts.

Table 2.1: Balance at baseline - household treatment

	Control	Treated	Mean comparison test		
	$Quinoa_{i,t} = 0$	$Quinoa_{i,t} = 1$	Diff	se	t-stat
HH_size	4.40	4.39	0.0123	0.04	0.28
N.children	1.69	1.70	-0.00726	0.03	-0.23
Within the HH					
Avg age	34.53	34.08	0.451	0.35	1.27
Avg years of schooling	3.68	3.97	-0.286***	0.06	-4.93
Unemployed, % *	8.62	6.76	1.856***	0.31	5.89
Household head					
Male	0.83	0.82	0.00145	0.01	0.20
Age	51.40	50.79	0.608*	0.30	2.04
Years of schooling	4.42	4.88	-0.458***	0.08	-5.71
Daily kcal per capita	2071.38	2261.52	-190.1***	20.13	-9.45
Daily kcal from carbs, %	74.36	74.47	-0.104	0.14	-0.72
Daily kcal from protein, %	11.18	11.75	-0.573***	0.06	-10.39
Daily kcal from fat, %	16.60	15.80	0.799***	0.12	6.75
Diet Quality Index	56.03	57.62	-1.598***	0.15	-10.37
Net total income	1880.67	1752.79	127.9**	42.29	3.02
Net money income	1206.14	1038.95	167.2***	36.89	4.53
Donations received	138.44	130.40	8.048	6.17	1.30
Agricultural income	307.45	261.53	45.92**	14.73	3.12
Expenditure	1653.77	1620.04	33.73	24.47	1.38
of which food expenditure, %	52.27	54.62	-2.352***	0.32	-7.42
of which autoconsumption, %	30.96	35.97	-5.007***	0.40	-12.63
Female agricultural producers, %	20.71	24.40	-3.689***	0.78	-4.76
Land cultivated, ha	2.84	3.81	-0.975***	0.29	-3.38
of which rented, %	19.74	20.01	-0.264	0.72	-0.36
N.agr products	4.94	6.96	-2.021***	0.05	-42.84
District share agr producers, %	81.79	82.37	-0.580	0.43	-1.34
District share quinoa producers, %	8.73	54.35	-45.63***	0.35	-131.57
Observations	16,668	3,188			

All variables are in per capita measures and weighted by household survey weights. Monetary variables are in real terms, at 2004 constant PEN. The value of agricultural production is computed per producer.
* Individuals are considered unemployed if they had no employment in the past week, do not have a job lined up, and are not home workers. The unemployment rate measures share of working age (15-65) members of the household that are unemployed.

Table 2.2: Balance at baseline - district treatment

	Control	Treated	Mean comparison test		
	$Quinoa_{d,t0} = 0$	$Quinoa_{d,t0} = 1$	Diff	se	t-stat
HH_size	4.42	4.37	0.0504	0.03	1.50
N.children	1.68	1.69	-0.0171	0.02	-0.72
Within the HH					
Male, %	50.29	49.16	1.133***	0.34	3.38
Avg age	34.28	34.55	-0.272	0.27	-1.00
Avg years of schooling	3.91	3.69	0.223***	0.05	4.92
Unemployed, %	9.50	8.65	0.848**	0.26	3.28
Household head					
Male	0.83	0.82	0.0185***	0.01	3.34
Age	51.46	51.13	0.338	0.23	1.47
Years of schooling	4.61	4.50	0.110	0.06	1.76
Diet Quality Index	55.69	56.61	-0.923***	0.12	-7.73
Daily kcal per capita	1994.01	2161.39	-167.4***	15.46	-10.83
Daily kcal from carbs, %	73.75	74.61	-0.856***	0.11	-7.69
Daily kcal from protein, %	11.05	11.44	-0.389***	0.04	-9.13
Daily kcal from fat, %	17.24	16.09	1.153***	0.09	12.69
Net total income	1969.00	1845.78	123.2**	39.05	3.16
Net money income	1294.64	1161.91	132.7***	35.75	3.71
Donations received	152.79	132.76	20.03***	4.85	4.13
Agricultural income	388.44	248.80	139.6***	11.51	12.13
Expenditure	1702.06	1647.24	54.81**	19.30	2.84
of which food expenditure, %	51.23	53.10	-1.862***	0.25	-7.59
of which autoconsumption, %	30.52	30.88	-0.361	0.30	-1.18
Female agricultural producers, %	18.41	22.81	-4.405***	0.60	-7.32
Land cultivated, ha	3.20	2.97	0.237	0.23	1.03
of which rented, %	22.77	17.79	4.972***	0.56	8.89
N.agr products	5.01	5.40	-0.388***	0.04	-10.22
Observations	7,470	13,005			

All variables are in per capita measures and weighted by household survey weights. Monetary variables are in real terms, at 2004 constant PEN. The value of agricultural production is computed per producer.

* Individuals are considered unemployed if they had no employment in the past week, do not have a job lined up, and are not home workers. The unemployment rate measures share of working age (15-65) members of the household that are unemployed.

2.4.2 Parallel trends

To inspect the validity of the parallel trends assumption I implement an event study approach with treatment leads and lags. Namely, I estimate the following regression for both the household and district level treatments:

$$Y_i = \alpha + \sum_{y=2004}^{2012} \beta_{2,y} \mathbb{1}(year = y) * Quinoa + \delta_2 X_i + \zeta_{2,t} + \eta_{2,m} + \theta_{2,p} + \lambda_{2,s} + \epsilon_{2,i} \tag{2.3}$$

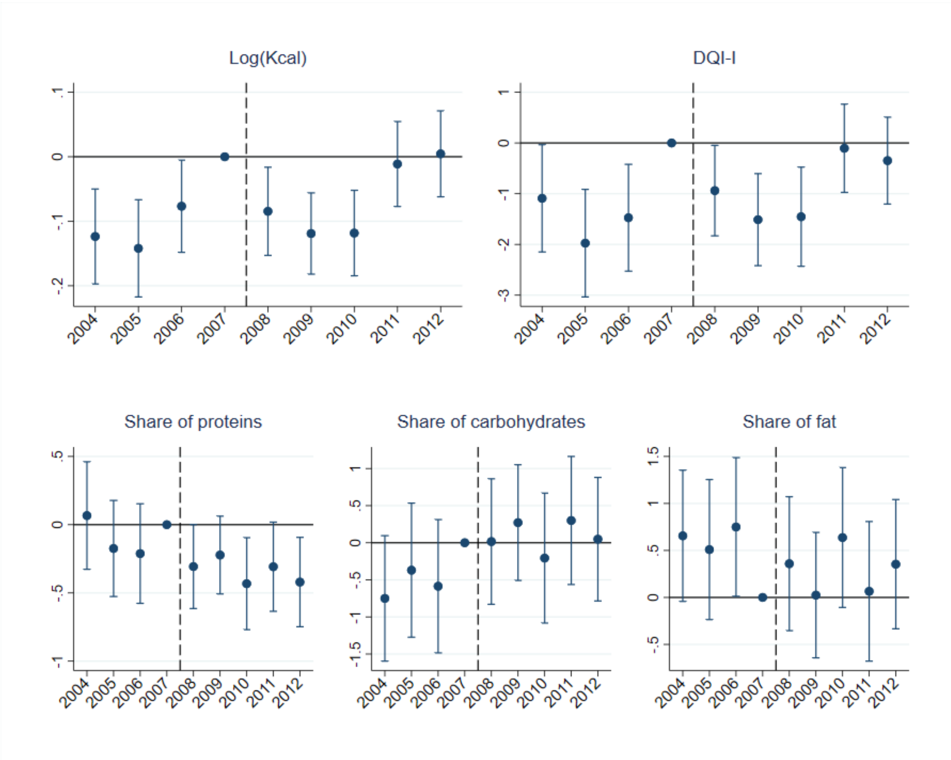
where $Quinoa$ corresponds respectively to $Quinoa_i$ or $Quinoa_{d,0}$.

Figure 2.4 plots the evolution of the main variables of interest, distinguishing between quinoa-farming and other farming households. and Panel 2.4a plots the treatment-year interaction coefficients from Eq.(2.3), with year 2007 as base level, and Panel 2.4b reports the marginal means for the treated and control households. Notably, per capita caloric intake and Diet Quality Index do not seem to follow a parallel trend across the two groups before 2008. In particular, the inspection of the marginal means in Panel 2.4b highlights that the difference in pre-trends originates from quinoa-farming households. This group shows increasing per-capita caloric consumption in 2006 and 2007; interestingly, this trend seems to invert exactly from 2008. Section 2.B of the Appendix reports the corresponding plots for per capita caloric consumption by food group and the diet quality subindices and further explores possible explanations. The spike in caloric intake appears to be related to a trend in grains, legumes, and vegetable consumption; however, it is difficult to pinpoint an exact cause for this phenomenon.[21] The caloric composition of the diet instead, measured by the share of total calories accounted for by protein, carbohydrates, and fat consumption, shows an evolution more compatible with the assumption of parallel pre-trends, as the lead estimates in Panel 2.4a are not statistically different from zero. The evolution of the main variables after the price shock is also telling: per capita caloric intake first flattens out, then starts increasing for both groups from 2011; overall diet quality instead increases for both groups from 2009 onwards. This similarity of pattern across groups after the increase in the price of quinoa might suggest the existence of a positive spillover effect on the control group.

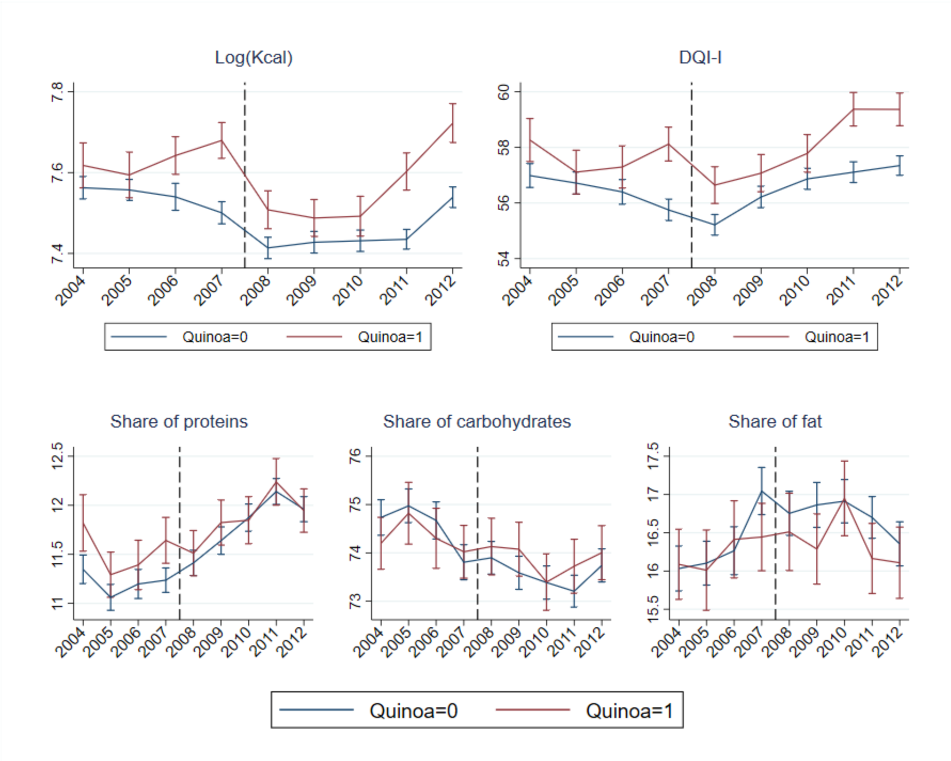
21. This pattern is not explained by obvious candidates such as self-consumption or outliers in the caloric intake distribution. It is also not explained by changes in sampling design, which in any case should affect both the treated and control units.

Figure 2.4: Parallel trends - individual treatment

(a) Event study coefficients



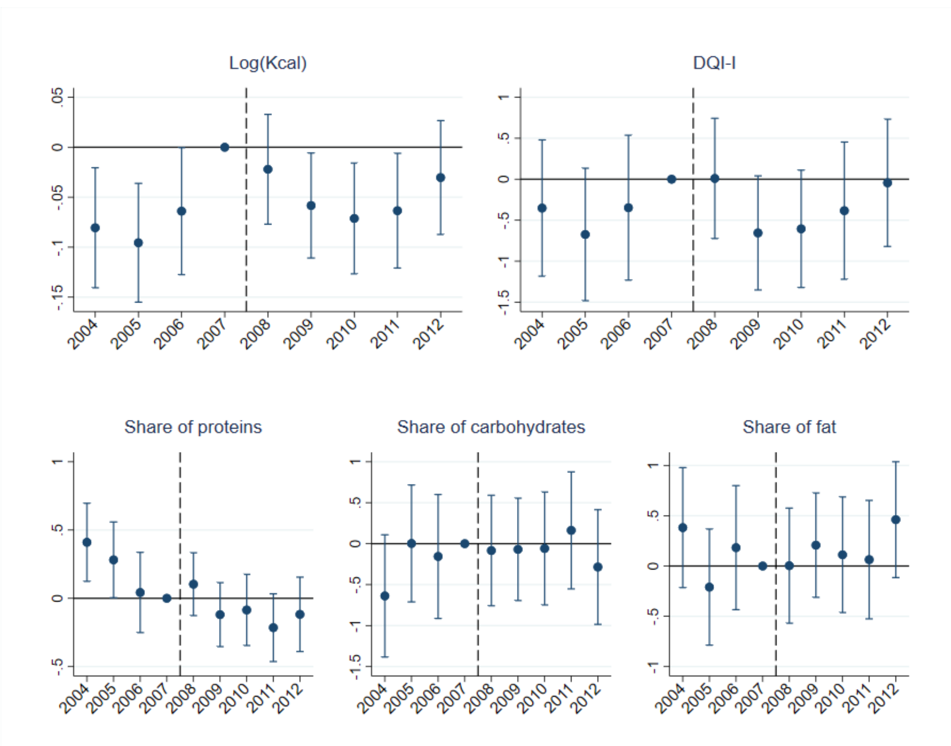
(b) Marginal means



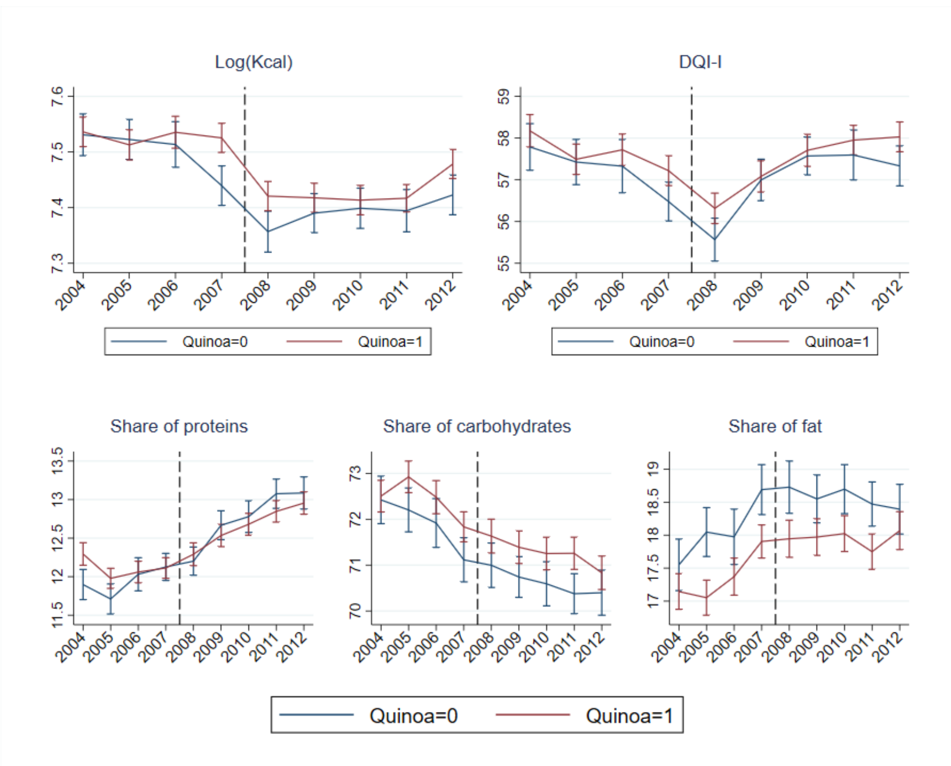
Note: 95% confidence intervals reported.

Figure 2.5: Parallel trends - district treatment

(a) Event study coefficients



(b) Marginal means



Note: 95% confidence intervals reported.

Figure 2.5 replicates the same plots for the aggregate treatment, i.e. comparing districts with and without quinoa production in the pre-boom period. From the lead estimates in Panel 2.5a, I observe a significant change in the difference in caloric intake and the calorie share of proteins in 2007 across treated and control districts, but overall diet quality and the other variables follow a parallel pre-trend. An inspection of the marginal means in Panel 2.5b reveals that the diverging behavior of total caloric intake is likely driven by the fact that caloric intake of the control districts starts declining in 2007, while it does not in the treated districts.

Overall, the failure to satisfy the parallel pre-trends assumption in some cases puts some limitations on the interpretation of estimates. For the household treatment analysis, this entails that I cannot interpret the DID estimates for the overall nutrition outcomes, caloric intake, and the diet quality index, causally. An inspection of their estimates, however, can still be informative in terms of the evolution of these variables over time. Moreover, I can still make a causal statement regarding the evolution of diet nutrient composition, in terms of caloric share. Given the concern of spillover effects within districts and the graphical evidence for parallel trends in diet quality, the following interpretation of results will rely primarily on the estimates for the district treatment.

2.5 Results

Table 2.3 reports estimates for the analysis considering the household level treatment, i.e. comparing quinoa farmers and other agricultural farmers. Panel 2.3a investigates the effect of the quinoa price increase on caloric intake, macronutrient composition, and diet quality. Results in columns (1) and (5) show that the onset of the quinoa boom does not correlate with higher per capita nutritional intake or diet quality in quinoa farming households, as the estimates are not economically nor statistically significant. However, the composition of caloric intake changes: columns (2-4) highlight that the share of calories from proteins consumed by quinoa-farming households significantly decreases by 0.25 percentage points (p.p.) after the quinoa price boom, at the 1% level (95% CI [-0.50;-0.07]), and the share from carbohydrates increases by 0.47 p.p. at the 5% level (95% CI [0.03;0.89]). The share of calories from fat consumption, instead, remains unaffected.[22] Moreover, Panel 2.3b investigates the effect of the quinoa boom on per capita food consumption, distinguishing across food categories. The price shock brings about a 10% reduction in caloric consumption of animal products, which is statistically significant at the 1% level (95% CI [-0.17;-0.03]), in favor of a 5% and 13% increase in the intake of calories from legumes and quinoa, respectively (95% CIs [0.005;0.07] and [0.048;0.217]). This pattern is likely responsible for the observed redistribution of caloric intake from proteins to carbohydrates observed in Panel 2.3a. Overall, this evidence points to the fact that, although the overall caloric intake of quinoa farming households was not affected by the increase in quinoa prices, this might be the result of offsetting changes in diet composition.

Table 2.4 reports the result for the comparison across quinoa-farming and other agricultural districts. Caloric intake and diet quality do not change significantly across districts of the same province with different agricultural specializations, but the share of calories from protein significantly decreases by 0.29 percentage points in quinoa-farming districts after the quinoa boom (95% CI [-0.14;-0.44]). Panel 2.4b suggests that this might be originated from lower consumption of protein-intensive food groups, such as animal products and quinoa. Though statistically significant, the estimated decrease in protein intake is of limited magnitude and indeed it does not affect overall diet quality.

Overall, these results point to a limited effect of the rise in quinoa prices on the nutrition of local Peruvian farmers, at least in the short-term, as neither total caloric consumption nor overall diet quality increase in either estimation approach.[23] Table 2.3 brings evidence that the consumption of quinoa increased among quinoa farmers compared to other farming households, which might be the result of multiple channels. First, quinoa farmers might have an increased taste of quinoa, once they become more informed of its nutritional qualities; this however seems at odds with the findings by Stevens (2017). Second, the increase in quinoa prices might have reduced the consumption of quinoa by its local purchasers, i.e. households that do not produce it themselves. This hypothesis is supported by the fact that, when comparing quinoa-producing districts to other agricultural districts

22. Note that the calories from proteins, carbohydrates, and fat do not necessarily add up to the total of calories consumed. However, these three constitute the main components of caloric intake: for 99% of observations, the sum of calorie share from these macronutrients corresponds to 100.

23. Tables 2.9 and 2.10 in the Appendix replicate this analysis without the use of survey weights, which results in qualitatively similar estimates.

Table 2.3: Household treatment

(a) Nutritional intake

	(1) Log(Kcal)	(2) Protein %	(3) Carb %	(4) Fat %	(5) Diet Quality
Quinoa_i, t	0.095***	0.310***	-0.167	-0.140	1.216***
	(0.015)	(0.071)	(0.165)	(0.141)	(0.218)
Quinoa_i, t*Post_08	0.015	-0.248**	0.466*	-0.169	0.295
	(0.019)	(0.092)	(0.220)	(0.188)	(0.258)
Obs.	44988	44988	44988	44988	44988

(b) Per capita log(kcal) - food groups

	(1) Animal	(2) Grains	(3) Legumes	(4) Quinoa	(5) Veg & Fruit	(6) Other
Quinoa_i, t	0.143***	0.073***	0.115***	0.191***	0.062*	0.031
	(0.029)	(0.016)	(0.028)	(0.030)	(0.025)	(0.021)
Quinoa_i, t*Post_08	-0.100**	0.046*	-0.051	0.133**	0.010	-0.075**
	(0.036)	(0.021)	(0.036)	(0.043)	(0.029)	(0.029)
Obs.	42560	44785	36521	12692	44494	43939

Panel a: diet composition is measured as the percentage of total caloric intake accounted for by proteins, carbohydrates, and fat consumption. Panel b: all dependent variables are in logarithm. The dependent variables measure daily per capita food consumption in grams.
Households are weighted by their survey weight. All regressions include year, province, population and survey month dummies. Standard errors in parentheses are clustered at the conglomerate-sampling unit level.
* $p < .05$, ** $p < .01$, *** $p < .001$

Table 2.4: District treatment

(a) Nutritional intake

	(1) Log(Kcal)	(2) Protein %	(3) Carb %	(4) Fat %	(5) Diet Quality
Quinoa_{d,t_0}	0.036*	0.144	0.556**	-0.721***	0.487*
	(0.015)	(0.075)	(0.174)	(0.135)	(0.206)
Quinoa_{d,t_0}*Post_{08}	-0.006	-0.292***	0.097	0.149	-0.164
	(0.015)	(0.075)	(0.186)	(0.149)	(0.214)
Obs.	71840	71840	71840	71840	71840

(b) Per capita log(kcal) - food groups

	(1) Animal	(2) Grains	(3) Legumes	(4) Quinoa	(5) Veg & Fruit	(6) Other
Quinoa_{d,t_0}	0.037	0.059***	0.060*	0.120***	-0.006	0.010
	(0.028)	(0.017)	(0.024)	(0.034)	(0.027)	(0.022)
Quinoa_{d,t_0}*Post_{08}	-0.062*	0.011	-0.047	-0.094*	-0.019	-0.035
	(0.027)	(0.017)	(0.026)	(0.038)	(0.024)	(0.022)
Obs.	68204	70882	57001	21889	70554	70050

Panel a: diet composition is measured as the percentage of total caloric intake accounted for by proteins, carbohydrates, and fat consumption. Panel b: all dependent variables are in logarithm. The dependent variables measure daily per capita food consumption in grams.
Households are weighted by their survey weight. All regressions include province, population, year and survey month dummies. Standard errors in parentheses are clustered at the conglomerate-sampling unit level.
* $p < .10$, ** $p < .05$, *** $p < .01$

within the same province in Table 2.4, caloric intake from quinoa consumption decreases.

2.6 Conclusion

Building on previous research, this paper investigates the effect of the sudden increase in the price of quinoa on Peruvian households in agricultural areas. I compare the effect of the quinoa boom of 2009 across districts specialized in quinoa production and agricultural districts not engaged in quinoa production before the price increase. To this end, I use yearly household survey data from the Peruvian Statistics Office from 2004 to 2012.

The results point to the fact that even though the price increase is substantial and reflects on the producer prices paid in local markets, there is not a significant effect on the farmers' diet, at least in the short term. Quinoa farming households retain a caloric intake and diet quality comparable to that of farmers of other crops after the price increase, although the composition of their diet changes trading lower protein for higher carbohydrates consumption. Moreover, overall diet quality, as measured by the DQI index, does not significantly change. These estimates are however potentially subject to selection into treatment, as farmers that have a higher caloric intake to start with might turn to quinoa production following the price boom.

When comparing households across districts that traditionally engage in quinoa production, there appears to be no significant difference in overall caloric intake and diet quality after the quinoa price boom. Quinoa-producing districts show a decrease in protein consumption after the price increase, which however is of limited magnitude.

Overall, the results point to the fact that the quinoa boom did not bring about a substantial improvement or worsening of diet quality for its traditional producers.

Future research could further examine the dynamics of income and quinoa production following the increase in demand for quinoa, and changes in quinoa consumption patterns. Moreover, the case of quinoa provides a privileged platform to investigate the spillover effects of the price increase of a specialized good and its potential effect on education and gender equality.

References

Alem, Y., and M. Söderbom. 2012. "Household-level consumption in urban Ethiopia: the effects of a large food price shock." *World Development* 40 (1): 146–162.

Anríquez, G., S. Daidone, P. Karfakis, and E. Mane. 2008. "Methodological Note: Estimating the Nutritional Impacts of Soaring Food Prices at the Household Level." *FAO, Rome.*

Appel, L. J., D. Baker, O. Bar-Or, K. Minaker, R. Morris, L. Resnick, M. Sawka, S. Volpe, M. Weinberger, and P. Whelton. 2005. "Dietary reference intakes for water, potassium, sodium, chloride, and sulfate." *Washington, DC: Institute of Medicine.*

Attanasio, O., V. Di Maro, V. Lechene, and D. Phillips. 2013. "Welfare consequences of food prices increases: Evidence from rural Mexico." *Journal of Development Economics* 104:136–151.

Bedoya-Perales, N. S., G. Pumi, A. Mujica, E. Talamini, and A. Domingos Padula. 2018. "Quinoa expansion in Peru and its implications for land use management." *Sustainability* 10 (2): 532.

Bellemare, M., J. Fajardo-Gonzalez, and S. R. Gitter. 2018. "Foods and fads: the welfare impacts of rising quinoa prices in Peru." *World Development* 112:163–179.

Blythman, J. 2013. "Can vegans stomach the unpalatable truth about quinoa." *The Guardian* 16 (01).

Carimentrand, A., A. Baudoin, P. Lacroix, D. Bazile, and E. Chia. 2015. "Quinoa trade in Andean countries: opportunities and challenges for family."

Deaton, A. 1989. "Rice prices and income distribution in Thailand: a non-parametric analysis." *The Economic Journal* 99 (395): 1–37.

Dreibus, T. 2014. "Quinoa rides the'superfoods' wave." *Wall Street Journal.*

Ecker, O., K. Weinberger, and M. Qaim. 2010. "Patterns and determinants of dietary micronutrient deficiencies in rural areas of East Africa." *African Journal of Agricultural and Resource Economics* 4 (311-2016-5545): 175–194.

FAO. 1997. "FAOSTAT statistical database."

———. 2015. "Catalogue of commercial varieties of quinoa in Peru."

Friedman-Rudovsky, J. 2012. “Quinoa: The dark side of an Andean superfood.” *Time Magazine.*

Gamboa, C., M. Schuster, E. Schrevens, and M. Maertens. 2017. *The Quinoa Boom and the Welfare of Smallholder Producers in the Andes.* Technical report.

Harttgen, K., S. Klasen, and R. Rischke. 2016. “Analyzing nutritional impacts of price and income related shocks in Malawi: Simulating household entitlements to food.” *Food Policy* 60:31–43.

Harvard T.H. Chan School of Public Health. 2019. “Superfoods or Superhype?” Available at https://www.hsph.harvard.edu/nutritionsource/superfoods/, *The Nutrition Source.*

Hasan, S. A. 2016. “The impact of the 2005–2010 rice price increase on consumption in rural Bangladesh.” *Agricultural Economics* 47 (4): 423–433.

Institute of Medicine, N. I. Krinsky, G. Beecher, R. Burk, A. Chan, j. J. Erdman, R. Jacob, et al. 2000. “Dietary reference intakes for vitamin C, vitamin E, selenium, and carotenoids.” *Institute of Medicine, US National Academy of Sciences.*

Institute of Medicine, R., Russell, J. L. Beard, R. J. Cousins, J. T. Dunn, G. Ferland, K. Hambidge, S. Lynch, J. Penland, A. Ross, B. Stoecker, et al. 2001. “Dietary reference intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc.” *Institute of Medicine, US National Academy of Sciences.*

James, L. E. A. 2009. “Quinoa (Chenopodium quinoa Willd.): composition, chemistry, nutritional, and functional properties.” *Advances in food and nutrition research* 58:1–31.

Kim, S., P. S. Haines, A. M. Siega-Riz, and B. M. Popkin. 2003. “The Diet Quality Index-International (DQI-I) provides an effective tool for cross-national comparison of diet quality as illustrated by China and the United States.” *The journal of nutrition* 133 (11): 3476–3484.

Loyer, J. 2016. “The social lives of superfoods.” PhD diss., University of Adelaide, School of Humanities.

McDonell, E. 2015. “Miracle foods: Quinoa, curative metaphors, and the depoliticization of global hunger politics.” *Gastronomica: The Journal of Critical Food Studies* 15 (4): 70–85.

Medicine, I. of. 1997. “Dietary reference intakes for calcium, phosphorus, magnesium, vitamin D, and fluoride.” *Institute of Medicine, US National Academy of Sciences.*

Mghenyi, E., R. J. Myers, and T. Jayne. 2011. “The effects of a large discrete maize price increase on the distribution of household welfare and poverty in rural Kenya.” *Agricultural Economics* 42 (3): 343–356.

Ministerio de Salud, n. 2014. “Bases Técnicas de las guías alimentarias para la población Boliviana.”

Ministerio de Salud de la Nación. 2016. “Guías Alimentarias para la Población Argentina.” Buenos Aires, Argentina.

Mišečka, T., P. Ciaian, M. Rajčániová, and J. Pokrivčák. 2019. “In search of attention in agricultural commodity markets.” *Economics Letters* 184:108668.

Montero, C., and C. Romero. 2017. “Análisis económico de la producción nacional de la quinua.” *Ministerio de Agricultura y Riego, Lima, Perú.*

Mujica, A. 2015. “El origen de la quınoa y la historia de su domesticación.” *Tierra Adentro* 108:14–17.

Nowak, V., J. Du, and U. R. Charrondière. 2016. “Assessment of the nutritional composition of quinoa (Chenopodium quinoa Willd.)” *Food chemistry* 193:47–54.

Ofstehage, A. 2012. “The construction of an alternative quinoa economy: Balancing solidarity, household needs, and profit in San Agustın, Bolivia.” *Agriculture and human values* 29 (4): 441–454.

Padulosi, S., K. Amaya, M. Jäger, E. Gotor, W. Rojas, and R. Valdivia. 2014. “A holistic approach to enhance the use of neglected and underutilized species: the case of Andean grains in Bolivia and Peru.” *Sustainability* 6 (3): 1283–1312.

Repo-Carrasco, R., C. Espinoza, and S.-E. Jacobsen. 2003. “Nutritional value and use of the Andean crops quinoa (Chenopodium quinoa) and kañiwa (Chenopodium pallidicaule).” *Food reviews international* 19 (1-2): 179–189.

Reyes García, M., I. Gómez-Sánchez Prieto, and C. Espinoza Barrientos. 2017. “Tablas peruanas de composición de alimentos.” Lima.

Romero, C. 2015. “Quinua Peruana Situación Actual y Perspectivas en el Mercado Nacional e Internacional al 2015.” *Informe Técnico. Lima-Perú: Ministerio de Agricultura y Riego, Dirección de Estudios Económicos e Información Agraria.*

Romero, S., and S. Shahriari. 2011. “Quinoa’s global success creates quandary at home.” *The New York Times* 19.

Stevens, A. W. 2017. “Quinoa quandary: Cultural tastes and nutrition in Peru.” *Food policy* 71:132–142.

Trumbo, P., S. Schlicker, A. A. Yates, and M. Poos. 2002. “Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein and amino acids.” *Journal of the Academy of Nutrition and Dietetics* 102 (11): 1621.

US Department of Agriculture. 2016. “USDA National Nutrient Database for Standard Reference.” *Agricultural Research Service, Nutrient Data Laboratory* Release 28.

APPENDIX

Appendix 2.A Descriptives

Subsection 2.A.1 Summary statistics

Table 2.1 compares characteristics across these groups at baseline. This section provides summary statistics covering the entire window of observation, i.e. years between 2004 and 2012, again distinguishing households by treatment status. This portrays the characteristics of the households throughout the sample, thus potentially also encompassing the effects of the increase in the price of quinoa.

Table 2.5 reports summary statistics for the main variables across quinoa-farming and other farming households, using the appropriate sampling weights. The two groups are comparable in terms of average household size, number of children, gender composition, age, and years of schooling, but quinoa households appear to have a lower unemployment rate. Quinoa farmers show higher caloric intake and diet quality on average, but the distribution of calories across proteins, carbohydrates, and fat is similar across the two groups. In terms of income, the average quinoa farming household has a lower per capita income than other agricultural farmers, but similar total expenditure and food expenditure. Quinoa farmers seem to be more inclined to auto consumption, which constitutes 40% of their food expenditure. Finally, both groups of farmers have a similar gender share, and farm around 4 hectares of land, of which 25% is rented. Quinoa producers tend to be geographically concentrated, as they live in districts where the majority of farmers produce quinoa. Table 2.6 reports summary statistics for the main characteristics of households across quinoa intensive and not-intensive districts, using the appropriate sampling weights.

Table 2.5: Summary statistics - agricultural farmers

	Control $Quinoa_{i,t} = 0$			Treated $Quinoa_{i,t} = 1$		
	mean	median	sd	mean	median	sd
HH_size	4.287	4	2.256	4.245	4	2.194
N. children	1.554	1	1.580	1.514	1	1.581
Within the HH						
Males, %	49.33	50	23.23	49.15	50	22.53
Unemployed, % *	30.23	0	62.95	24.70	0	56.76
Avg age	35.80	29.40	18.99	35.80	29.80	18.48
Avg years of schooling	3.868	3.400	3.132	4.321	4	3.177
Daily kcal per capita	1810.1	1668.6	922.4	1961.7	1819.6	894.1
Daily kcal from carbs, %	71.98	72.79	8.765	72.50	73.07	6.978
Daily kcal from protein, %	11.74	11.30	3.437	12.27	11.98	2.854
Daily kcal from fat, %	18.42	17.61	7.472	17.17	16.66	5.826
Diet Quality Index	65.94	66.13	9.117	67.65	67.75	8.155
Net total income	2318.7	1524.1	3017.2	2149.0	1504.6	2191.1
Net money income	1547.8	823.4	2686.3	1348.4	773.9	1847.7
Donations received	175.0	85.75	365.2	144.2	65.29	406.8
Agricultural income	299.2	132.4	7207.5	310.8	161.1	822.3
Expenditure	1966.0	1551.1	1578.3	1930.2	1590.9	1342.4
of which food expenditure, %	50.30	51.01	16.99	52.91	53.73	16.50
of which autoconsumption, %	34.00	31.29	22.62	40.71	39.86	20.63
Female agricultural producers, %	22.81	0	41.96	27.49	0	44.65
Land cultivated, ha	3.101	1	16.55	4.090	1.125	13.16
of which rented, %	23.96	0	38.49	25.43	0	38.66
N.agr products	5.226	5	2.948	7.144	7	2.738
District - agr producers, %	80.96	87.50	23.85	82.21	89.00	22.93
District - quinoa producers, %	9.628	0	16.83	56.40	57.14	27.52
Households		30,598			6,126	
Observations		37,957			7,050	

All variables are in per capita measures and weighted by household survey weights. Monetary variables are in real terms, at 2004 constant PEN. The value of agricultural production is computed per producer.
* Children are household members younger than 15.
* Individuals are considered unemployed if they had no employment in the past week, do not have a job lined up, and are not home workers. The unemployment rate measures share of working age (15-65) members of the household that are unemployed.

Table 2.6: Summary statistics, HH data by district treatment

	Control $Quinoa_{d,pre} = 0$			Treated $Quinoa_{d,pre} = 1$		
	mean	median	sd	mean	median	sd
HH_size	4.250	4	2.248	4.099	4	2.233
N. children	1.597	1	1.581	1.518	1	1.578
Men in the HH, %	50.01	50	23.46	48.45	50	24.27
Unemployed, %	30.34	0	62.86	28.93	0	62.39
Avg age in the HH	35.78	29.75	18.69	36.93	30.25	19.64
Avg years of schooling	3.923	3.500	3.113	3.966	3.500	3.260
Net total income	2390.7	1533.1	2995.8	2329.7	1559.1	3155.7
Net money income	1608.6	838.8	2666.1	1539.5	827.9	2857.1
Donations received	193.1	92.08	435.8	169.7	81.50	362.1
Agricultural income	682.7	272.7	2032.1	521.4	258.1	1415.6
Expenditure	2005.0	1562.9	1643.9	2009.0	1605.2	1561.6
of which food expenditure, %	49.89	50.67	16.95	51.44	52.23	16.90
of which autoconsumption, %	33.00	30.60	22.66	35.05	33.03	22.66
Agricultural producers, %	96.28	100	18.92	96.33	100	18.79
Female agricultural producers, %	20.63	0	40.46	25.71	0	43.71
Quinoa producers, %	1.734	0	13.05	27.44	0	44.62
of which female, %	18.24	0	38.68	26.99	0	44.39
of which sell quinoa, %	0.140	0	3.737	2.890	0	16.75
Land cultivated, ha	2.946	1	14.79	3.317	0.910	15.31
District share agr producers, %	82.17	90.43	23.47	79.88	87.50	24.75
District share quinoa producers, %	1.755	0	6.943	27.33	16.67	28.88
Daily kcal per capita	1773.6	1646.2	897.9	1913.4	1762.7	943.7
Daily carbs per capita	317.0	295.1	164.7	346.2	319.2	174.8
Daily fat per capita	37.19	33.49	22.39	37.40	33.61	21.97
Daily protein per capita	51.61	45.96	30.24	57.75	51.84	32.85
Diet Quality Index	65.76	66.07	9.234	66.85	67.00	8.839
Observations		16,141			24,747	
Households		12,583			19,487	
Districts		391			342	

All variables are in per capita measures and weighted by household survey weights. Monetary variables are in real terms, at 2004 constant PEN. The value of agricultural production is computed per producer.
* Children are household members younger than 15.
* Individuals are considered unemployed if they had no employment in the past week, do not have a job lined up, and are not home workers. The unemployment rate measures share of working age (15-65) members of the household that are unemployed.

Subsection 2.A.2 Sample selection - balance of characteristics

The present study restricts the sample to areas in the Sierra macro-region of Peru, particularly the rural districts. Table 2.7 reports summary statistics and mean comparison t-tests for the Sierra and other macro-regions, and within the Sierra across rural and urban areas. The Sierra has a higher concentration of farmers and quinoa farmers, while in other regions the share of quinoa farmers is negligible. On the other hand, the average per capita net income, caloric intake, and overall diet quality are significantly lower in the Sierra. Within the Sierra macro-region, rural areas show a higher concentration of farmers and quinoa farmers than urban areas, not surprisingly, but also statistically significantly lower per capita net income, caloric intake, and overall diet quality.

Table 2.7: Sample selection

(a) Macro regions

	Other regions	Sierra	Diff	T-stat
Agricultural producers, %	32.3	62.6	-30.259***	-126.75
of which quinoa producers, %	0.2	15.7	-15.512***	-71.80
Net total income	5326.8	3680.2	1646.526***	51.65
Daily kcal per capita	2132.7	1972.8	159.867***	30.25
Diet Quality Index	58.0	57.2	0.809***	17.08
Observations	88140	71904		

(b) Urban vs rural

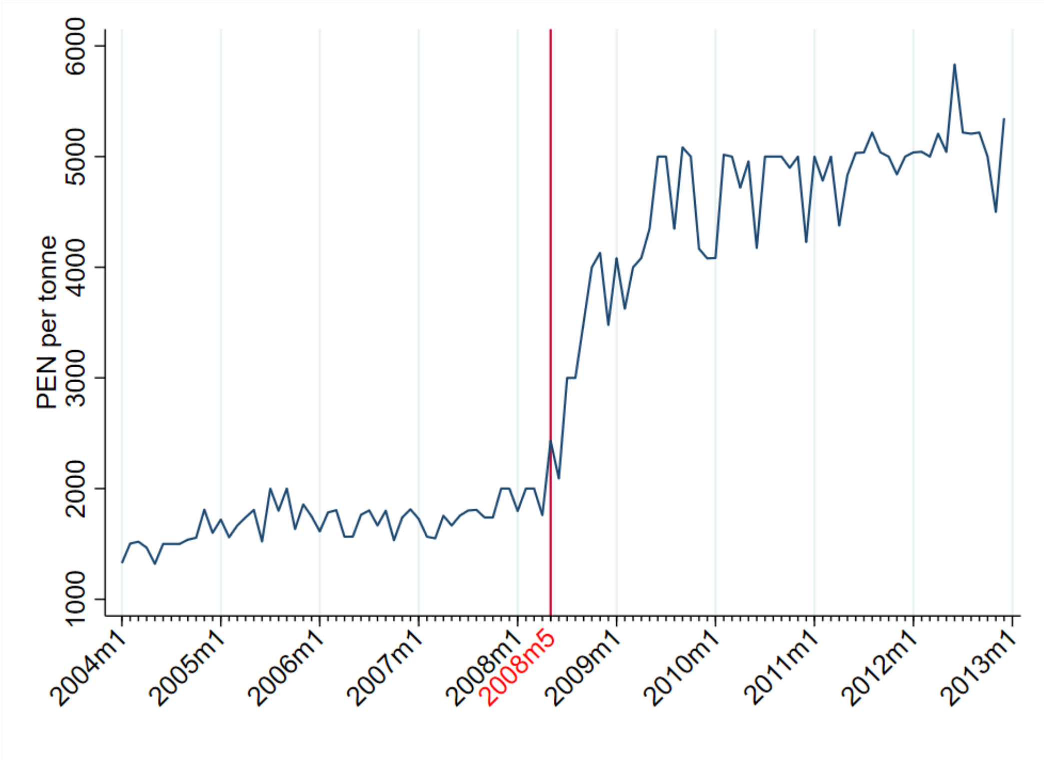
	Urban sierra	Rural sierra	Diff	T-stat
Agricultural producers, %	32.9	90.1	-57.147***	-195.99
of which quinoa producers, %	13.3	16.5	-3.165***	-8.04
Net total income	5467.8	2025.7	3442.117***	81.90
Daily kcal per capita	1976.0	1969.9	6.043	0.76
Diet Quality Index	58.8	55.7	3.060***	45.23
Observations	34562	37342		

Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$
All variables are in per capita measures; monetary variables are in real terms, at 2004 constant PEN.

Subsection 2.A.3 Quinoa - production and consumption

Figure 2.6 reports the monthly unit value of quinoa, as reported by farmers in the ENAHO survey. The graph highlights the sudden rise in the price of quinoa, as already illustrated by Figure 2.1, and allows to further pin down the start of the price boom around May 2008; the phenomenon intensifies from July 2008.

Figure 2.6: Quinoa unit value by month



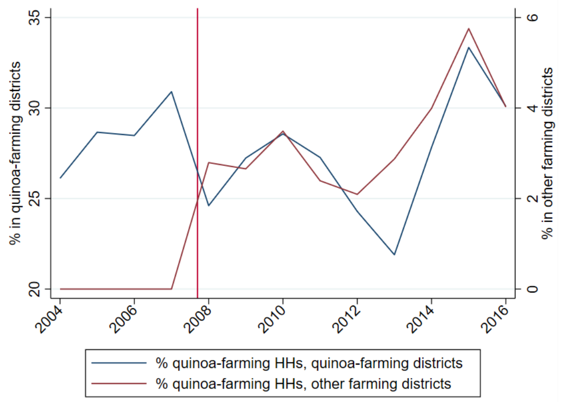
Source: ENAHO.

Quinoa production, in the sample considered, concentrates among small producers, where median production is 25 kg per household/year, while from 2014 onwards the number of large-scale producers increased, even in the mountainous areas. The use of produced quinoa varies with the size of the production: as production increases, the share of production dedicated to sales increases and the share consumed decreases, while seeding and other uses remain flat. After 2008, both the quantity of quinoa produced and the share of produced kg intended for sale increase.

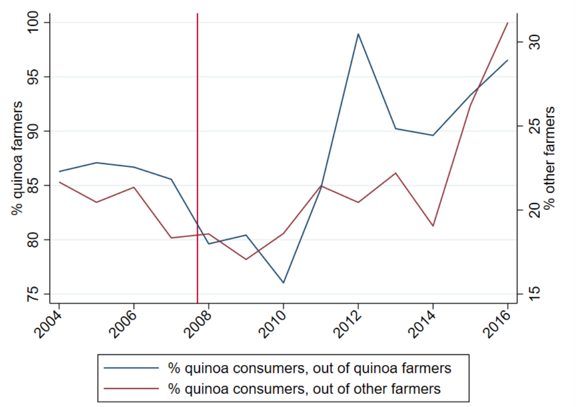
Figure 2.7 portrays the evolution of quinoa production and consumption for households residing in the Peruvian Highlands. Panel 2.7a reports the share of quinoa-farming households evolves over time, distinguishing between districts where quinoa farming is carried out or not before 2008. By definition, the share of quinoa producers is positive in quinoa-farming districts before the 2008 price increase, standing at around 25-30%, and zero in non-quinoa-farming districts. After 2008, I observe quinoa-farming households also in (previously) non-quinoa-farming districts, but this phenomenon is very limited - they represent 6% of households at most. In both types of districts, the share of quinoa-farming households noticeably increases from 2013, corresponding to the International Year of Quinoa, again suggesting that only then the quinoa production side starts adjusting to the change in prices (and also responding to the stimulation packages introduced, as described in Section 2.E). Panel 2.7b shows that, unsurprisingly, quinoa consumption is much more frequent among quinoa farmers than among other farmers.; panel 2.7c then delves into the composition of quinoa consumers in terms of quinoa production status. This suggests that the composition of quinoa-consuming households remains relatively stable over time: every year, households that consume quinoa are composed of 22% quinoa farmers, 30% other farmers, and the remaining of non-agricultural households. Among quinoa farmers, the majority of households are net consumers of quinoa. Finally, Panel 2.7d explores the intensive margin of quinoa consumption. Quinoa consumption, measured in daily grams per capita, remains almost constant for quinoa producers during the first years of the quinoa boom, only to decrease dramatically in 2012, while it decreases for the other farmers. Moreover, note that quinoa consumption constitutes on average 5% of the daily kcal intake for quinoa producers, and only 2% for other farmers. We can conclude that although it is culturally important, quinoa is not a staple food.

Figure 2.7: Quinoa production and consumption

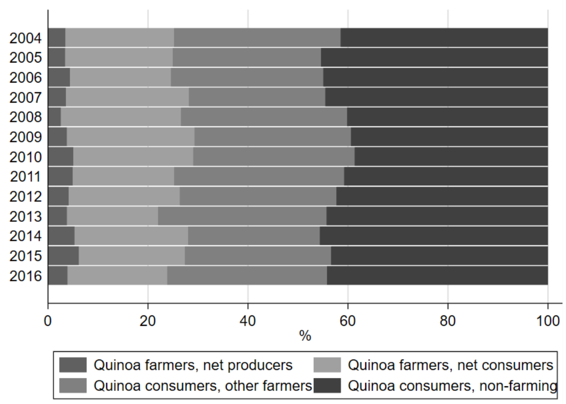
(a) Share of quinoa producers



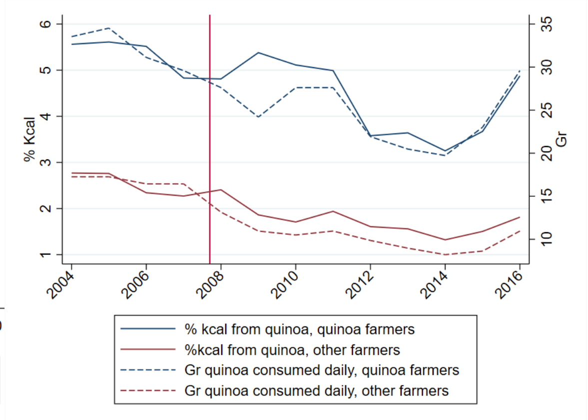
(b) Share of quinoa consumers



(c) Quinoa consumers: composition



(d) Median consumption



Appendix 2.B Parallel trends

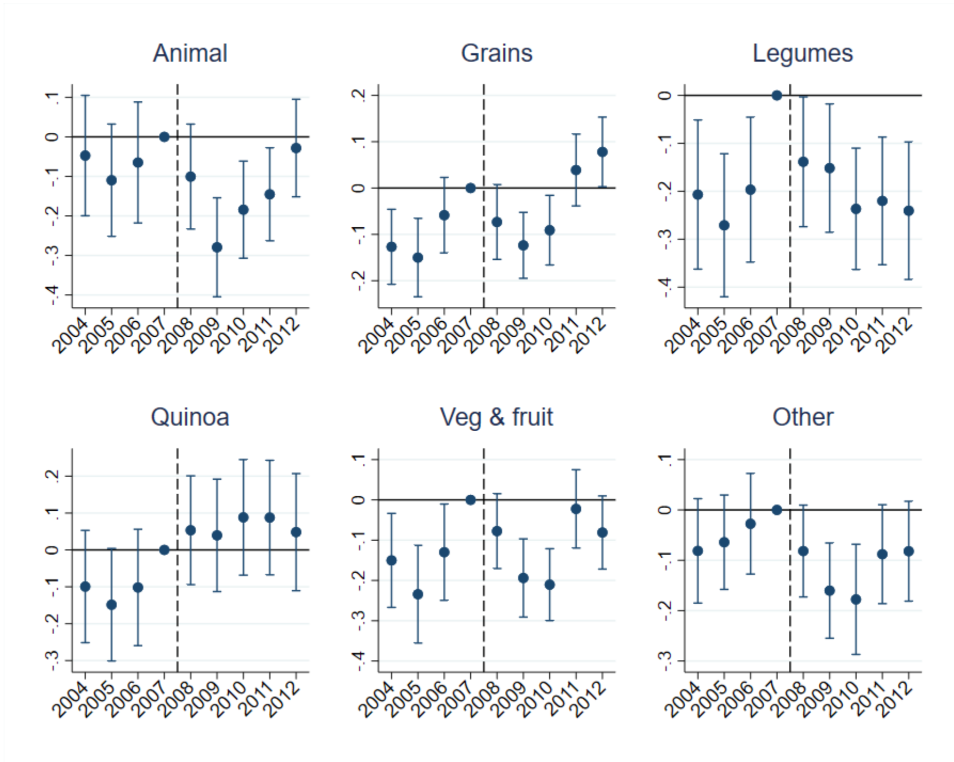
Figure 2.8 unpacks the evolution of nutritional intake over time into the consumption of different food groups and the components of the Diet Quality Index across quinoa farmers and other farmers. From Panel 2.8a we can identify the sudden rise in the consumption of grains, legumes, vegetables and fruits in 2007 as the driving force behind the peak in overall per capita caloric intake. The diet quality subindices further indicate that the gap between quinoa and other farmers in both diet adequacy, moderation, and variety is significantly larger in 2007 than in previous years. These patterns are not explained by obvious candidates such as self-consumption or outliers in the caloric intake distribution. Similarly, they are not related to changes in survey design, which would affect both the control and treated group.

Figure 2.9 replicates this exercise across districts. Again, it is an increase in the consumption of grains, legumes, vegetables and fruits in 2007 the likely cause of the rise in overall per capita caloric intake. As a consequence, diet adequacy and variety in 2007 increased faster in traditionally quinoa farming districts than elsewhere, while the gap in diet moderation decreased and diet balance followed a parallel trend.

Finally, Table 2.8 addresses potential self-selection into quinoa farming across the years 2006-2008, by performing pair-wise yearly t-test comparisons. All in all, the 2007 sample of treated households does not appear to be systematically different in terms of household characteristics and income. In particular, in 2007 quinoa-farming households have on average a statistically significantly lower share of male household members than in 2006, while income is comparable, although in 2007 they receive more donations (both from public institutions, churches, and private agents). In terms of expenditure, in 2007 quinoa farming households spend more per capita and in particular have a 14 p.p. higher share of food expenditure coming from auto consumption. In 2008 instead, the gender composition of households returns to the 2006 levels, while income rises compared to 2007, most likely as a consequence of the quinoa price increase

Figure 2.8: Parallel trends - individual treatment

(a) Food groups



(b) DQI subindices

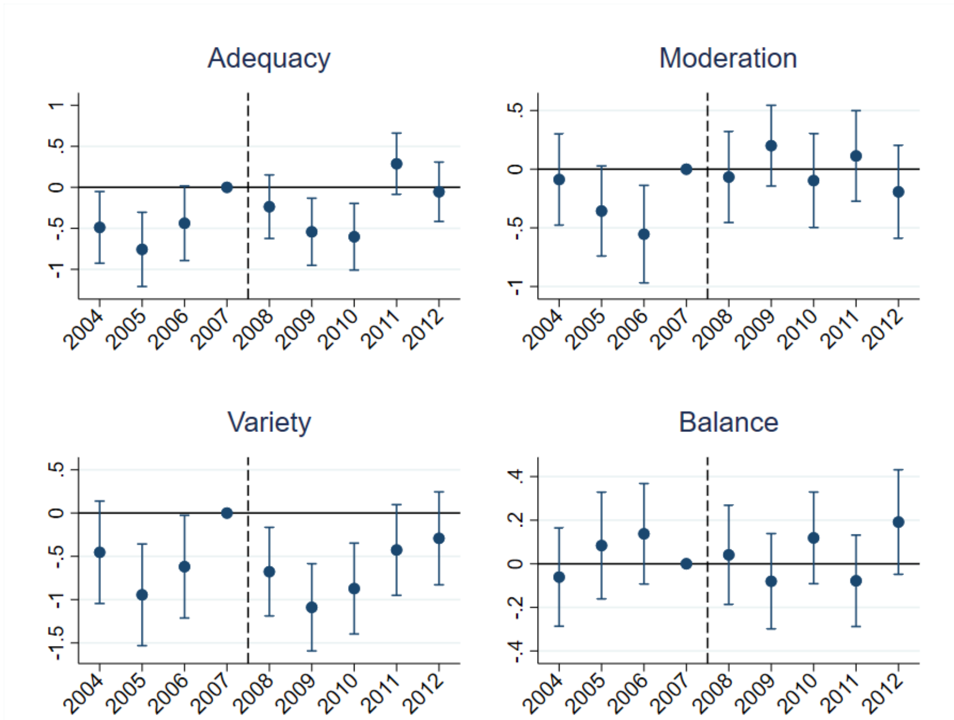
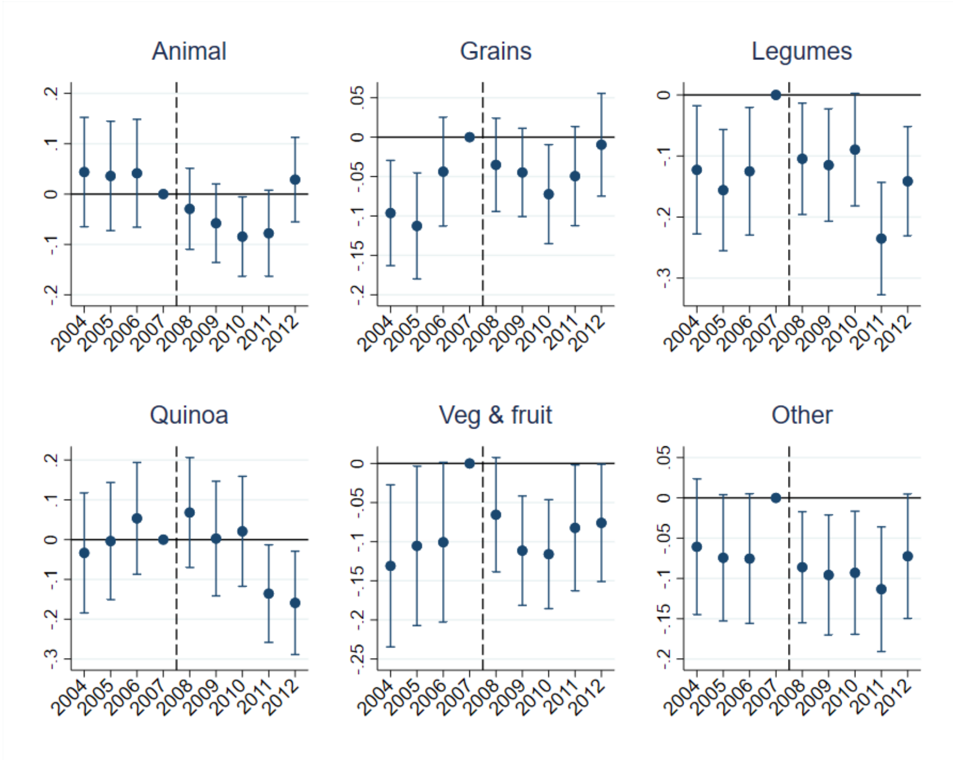


Figure 2.9: Parallel trends - district treatment

(a) Food groups



(b) DQI subindices

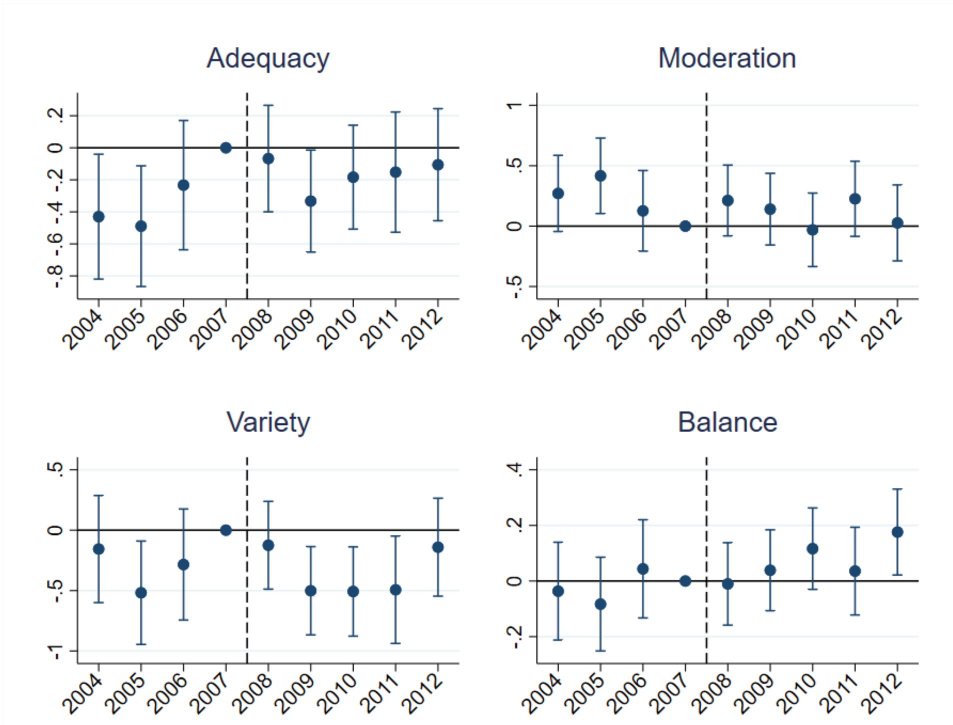


Table 2.8: Sample characteristics in 2006-2008, quinoa producers - Sierra

	2006	2007	Diff: 2006-2007	2008	Diff: 2007-2008
HH_size	4.37	4.36	0.014	4.42	-0.061
N. children	1.65	1.72	-0.072	1.64	0.082
Male, %	50.16	48.29	1.867	50.26	-1.961
Avg age	34.74	33.33	1.408	34.47	-1.138
Avg years of schooling	4.18	4.05	0.131	4.14	-0.093
Unemployed, %	5.84	5.19	0.650	6.19	-1.001
Male	0.84	0.80	0.046*	0.84	-0.043*
Age	50.91	50.17	0.739	51.04	-0.869
Years of schooling	5.16	5.05	0.104	4.82	0.235
Net total income	1728.24	1805.71	-77.465	2132.89	-327.179**
Net money income	1034.83	1077.36	-42.536	1356.02	-278.657**
Donations received	104.10	173.25	-69.146*	174.71	-1.464
Agricultural income	303.69	289.92	13.773	324.97	-35.044
Expenditure	1609.09	1737.64	-128.551*	1822.00	-84.364
of which food expenditure, %	53.29	55.72	-2.437**	52.70	3.020***
of which autoconsumption, %	30.08	44.13	-14.051***	48.41	-4.287***
Female agricultural producers, %	23.24	26.43	-3.194	21.17	5.258*
Land cultivated, ha	4.98	5.07	-0.089	2.79	2.283
of which rented, %	18.97	21.79	-2.821	18.97	2.812
N.agr products	6.78	7.30	-0.515***	7.28	0.021
Observations	723	874		751	

Standard errors in parentheses. * $p < .10$, ** $p < .05$, *** $p < .01$

Appendix 2.C Other results

Tables 2.9 and 2.10 report the main results without weighting households by their survey weights. The estimates are qualitatively the same, where the caloric intake and diet quality do not significantly change after the quinoa price increase, neither for the household or district treatment. Estimates regarding diet composition are somewhat different but also tell a similar story.

Appendix 2.D Nutrition estimation

Subsection 2.D.1 Estimation of food and nutrient consumption

Household food consumption is measured using a two-week recall survey in the ENAHO survey, which includes food obtained from own production, purchases, gifts, or in-kind transfers. To calculate nutrient consumption amounts from the reported food quantities, I employ the food-nutrient conversion tables published by the Peruvian Center for Food and Nutrition (Reyes García, Gómez-Sánchez Prieto, and Espinoza Barrientos 2017). For foods that have no corresponding entry in the Peruvian nutrition data, specifically hard liquor and sweets, I use data from the USDA National Nutrient Database for Standard Reference (US Department of Agriculture 2016). Moreover, information on fat content is obtained from the Argentinean Dietary Guidelines (Ministerio de Salud de la Nación 2016), allowing to distinguish between saturated, monounsaturated, and polyunsaturated fatty acids. In total, I match a list of 661 unique food items.[24][25]

For each household, I compute quantities and nutrient consumption for the overall diet and six basic food groups (grains and tubers, quinoa, legumes, vegetable and fruit, animal products, and other meal complements). This allows measuring the overall caloric intake, but also the consumption of macronutrients such as protein, fat, and carbohydrates [26], and micronutrients such as vitamins and minerals. In addition, I

24. This is a more detailed food basket than that considered by previous literature, as the analysis by Stevens (2017) is limited to 160 items. This is partially, but not exclusively, related to the fact that I do not exclude drinks from the analysis. This choice is motivated by the assumption that sugary drinks can constitute a substantial portion of a Peruvian household's caloric intake.

25. For some foods, data were imputed based on the nutrient composition of similar food items. For instance, some specific fish species were imputed as having the same macronutrient content as the average fresh fish; the same procedure was followed for different types of eggs.

26. Carbohydrates refer to available carbohydrates, i.e. the non-fiber portion of carbohydrates.

compute the caloric share of each macronutrient, by converting the consumed amount from grams to calories.[27] To account for household size, I consider the daily consumption in grams of the average household member, dividing the total household consumption by the number of individuals in the household.[28] Since food consumption is measured at the household level, I assume that food is distributed within the household according to the needs of each member.

Table 2.9: Household treatment, no survey weights

(a) Nutritional intake

	(1) Log(Kcal)	(2) Protein %	(3) Carb %	(4) Fat %	(5) Diet Quality
Quinoa_i,t	0.107***	0.370***	-0.103	-0.211	1.450***
	(0.013)	(0.067)	(0.152)	(0.127)	(0.192)
Quinoa_i,t*Post_08	-0.004	-0.300***	0.326	0.001	-0.001
	(0.017)	(0.089)	(0.202)	(0.168)	(0.230)
Obs.	44988	44988	44988	44988	44988

(b) Per capita log(kcal) - food groups

	(1) Animal	(2) Grains	(3) Legumes	(4) Quinoa	(5) Veg & Fruit	(6) Other
Quinoa_i,t	0.146***	0.090***	0.149***	0.209***	0.058*	0.008
	(0.027)	(0.014)	(0.026)	(0.028)	(0.023)	(0.020)
Quinoa_i,t*Post_08	-0.096**	0.023	-0.078*	0.074	0.005	-0.074**
	(0.033)	(0.019)	(0.032)	(0.039)	(0.027)	(0.028)
Obs.	42560	44785	36521	12692	44494	43939

Panel a: diet composition is measured as the percentage of total caloric intake accounted for by proteins, carbohydrates, and fat consumption. Panel b: the dependent variables measure the ratio of daily protein, carbohydrate, fiber intake to the respective recommended allowances. Panel c: all dependent variables are in logarithm. The dependent variables measure daily per capita food consumption in grams.

All regressions include year, province, population and survey month dummies. Standard errors in parentheses are clustered at the conglomerate-sampling unit level.

* $p < .05$, ** $p < .01$, *** $p < .001$

27. Following USDA data, carbohydrates provide 4 calories per gram, protein provides 4 calories per gram, and fat provides 9 calories per gram.

28. This includes long-term guests that stay for longer than 20 days, whose food consumption would be reflected in the 2-week recall food data

Table 2.10: District treatment - no survey weights

(a) Nutritional intake

	(1) Log(Kcal)	(2) Protein %	(3) Carb %	(4) Fat %	(5) Diet Quality
Quinoa_{d,t_0}	0.045**	0.055	0.704***	-0.751***	0.314
	(0.015)	(0.078)	(0.164)	(0.123)	(0.197)
Quinoa_{d,t_0}*Post_{08}	-0.005	-0.341***	-0.124	0.356**	-0.290
	(0.016)	(0.076)	(0.169)	(0.131)	(0.197)
Obs.	71840	71840	71840	71840	71840

(b) Per capita log(kcal) - food groups

	(1) Animal	(2) Grains	(3) Legumes	(4) Quinoa	(5) Veg & Fruit	(6) Other
Quinoa_{d,t_0}	-0.025	0.072***	0.079***	0.116***	-0.049	-0.012
	(0.029)	(0.016)	(0.024)	(0.031)	(0.027)	(0.021)
Quinoa_{d,t_0}*Post_{08}	-0.035	0.002	-0.062**	-0.095**	-0.008	-0.019
	(0.024)	(0.017)	(0.023)	(0.033)	(0.022)	(0.021)
Obs.	68204	70882	57001	21889	70554	70050

Panel a: diet composition is measured as the percentage of total caloric intake accounted for by proteins, carbohydrates, and fat consumption. Panel b: the dependent variables measure the ratio of daily protein, carbohydrate, fiber intake to the respective recommended allowances. Panel c: all dependent variables are in logarithm. The dependent variables measure daily per capita food consumption in grams.
All regressions include province, population, year and survey month dummies. Standard errors in parentheses are clustered at the conglomerate-sampling unit level.
* $p < .10$, ** $p < .05$, *** $p < .01$

These estimates of food and nutrient consumption are based on retrospective answers, which are subject to measurement error. Concerns of low data accuracy due to the 2-week recall period are limited, since every year about 95% of food purchases reported in the ENAHO survey happen with bi-weekly frequency or more, while only the remaining 5% is represented by monthly or annual purchases.[29] While, on the one hand, this approach overestimates food consumption by not accounting for food waste, on the other hand, it underestimates it by assuming that the food basket of each household is limited to the foods they reported obtaining in the previous two weeks, with a constant consumption throughout the year: the overall net measurement error is ambiguous. Measurement error generated by seasonality in food consumption should be a minor concern, as the survey runs continuously throughout the year. In addition, the use of month of survey dummies in the regression helps to control for the seasonality of food consumption.

Subsection 2.D.2 Estimation of diet quality

To have a comprehensive measure of diet quality, I compute the Diet Quality Index - International (henceforth DQI-I) developed by Kim et al. (2003). This measure considers four main aspects of diet quality: food variety, adequacy of nutrient intake, moderation of unhealthy nutrients, and overall diet balance. The final index is additive across these dimensions and ranges from 0 to 100, with higher values indicating a higher quality diet. The final diet quality index is essentially a weighted average of the subindices: diet adequacy is the main component (worth up to 40 points), followed by diet moderation (up to 30 points), diet variety (up to 20 points), and diet balance (up to 10 points). For a detailed description of the construction of the index, see Table 1 of Kim et al. (2003).

The adequacy subindex evaluates the intake of dietary elements that must be supplied sufficiently to guarantee a healthy diet, assigning scores based on the percentage attainment of the recommended intakes on a continuous scale. Recommended intakes for vegetables, fruit, and grains are measured as the number of servings consumed per day, while grams per day are used for fiber intakes, and the score assignment distinguishes between three levels of energy intake (below 1700, below 2700, above 2700 kcal/day): for individuals with a lower caloric intake, a lower number of servings is sufficient to obtain the maximum adequacy score. Intake of protein is considered

29. These summary statistics are likely to reflect the actual frequency of purchase: monthly or annual purchases can be reasonably assumed to be more salient or sizeable than frequent purchases, so there is no reason why they would have a higher probability of being forgotten and thus not reported in the survey.

adequate when the proportion of total energy intake that comes from protein is higher than 10%. Finally, the adequate intake of iron, calcium, and vitamin C is derived by comparing consumption to the Recommended Daily Amount (RDA). I use the standard recommendations compiled by the US Institute of Medicine (Institute of Medicine et al. 2000; Institute of Medicine et al. 2001; Medicine 1997) to calculate household-specific reference recommendations, taking household size, age, and sex composition into consideration.

The moderation subindex evaluates the consumption of nutrients that are unhealthy when consumed in large quantities, such as different types of fat or empty-calorie foods. Empty calorie foods are foods that provide only energy but insufficient nutrients, identified by a total nutrient density lower than one. I compute the ratio of consumption to Recommended Daily Amount (RDA) for each nutrient[30] , then take the sum of these ratios and divide it by the ratio of consumed to recommended calories to obtain a measure of nutrient density. When this measure takes a value on one the consumption of overall nutrients is balanced with respect to the energy intake, compared to the recommended levels; a value lower than one indicates that the food is nutritionally poor, compared to the number of calories it provides.

The food variety score considers the variety of food groups and protein sources consumed, taking into account whether the individual consumed at least one serving per day of each group.[31]

Finally, to assess the overall balance of the diet, the index considers the proportionality in energy sources and fatty acid composition. In particular, it considers the macronutrient ratio, i.e. the distribution of consumed calories across consumed carbohydrates, protein, fat, and the distribution of fatty acids consumed (polyunsaturated, monounsaturated, saturated).

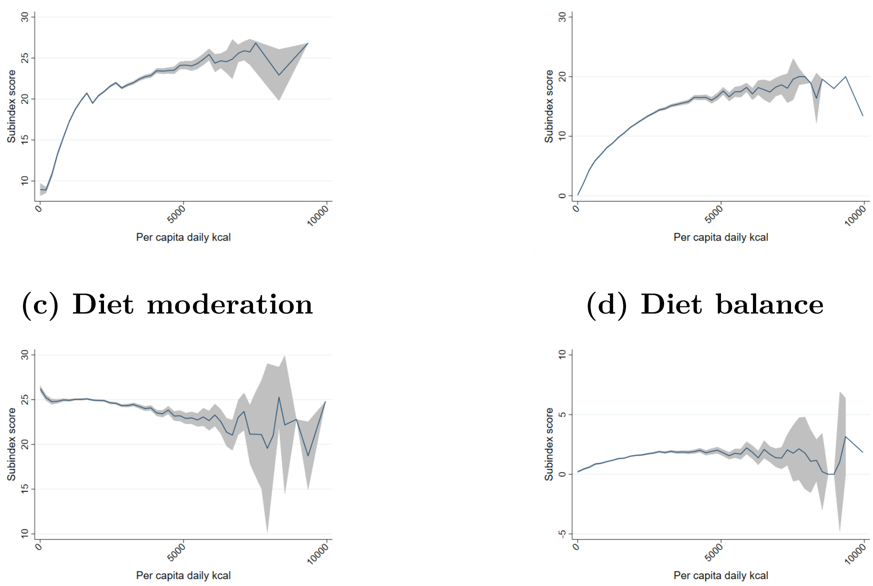
30. I consider the following nutrients, which are the ones that can be measured using the Peruvian conversion tables: protein, calcium, iron, phosphorus, zinc, potassium, vitamin A and C, thiamin, riboflavin, niacin, folate.

31. Serving sizes are taken from the Bolivian nutritional guidelines (Ministerio de Salud 2014).

Subsection 2.D.3 Diet index and caloric intake

Figure 2.10 illustrates the relationship between the diet quality subindices and per capita daily caloric intake. Taken together, the figures confirm that the DQI-I index is a reasonable measure of diet quality. Diet adequacy and variety increase with caloric intake until they reach a plateau. Unsurprisingly, diet moderation decreases with daily caloric intake. Finally, diet balance follows an inverted U-shape: it initially rises with caloric intake, and starts decreasing for excessive levels of daily kcal.

Figure 2.10: DQI-I subindices and caloric intake



(a) Diet adequacy **(b) Diet variety**

(c) Diet moderation **(d) Diet balance**

Notes: the graph reports local polynomial smooth plots using the epanechnikov kernel and the corresponding 95% confidence interval.

Appendix 2.E Institutional initiatives

Nutritional outcomes and poverty rates of Peruvian households have long been under the attention of policy makers, with both governmental and international organizations engaging in a sustained effort to improve the situation.

Several institutional initiatives were active in rural areas before the quinoa demand boom. These programs generally correspond to initiatives aimed at improving nutritional quality and reducing income inequality, and those targeted at rural development. Among these, the conditional cash-transfer program Juntos (from 2005), the National Food Assistance program (PRONAA, from 1992), and the Good Start in Life Program are worth mentioning. Although not aimed specifically at quinoa farmers, these programs could have a confounding effect on the estimates, as they could have disproportionately affected quinoa farmers. In particular, quinoa-producing districts might have been overrepresented in the intervention sample insofar as they correspond to the poorest districts. Moreover, the impact of nutrition-oriented programs might have been heterogeneous due to the higher availability of nutrient-dense foods such as quinoa. To this end, Table 2.11 performs mean test comparisons in poverty and assistance rates between quinoa-farming and other farming households before 2008. Quinoa-farming households show higher poverty rates, but the two groups do not exhibit significant differences in terms of reception of government cash transfers, participation in the Juntos program, and food assistance rates. Secondly, there were several initiatives targeting agricultural producers, such as financing through the development bank AGROBANCO; projects funded by the rural development agency AGRORURAL, some of which focused specifically on Andean rural populations; or the development of farming business plans through the competitiveness program AGROIDEAS (from 2010). Unfortunately, participation in these programs is not measured by the ENAHO survey.

Efforts aimed at improving nutritional quality and rural development intensified after the rise in demand for quinoa. However, most of these programs became operative in the second phase of the boom, corresponding to the second spike in prices in 2013. In 2011, the FAO and the UN Assembly designated 2013 the International Year of Quinoa (IYQ). The initiative aimed at increasing the visibility of quinoa and its potential for achieving global food security, the development of sustainable production systems, and the appreciation of indigenous communities as keepers of this grain. The activities planned during the IYQ ranged from scientific conferences and innovation competitions to regional meetings of quinoa producers aimed at understanding the market opportunities and limitations of the existing production chain and the establishment of organizations of quinoa farmers. Moreover, quinoa consumption was promoted through food festivals and the diffusion of recipe books focused on quinoa and its nutritional value, both at the national and international level, including the inclusion of quinoa in the school meal program Qali Warma. By the end of 2013, the agricultural bank AGROBANCO had funded quinoa farmers for more than 27 million PEN. The information sharing and gathering that occurred during the IYQ activities then served as a solid foundation for the design of tailored development programs.
Consequently, most of these initiatives only became operative in the second phase of the boom, corresponding to the second spike in prices in 2013. These programs were

aimed at helping quinoa producers bridge the gap between producers and international markets, by selecting appropriate varieties and improving agricultural techniques, but also educating local populations with regard to the nutritional value of quinoa. Programs aimed at improving the production and quality of Andean grains (Mi Riego, Buena Siembra) started operating in 2013; the ILO/FAO sponsored program Economic inclusion and sustainable development of Andean grain producers in Ayacucho and Puno started in 2014; and in 2015 the Andean Grains program, supported by the Sustainable Development Goals Fund.

Table 2.11: Assistance rates at baseline, Peruvian Highlands

	Quinoa-farming	Other farming	Diff
Below poverty threshold, %	76.55	78.17	-1.614*
Receiving public cash transfers, %	71.67	72.62	-0.946
Receiving cash transfers through Juntos, %	7.75	7.56	0.186
Receiving food aid, %	61.69	62.01	-0.317
Observations	16668	3188	

Chapter 3

Executive Gender Quotas and Social Services: Evidence from Italy

joint with Alice Dominici and Olivia Masi

3.1 Introduction

Does female representation alone increase the provision of public services that would otherwise burden women in their private lives? Previous literature has shown that the cost of child and elderly care is disproportionately shouldered by women (OECD et al. 2014). Does then the vested interest of women in these services translate into policy outcomes, once they obtain political representation?

In this paper, we causally identify the effect of increased female representation within local executive bodies in small Italian municipalities, relying on exogenous variation provided by the introduction of Law n. 56/2014. We first investigate whether gender quotas effectively translated into higher female representation. Next, we test whether this increased female representation generated higher public spending on social services, with a particular focus on childcare and elderly care provision. This law mandated that no gender represent less than 40% of the municipal executive in municipalities with more than 3,000 residents. Using a difference-in-discontinuities (henceforth diff-in-disc) approach à la Grembi, Nannicini, and Troiano (2016), we find that the share and the number of women in municipal executives increased in line with the mandate, thus confirming the enforcement of the law. However, we find no significant increase in current and capital expenses on social services in general, or on specific subgroup items such as childcare and elderly assistance.

The study of gender quota policies can be of particular interest in contexts with low female participation and with rationed social services, i.e. with low provision relative to the resident population, such as Italy. Italy is characterized by low female participation in the labor market and by an uneven division of housework between men and women (Del Boca et al. 2021). Several empirical studies on Italy have pointed out that facilitating access to early childcare services would be very effective in allowing households to reconcile family and work, and incentivizing women's participation in the labor market (Del Boca 2002; Bratti, Bono, and Vuri 2005; Del Boca, Locatelli, and Vuri 2005; Del Boca and Vuri 2007; Del Boca and Sauer 2009). Moreover, Giorgetti and Picchio (2021) highlight that in Italy, as in many countries where childcare is rationed, female labor participation responds more strongly to the provision of public childcare, rather than the price of private services (Wrohlich 2004; Del Boca and Vuri 2007; Vandelannoote et al. 2015). Of particular importance is the fact that in Italy local executive authorities (municipalities) are in charge of policies concerning early childcare services. They constitute the most granular level of local government, and often they are the direct providers of childcare services (Giorgetti and Picchio 2021).

This paper relates to a literature on the role of political gender representation in shaping local policies; Hessami and Fonseca (2020) provide a literature review on the matter. In India, higher female political representation has been found to increase the provision of public goods and infrastructure (Bhalotra and Clots-Figueras 2014; Chattopadhyay and Duflo 2004). Clots-Figueras (2011) shows that these effects largely rely on politicians' socio-economic background on top of gender: lower caste female politicians favor redistributive laws and invest more in health and early education, while female politicians from higher castes reduce social expenditure.[1] In contrast, in high-income countries higher female representation has not been found to affect the overall size and composition of municipal spending (Ferreira and Gyourko 2014; Bagues and Campa 2021). In the case of Italy, Rigon and Tanzi (2012) find no correlation between female representation in city councils and municipal expense, and Baltrunaite et al. (2019) show that a large increase in the share of female councilors, induced gender quotas on candidates lists, did not significantly affect the structure of public expenditure. Our work adds to these previous studies by adding evidence from a high-income country and by considering an exogenous shock to female representation within local executive

1. In the same context, female representation in political bodies has been found to affect a wide array of other outcomes such as education (Clots-Figueras 2012), voter attitudes (Beaman et al. 2009), political participation (Bhalotra, Clots-Figueras, and Iyer 2018), and reporting of crimes against women (Iyer et al. 2012).

committees rather than electoral candidates. Although already studied in the context of developing countries (Chattopadhyay and Duflo 2004), imposing gender quotas on local executive bodies is an uncommon policy choice in high-income countries. According to the IDEA Gender Quotas Database (2021), other European countries have put in place either gender quotas only for electoral candidates (Portugal, Spain, France) or rely on voluntary gender quota policies from the individual political parties (Germany, Austria, Netherlands, Norway).

In addition, recent literature has shown that female representation can induce changes in parliamentary deliberations and specific policy choices in gender-sensitive topics that may not be reflected in the overall composition of public spending, though the evidence is mixed (Lippmann 2020; Gago and Carozzi 2021). For instance, Hessami and Baskaran (2019) find that a female victory accelerates the expansion of public childcare provision by 40% and increases discussion of childcare policies during council meetings. We align with this literature by focusing on expenditure on gender-sensitive policies, and extend the scope of the analysis by considering not only childcare but also elderly care and other elements of publicly provided social services, that have been overlooked by previous studies.

More generally, this paper relates to the literature on the effects of gender quotas on decision making and economic performance, also in firm settings. Evidence considering gender quotas for top company positions is not conclusive (Ahern and Dittmar 2012; Matsa and Miller 2013; Eckbo, Nygaard, and Thorburn 2016; Ferrari et al. 2021). In this regard, this paper contributes to the ongoing discussion on the effectiveness of gender quotas and their consequences.

Finally, this work is naturally related to studies that examine the same reform (Law n. 56/2014). Broberg and Panizza (2021) investigate the effect of the same reform on politicians' incentives, focusing on the effect of third mandate eligibility on local expenditures. They find that mayors in their first or second term at the time of the reform adjust revenues and expenditures differently, respectively investing in medium-term versus short-term spending items. However, they do not find a statistically nor economically significant effect of the term limit extension on social expenditure. De Benedetto and De Paola (2019) find that the extension of the mayoral term limit induced lower electoral participation, driven by a decrease in political competition.

The remainder of the paper proceeds as follows. Section 3.2 describes the institutional background and the data employed. Section 3.3 explains the conceptual framework, and Section 3.4 discusses our empirical strategy. Section 3.5 presents the main results and identification checks. Finally, Section 3.7 concludes.

3.2 Institutions and Data

3.2.1 Institutional Framework

The electoral system currently regulating municipal elections in Italy was introduced in 1993. Each municipality is composed of a mayor (*Sindaco*), an executive body (*Giunta*), and an elected city council (*Consiglio Comunale*). The mayor is directly elected with a simple majority rule and its electoral list is assigned two-thirds of council seats. The mayor then appoints the members of the executive council (*Assessori*) and designates each as responsible for a specific aspect of municipal affairs. In municipalities with less than 15,000 residents, these executive councilors need to be chosen among elected council members, unless otherwise provided by the municipal charter. They act as local ministers in their specific area of jurisdiction, proposing new bylaws and supervising the corresponding branch of local government. The majority of bylaw proposals then need to be approved by the city council with a public vote during a council session before they are adopted.

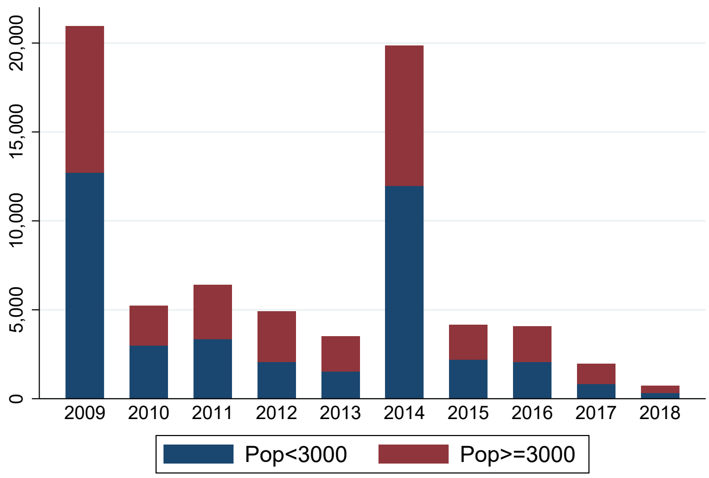
A term lasts five years, with early termination in case of resignation of the mayor or a vote of no confidence from the city council. Although all municipalities held their first election together in 1946, the occurrence of early terminations over time has generated a staggered election cycle, as shown by Figure 3.1.[2] Italian municipalities manage about ten percent of total public expenditure and are in charge of a wide range of services, including welfare, social services, waste management, municipal police, and infrastructure (Grembi, Nannicini, and Troiano 2016).

2014 Reform. In this paper, we focus on an exogenous change in female representation in the local executive body in Italy that was introduced by National Law n. 56/2014. This law had the general purpose of reordering the territorial organization of the country, and in doing so introduced a wide range of changes. Relevant to our research question, it mandated that no gender represent less than 40% of the municipal executive committee (including the mayor) in municipalities with more than 3,000 residents. In addition, it introduced the possibility of a third mandate for mayors of municipalities

2. Nowadays, early terminations are a relatively rare occurrence: between 2000 and 2018, only 7% of municipalities did not complete their five-year term.

with less than 3,000 residents and changed the number of seats in the council and executive committee differentially by population size. How we deal with these competing policies for identification is discussed in Section 3.4.1. The law entered into force in April 2014, with each municipality implementing it at their next election. This timing of implementation is plausibly exogenous because it is determined by the previous election cycle, i.e. pre-determined with respect to the policy change. Figure 3.1 shows the distribution of election timing across treatment groups, considering municipalities between 1000 and 5000 residents.

Figure 3.1: Histograms of the election timing



Among municipalities of population between 1000 and 5000 residents.

3.2.2 Data Sources and Sample Selection

We employ three main sources of data, namely the Local Administrators Registry, the Population Census, and a database of municipal finances collected by the Ministry of Interior.

Data on the composition of local executives come from the Local Administrators Registry collected by the Ministry of the Interior.[3] The female share in the local executive captures the average share of female members of the executive committee within a mandate, including new appointments and resignations. The resident population is measured from the 2001 and 2011 Censuses, making the 2011 Census population the relevant measure for the application of the gender quota policy.

Our main dependent variables are budgetary outcomes of a sub-sample of Italian municipalities in the period 2009-2018, from a database of municipal finances collected

3. Anagrafe degli Amministratori locali e regionali.

by the Ministry of Interior. We focus on expenditure on social services, particularly four expense groups: elderly care; services for persons with disabilities; childcare services; and other assistance such as services for families or assistance to marginalized groups. The main analysis considers expenses in the accrual account only, as these are more closely linked with the current administration; in the Appendix, we report the results for total accrued expenses.[4]

Our dataset also contains additional information on municipalities, such as total or rural area size, and individual characteristics of the elected mayor and council. Finally, we also have data from ISTAT on the number of authorized childcare places per child of age 0-2, the number of nurseries, and the share of expenses covered by the municipality for such services, which however only covers years from 2013 to 2016.

Our sample considers Italian municipalities with less than 5,000 residents and with local elections in 2014. The population restriction is justified by the fact that above the 5,000 residents cutoff different rules apply concerning various policies, as described in Section 3.2.1. Regression results refer to a further selected sample in the neighborhood of the 3,000 residents cutoff, using an optimally-selected bandwidth for each specification. In addition, we restrict our focus to municipalities that participate in the 2014 election cycle. These are municipalities where the gender quota policy had to be implemented shortly after it entered into force (April 2014), since elections took place in May 2014, thus attenuating any concerns of anticipation behavior.[5] Moreover, a quick look at Figure 3.1 reveals that this election cycle is the most populated one: 68% of municipalities within a 5-year cycle and with less than 5000 residents have elections in 2014 (and, in the previous election cycle, in 2009).

The final sample thus covers 3,052 municipalities, with less than 5,000 residents and local elections in 2009 and 2014, observed yearly between 2009 and 2018.

3.2.3 Descriptive Statistics

Table 3.1 reports some descriptive statistics for the final sample described above. Note that this includes all municipalities with less than 5,000 residents and local elections in

4. Total expenses are divided into the accrual and residual account. Whereas the accrual account records expenses relative to the current budgeting period (in this case, a year), the residual account records expenses from previous years that are still due.

5. A remaining concern is that in this subsample candidate mayors might not have had sufficient time to screen the best female candidates for their executive council, given the short amount of time available before elections.This would imply that the pool of female candidates in the 2014 elections is of lower quality than in the following election cycles, where candidate lists were more carefully scrutinized. This can be checked by analyzing the characteristics of candidates across different election cycles.

2009 and 2014 observed between 2009 and 2018, while the regression results are based on a more local sample selected around the threshold of 3,000 residents by data-driven bandwidth selection methods.

With an average resident population of $1,793$ inhabitants, the municipalities under study have an average municipal council and executive size of 12 and 4 members, respectively. The share of female representation in both institutions is around 23-25%, with a large standard deviation of up to 20 percentage points. At the executive council level, this corresponds to having around one female member. The average per capita revenue and expenditure by municipalities both stand at around $1,700$, with substantial dispersion. Regarding expenditures in the social sector, these consist mostly of current expenses, with a mean expense of 51 euros per resident, while capital expenses stand at around 6 euros per resident; moreover, both current and capital expenses exhibit large standard deviations. The level of current expenditures varies across different categories of social services: general assistance and elderly care are the highest with a mean of respectively 28 and 10 euros per capita, while for childcare and services for persons with disabilities the expense is only 1.41 and 4.45 per resident, respectively, potentially reflecting the population composition. Capital expenditures in specific services have a smaller magnitude than the respective current expenditures, with elderly care and other assistance again exhibiting higher per capita expenses than other categories. Overall, the large standard deviation of different spending items informs us that expenditures can vary widely among different municipalities and different periods.

Table 3.2 compares characteristics across the control and the treatment group in the years preceding the reform. The covariates we consider are the share of women in municipal executive and in the council, the share of population between 0 and 2 years old, the share of population above 65 years old, the area in hectares, the urban area in hectares, the per capita total expenditures, the per capita total revenues, and the per capita fiscal revenues. Pooling together the years 2011-2013, we first compute the average characteristics for the control and treated groups. Then, we compute the absolute standardized difference between them. As a rule of thumb, this standardized difference is considered significant if higher than 0.10. Based on this, the differences displayed in Table 3.2 are generally not significant, except for the share of population between 0 and 2 years old and the share of population above 65 years old. The treated municipalities have slightly more children and fewer elderly residents than the control group.

Table 3.1: Summary statistics

	Mean	Median	SD
Resident population, 2011 Census	1793.6	1486.0	1281.8
Municipal council size	12.5	13.0	2.2
Share of women in municipal council	25.3	25.0	12.3
Municipal executive size	4.0	4.0	1.2
Number of women in municipal executive	0.9	1.0	0.8
Share of women in municipal executive	23.3	20.0	20.1
Total revenues, per capita	1713.4	1235.3	1928.3
Total expenditure, per capita	1709.2	1235.3	1933.7
Current exp. in social services, aa, per capita	51.0	35.6	84.3
Current exp. in childcare, aa, per capita	4.4	0.0	12.5
Current exp. in serv. for disabled, aa, per capita	1.4	0.0	6.3
Current exp. in elderly care, aa, per capita	10.4	0.0	70.6
Current exp. in other assistance, aa, per capita	28.3	22.7	40.8
Capital exp. in social services, aa, per capita	5.6	0.0	43.8
Capital exp. in childcare, aa, per capita	0.2	0.0	2.8
Capital exp. in serv. for disabled, aa, per capita	0.2	0.0	10.9
Capital exp. in elderly care, aa, per capita	0.9	0.0	18.2
Capital exp. in other assistance, aa, per capita	1.3	0.0	33.7
N. of observations		30311	
N. of municipalities		3052	

Notes: * aa stands for accrual account

Table 3.2: Pre-treatment characteristics' balance

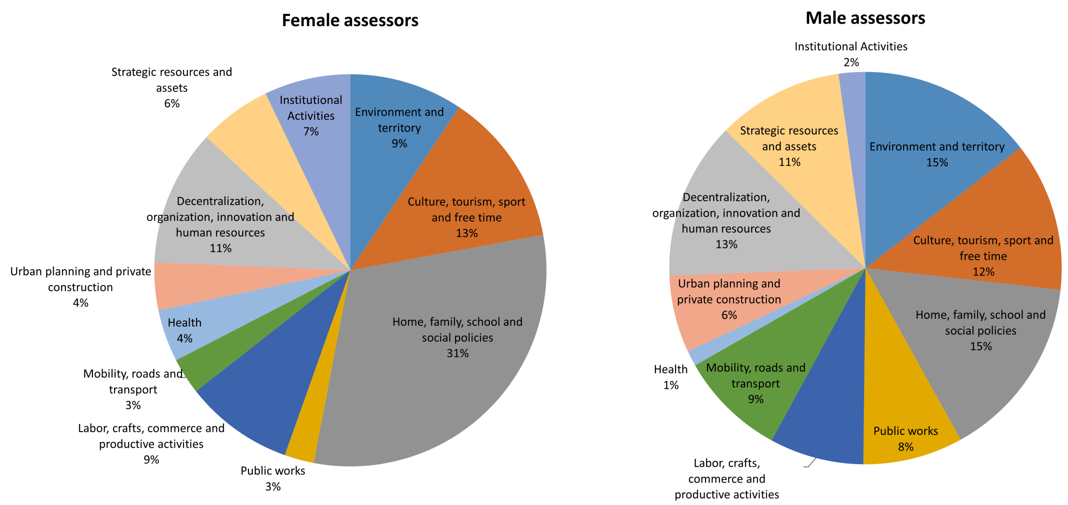
Covariate	Absolute st. distance	Control	Treated
Share of women in municipal executive	0.053	19.0	20.3
Share of women in municipal council	0.061	21.5	22.4
Share of population 0-2 y.o.	0.126	2.7	2.8
Share of population above 65 y.o.	0.147	21.9	21.2
Area in hectares	0.019	3307.8	3038
Urban area in hectares	0.04	1848.9	425.8
Total expenditures, per capita	0.077	1216.3	1157
Total revenues, per capita	0.075	1209.8	1151.3
Fiscal revenues, per capita	0.043	530.8	522.5

Notes: Averaging the characteristics of treated units (municipalities with population above 3000 inhabitants) and control units (municipalities with population below 3000 inhabitants) over the period 2011-2013, the Absolute Standardized Difference (ASD) for variable X is computed as: $ASD \equiv \frac{|\bar{X}_T - \bar{X}_C|}{\sqrt{Var_T(X)+Var_C(X)}}$, where T and C indicate the treated and control group, respectively.

3.3 Conceptual Framework

As female representation in institutions increases, there are various channels through which we might expect them to affect policy choices, ranging from preferences for redistribution (Chattopadhyay and Duflo 2004; Alesina and La Ferrara 2005; Funk and Gathmann 2015) to different management of resources.[6] Gender differences in policy choices may also stem from a gender gap in rent-extracting behavior, though causal evidence is lacking for high-income countries (Hessami and Fonseca 2020). In the case of Italy, there is evidence that female members of the executive council are disproportionately assigned to family and social policy matters within the executives of large municipalities (ANCI 2015).[7] This fact is summarized by Figure 3.2, and might be the result of either self-selection of female politicians or power relations within executive teams. Unfortunately, there is no coherent data on the attribution of administrative branches to members of the executive council, so these results cannot be directly extended to small municipalities at the moment. We might however expect this phenomenon to be even more salient in smaller municipalities.

Figure 3.2: Assignment to policy departments, by gender of executive councillors



On the other hand, Rigon and Tanzi (2012) find no significant relationship between female representation in city councils and municipal expenses, although this is only correlational evidence. They explain this result with the median voter theorem:

6. Funk and Gathmann (2015) document heterogeneous preferences for policies across genders for Switzerland with women favoring policies supporting the environment, persons with disabilities, and public healthcare. Other evidence also shows that female political representation in national parliaments influences climate change policy outcomes (Mavisakalyan and Tarverdi 2019).

7. In 2015, the National Association of Italian Municipalities (ANCI) collected data on members of the executive council for 136 large municipalities.

politicians' preferences and personal characteristics do not matter in public choices, only voter preferences are relevant to this end. In this regard, gender quota policies can only have an impact on policy outcomes insofar as they affect voter turnout.

Finally, increases in female representation could also affect governments' mode of operation. Studies on Italy show that female mayors have shorter appointments when the council is dominated by male politicians (Gagliarducci and Paserman 2012), but higher female representation in municipal councils increases government stability (Acconcia, Ronza, et al. 2021). Accettura and Profeta (2021) also show that female politicians have a different political budget cycle, ending up in fewer deficit pre-elections, by comparing mixed-gender close races in Italy. This evidence again suggests that the effect of gender quotas might extend to other dimensions beyond the magnitude of expenditures.

3.4 Empirical Strategy

We exploit an exogenous change in female representation in the municipal executive committee in Italy, mandated by Law n. 56/2014 described in Section 3.2.1. The enforcement of the gender quota for municipalities with more than 3,000 residents only allows us to use a difference-in-discontinuities (henceforth diff-in-disc) approach à la Grembi, Nannicini, and Troiano (2016). The reason for this methodological choice is that the same population threshold had been used previously for the introduction of different policies. Estimating a simple regression discontinuity (RD), therefore, would identify the compound effect of all policies changing discontinuously at that threshold. In a nutshell, the intuition behind the diff-in-disc estimator is that it subtracts the confounding effect of all previous policies by taking the difference between two RDs estimated at different points in time. In the rest of this section, we define the estimator formally in our setting.

3.4.1 Confounding Policies and Treatments' Definition

Our analysis focuses on the period between 2009 and 2018, employing yearly data on municipalities with elections in 2009 and 2014. The 2014 administrative elections were the first ones following the introduction of gender quotas in municipal executive committees, by Law n. 56/2014, for municipalities with a population above 3000 inhabitants. The same law, however, also introduced a second policy, namely the possibility for mayors to be re-elected for a third mandate for municipalities with a population below 3000 inhabitants. Note that the 2014 law also increased the number of council seats for municipalities above 3000 inhabitants. Finally, the reform under study also modified council and executive size, thus affecting our policy-relevant variable, the female executive share. Therefore, throughout our discussion, we also consider how the absolute number of women in the executive changes, showing that the observed change in the executive female share is not driven only by changes in the denominator. Moreover, the reform is such that the number of women necessary to comply with the 40% gender share rule only changes for the control group and not the treated group, which would if anything attenuate our estimates. The reported evidence shows that instead, there is limited (voluntary) take-up of the gender quota policy by control municipalities, reassuring us that this is not the case. [8]

For the sake of clarity, we will follow the same notation used by Grembi, Nannicini, and Troiano (2016). Formalizing the information above, let us denote:

- P_{it} the population of municipality i in year t. The policy population cutoff is defined as $P_c \equiv 3000$;
- $W_{it} \in \{0,1\}$ the binary indicator summarizing all of confounding policies implemented at the threshold P_c before 2014 (i.e., as of 2009). $W_{it} = 1$ denotes municipalities above the population threshold;
- $M_{it} \in \{0,1\}$, the confounding policy of the mayor's third mandate introduced in 2014. $M_{it} = 1$ if municipality i's mayor at time t is allowed to have a third mandate (i.e. it is observed to be equal to 1 if $P_{it} < 3000$ and $t \geq 2014$, 0 otherwise);
- $G_{it} \in \{0,1\}$, the gender quotas' policy of interest. In particular, $G_{it} = 1$ if municipality i at time t is mandated to have at least 40% city council seats assigned to women (i.e. it is observed to be equal to 1 if $P_{i,2014} \geq 3000$, 0 otherwise). Clearly, since these policies apply on opposite sides of the cutoff, we have that for $t = 2014$ $G_{it}^{obs} = 1 - M_{it}^{obs}$.

8. Section 3.B.1 further elaborates on the legislation regarding council and executive size for different population thresholds, and how the specifics of the 2014 reform affect our estimation strategy.

3.4.2 Potential Outcomes, Assumptions, and Estimator

Given the treatments outlined above, we define average potential outcomes as $Y_{it}(W_{it}, M_{it}, G_{it})$. This is the average amount of women in municipal city councils (as detailed below, both in percentage and absolute terms), in municipality i and year t, that corresponds to a set of values (observed or unobserved) of W_{it}, M_{it}, G_{it}.

Finally, following again Grembi, Nannicini, and Troiano (2016)'s notation, let $\tilde{Y}_i$ denote outcomes when $t < 2014$, so that we can avoid to use subscript t, and let superscripts $(+, -)$ denote limits as the population approaches the cutoff P_c, such that:

$$\begin{aligned} Y^+ &\equiv \lim_{P_{it}\to P_c^+} E\left[Y_{it} \mid P_{it} = P_c, t \geq 2014\right] \qquad & \tilde{Y}^+ &\equiv \lim_{P_{it}\to P_c^+} E\left[Y_{it} \mid P_{it} = P_c, t < 2014\right] \\ Y^- &\equiv \lim_{P_{it}\to P_c^-} E\left[Y_{it} \mid P_{it} = P_c, t \geq 2014\right] \qquad & \tilde{Y}^- &\equiv \lim_{P_{it}\to P_c^-} E\left[Y_{it} \mid P_{it} = P_c, t < 2014\right] \end{aligned} \tag{3.1}$$

Table 3.3 summarizes, by population and year, which average potential outcomes are actually observed in the neighborhood of P_c.

Table 3.3: Observed average potential outcomes, around threshold P_c

	t<2014	t$\geq$ 2014
$P_{it} \geq P_c$	$\tilde{Y}^+$= E$[\tilde{Y}_i(1,0,0)]$	$Y^+ = E[Y_i(1,0,1)]$
$P_{it} < P_c$	$\tilde{Y}^-$= E$[\tilde{Y}_i(0,0,0)]$	$Y^- = E[Y_i(0,1,0)]$

Notes: Potential outcomes follow the notation $Y_{it}(W_{it}, M_{it}, G_{it})$, where $W_{it} = 1$ denotes confounding policies above P_c before 2014; $M_{it} = 1$ denotes the third mandate policy introduced in 2014 below P_c; $G_{it} = 1$ denotes the presence of gender quotas, introduced in 2014 above P_c. Tildes denote potential outcomes before 2014, and (+,-) superscripts limit values. For instance, before 2014, municipalities above population cutoff P_c have $W_{it} = 1$, $M_{it} = 0$, $G_{it} = 0$.

Table 3.3 highlights that restricting to the year 2014 and estimating a simple RD identifies the effect of all three treatments since they all change discontinuously at P_c. However, conditioning on $t < 2014$, a simple RD identifies the average effect of passing from $W_{it} = 0$ to $W_{it} = 1$. In our empirical strategy, diff-in-disc takes the difference between the 2014 and the 2009 RD estimates, thus identifying the compound effect of M_{it} and G_{it}, the two policies introduced in 2014, under two assumptions introduced by Grembi, Nannicini, and Troiano (2016) which we report below for our setting. Then, in order to identify the effect of G alone, we will introduce a third ignorability assumption.

Assumption 1 (Continuity) *For all (both) periods, all average potential outcomes* $[\tilde{Y}_i(W_{it}, M_{it}, G_{it}) \mid P_{it} = P_c]$ *and* $[Y_i(W_{it}, M_{it}, G_{it}) \mid P_{it} = P_c]$*, with* $W_{it}, M_{it}, G_{it} = \{0, 1\}$ *are continuous in* P_{it} *at* P_c*. Namely,* $\forall\, W_{it}, M_{it}, G_{it}$*:*

$$\begin{aligned} &Y^+(W_{it}, M_{it}, G_{it}) = Y^-(W_{it}, M_{it}, G_{it}) = Y(W_{it}, M_{it}, G_{it}) \\ &\tilde{Y}^+(W_{it}, M_{it}, G_{it}) = \tilde{Y}^-(W_{it}, M_{it}, G_{it}) = \tilde{Y}(W_{it}, M_{it}, G_{it}) \end{aligned} \tag{A1}$$

Assumption 2 (Local parallel trends) *The effect of the confounding policy* W_{it} *at* P_c*, in the case of no compound treatment (when both* $M_{it}, G_{it} = 0$ *at the same time), is constant over time:*

$$\tilde{Y}_i(1, 0, 0) - \tilde{Y}_i(0, 0, 0) = Y_i(1, 0, 0) - Y_i(0, 0, 0) \tag{A2}$$

While assumption A1 is the usual RD continuity assumption, now extended to all potential outcomes, assumption A2 ensures the consistency over time of the treatment that will be subtracted by the diff-in-disc estimator, in our case W_{it}. Although formally untestable, we provide reassuring evidence of the validity of A2 in our setting in Section 3.5.3. Assumptions A1 and A2 alone allow to identify the compound effect of the policy changes introduced by the 2014 reform, by taking the difference between an RD around P_c estimated in 2014 and an RD around P_c estimated before 2014 (in 2009):

$$\begin{aligned} \hat{\tau}_{DD} \equiv& \left(Y^+ - Y^-\right) - \left(\tilde{Y}^+ - \tilde{Y}^-\right) \\ =& \left[Y^+(W_{it} = 1, M_{it} = 0, G_{it} = 1) - Y^-(W_{it} = 0, M_{it} = 1, G_{it} = 0)\right] \\ &- \left[\tilde{Y}^+(W_{it} = 1, M_{it} = 0, G_{it} = 0) - \tilde{Y}^-(W_{it} = 0, M_{it} = 0, G_{it} = 0)\right] \\ \overset{(A1-A2)}{=}& \left[Y(1, 0, 1) - Y(0, 1, 0)\right] - \left[\tilde{Y}(1, 0, 0) - \tilde{Y}(0, 0, 0)\right] \\ \overset{(A2)}{=}& \left[Y(W_{it}, 0, 1) - Y(W_{it}, 1, 0)\right] \end{aligned}$$

Clearly, the diff-in-disc estimator is not able to disentangle the effects of M_{it} and G_{it}. We therefore add a third identifying assumption to exclude the confounding effect of the third mandate policy, M_{it}. Section 3.A of the Appendix further elaborates on this point, and in Section 3.5.3 we provide evidence to support this assumption.

Define $M_{i,t-1} = \{1, 2\}$ as the running mandate of the mayor before the 2014 election (in practice, measured at the moment the mayor in charge until 2014 was elected, i.e. in 2009).

Assumption 3 (Irrelevance of Compound Treatments) *The third mandate policy M_{it} has no effect on social services expenditures and, therefore, the effect of gender quotas G_{it} on the outcome variables is the same regardless of M_{it}. Namely, $\forall\, W_{it}$:*

$$\begin{aligned}&Y(W_{it},1,1)-Y(W_{it},0,0)=Y(W_{it},0,1)-Y(W_{it},1,0)=\\&Y(W_{it},1,1)-Y(W_{it},1,0)=Y(W_{it},0,1)-Y(W_{it},0,0)\end{aligned} \tag{A3}$$

The intuition lies in the fact that, should M_{it} affect the outcome variables, whether directly or as the result of an interaction with gender quotas G_{it}, the regression discontinuity would estimate the compound effects of the two policies. This would lead to either negative or positive bias depending on the relative sign of the two effects. For example, if the gender quotas and the possibility of a third mandate both lead to higher social services expenditures, the regression discontinuity estimate in 2014 would be downward biased, since the two policies apply on the opposite sides of the population threshold. This ignorability assumption, arguably a strong one, rules out this possibility. Section 3.5.3 discusses it in more detail and provides an informal test of its validity.

3.4.3 Estimation

We restrict the sample to municipalities with an election in the interval $P_{it} \in [P_c - h, P_c + h]$, where h is the mean squared error optimal bandwidth (MSERD) described in Calonico, Cattaneo, and Titiunik (2014a) and Calonico, Cattaneo, and Titiunik (2014b). Following Grembi, Nannicini, and Troiano (2016), estimation is obtained by running the following regression model:

$$Y_{it} =\delta_0 + \delta_1 P_i^* + S_i(\gamma_0 + \gamma_1 P_i^*) + T_t[\alpha_0 + \alpha_1 P_i^* + S_i(\beta_0 + \beta_1 P_i^*)] + \epsilon_{it} \tag{3.2}$$

where S_i is a dummy for cities above 3,000 capturing treatment status, T_t an indicator for the post-treatment period, and $P_i^* = P_i - P_c$ the deviation from the population threshold. In this framework, β_0 is the parameter of interest, corresponding to $\hat{\tau}_{DD}$. It measures the post-reform discontinuity in Y_{it} for municipalities across the population threshold, thus identifying the effect of gender quotas under A1-A3. Y_{it} includes both current and capital expenditures per capita for a number of social services (we estimate the regression for each of them): (i) Social services, overall; (ii) Childcare; (iii) Services for the disabled; (iv) Elderly care; (v) Other forms of assistance. To show that the policy was effective, we also estimate the models with two indicators of female representation in municipal executive committees. The first indicator is the share of seats allocated to women (by the new policy, it cannot be less than 40%). The second indicator is the absolute number of women in the committees, which we include since the number of maximum seats in the municipal executive committee changed several times between 2011 and 2014, possibly introducing measurement error in the female executive share.[9] We will show that the observed change in the executive female share is not driven only by changes in the denominator, the total number of seats (for example, see Figure 3.3). Section 3.B.1 further elaborates on the legislation regarding council and executive size and reports a table summarizing the regulatory changes for different population thresholds.

3.5 Results

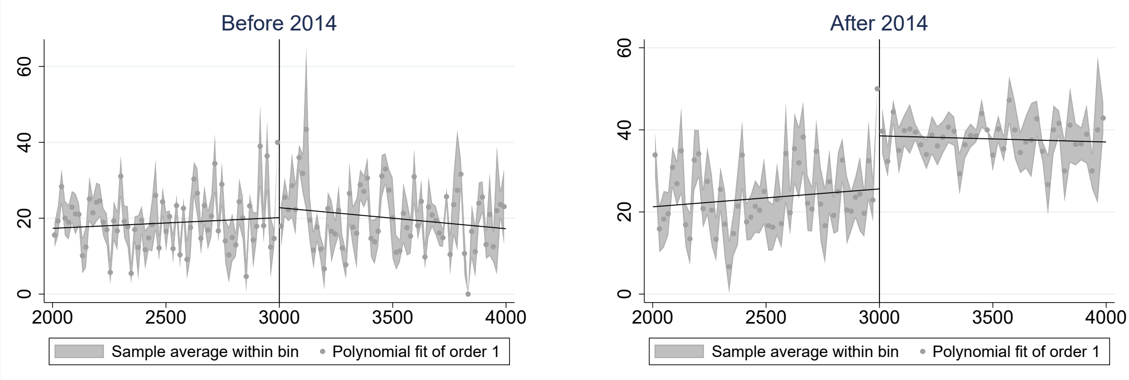
3.5.1 Share and Number of Women in Municipal Executive

We first evaluate the effectiveness of the 2014 reform in increasing gender representation within municipal executive committees, i.e. the degree to which it was enforced.

Figure 3.3 provides some graphical evidence, showing the share of women in municipal executive around the population threshold before and after 2014. Before 2014, there is no discontinuity around the population threshold for the share of women in municipal executive, while after 2014 there is: the reform concerns municipalities on the right of the population threshold and indeed we can see a jump at the threshold, meaning that the policy boosted the presence of women in municipalities with more than 3,000 inhabitants. To address the concern that changes in maximum executive size might be driving this result, we also investigate changes in the absolute number of women in the executive council. Figure 3.4 confirms that the observed change in the executive female share is not driven only by changes in the denominator, but by a clearly observable increase in the number of women.

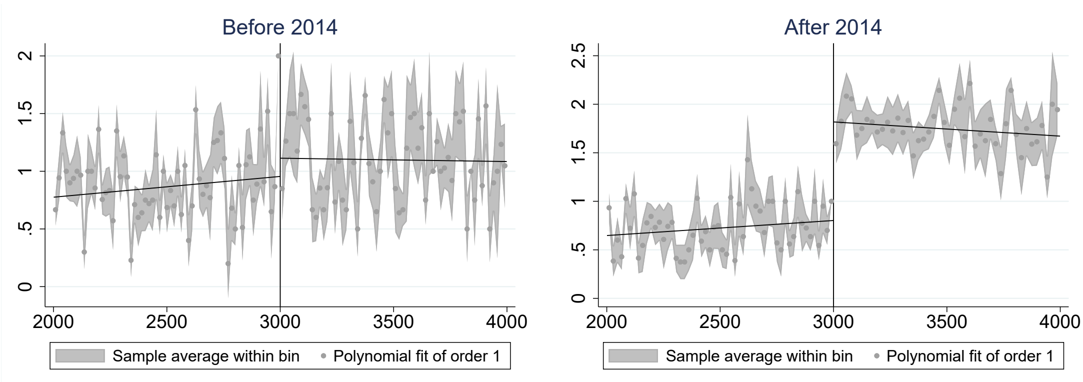
9. Changes in the size of the municipal council might also be relevant for social welfare expenditures, our outcome variables. However, previous literature has not reached an agreement regarding the relationship between legislature size and government spending at the local level, finding both positive and negative effects Alptekin et al. (2020).

Figure 3.3: Share of Women in Municipal Executive Councils: Before and After



Notes: Figure 3.3 shows the share of women in municipal executive councils around the population threshold before and after 2014. The figures are produced using a polynomial of order one, epanechnikov kernel and bin selection that specifies the mimicking-variance evenly spaced method using spacings estimators.

Figure 3.4: Number of Women in Municipal Executive: Before and After



Notes: Figure 3.4 shows the number of women in municipal executive councils around the population threshold before and after 2014. The figure is produced using a polynomial of order one, epanechnikov kernel and bin selection that specifies the mimicking-variance evenly spaced method using spacings estimators.

Table 3.4 reports estimates of the effect of the 2014 Law on the share and the total number of women in the municipal executive. The first three columns look at the share of women when applying respectively a first-, second-, or third-order polynomial. With the first-order polynomial, the reform increases the share of women by 7 percentage points (p.p.), statistically significant at the 1% significance level. With a second or third-order polynomial, the share of women increases by almost 9 p.p. with a 1% significance level. Hence, with all three orders of polynomial, the share of women increases significantly with a remarkable order of magnitude, i.e. an increase of between 30% and 35%. Notice that when one uses a different polynomial degree the optimal bandwidth, and subsequently the number of observations considered, change. Our preferred specification uses a quadratic local polynomial, in line with the suggestions of Gelman and Imbens (2019) against the use of high-order polynomials in RD designs. While confidence intervals may appear quite large and thus inconclusive of what the true effect of the policy is, note that in this context we use observations from the entire population of Italian municipalities around the threshold of 3000 inhabitants. Therefore, the information provided by confidence intervals is unclear, as they do not reflect sampling uncertainty over repeated experiments. Moreover, even the largest interval bound does not provide an economically significant variation in per capita expenditure. For instance, the education ministry estimates a yearly expense of 6027.55 EUR for each child enrolled in pre-school, and 6288.68 for elementary school (MIUR 2021). The last three columns look at the number of women in the municipal executive. With a first-order polynomial, the reform increases the number of women by 0.7 at 1% significance level, hence it increases by almost a woman. The higher-order polynomials of degrees 2 and 3 confirm the result. Therefore, both the share and the number of women increase with the reform, and this result does not depend on the order of polynomial used in the analysis.

Table 3.4: Share and number of female executive councillors: RD estimates by polynomial degree

	Share of females			N. females		
	(1)	(2)	(3)	(4)	(5)	(6)
	p=1	p=2	p=3	p=1	p=2	p=3
$S_{i,t} * T_{i,t}$	7.36***	8.83***	8.33***	0.70***	0.78***	0.76***
	(1.95)	(2.40)	(2.05)	(0.08)	(0.10)	(0.10)
Observations	6160	4119	5536	6160	4127	4287
CCT pop bandwidth	735	502	665	737	503	525
Mean Y	24.6	25.6	24.7	1.1	1.1	1.1
Mean N.assessori	4.4	4.4	4.4	4.4	4.4	4.4

Standard errors in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$
The population bandwidth for local regression is selected via MSE optimal criterion (MSERD). Columns (1), (2) and (3) show the 2014 RD estimates for the share of female executive councillors, with respectively polynomials of order one, two and three. Columns (4), (5) and (6) show the 2014 RD estimates for the number of female executive councillors, with respectively polynomials of order one, two and three.

3.5.2 Effect on Social Spending

In this section, we investigate the effect of the increase in the female share in local executives on municipal social services expenditure. In our main analysis, we focus on accrual account expenses, as these are more closely linked with the current administration; Section 3.A.2 in the Appendix reports the results for total accrued expenses.

Table 3.5 shows the results for social current spending in the accrual account. Column (1) reports the effect of the reform on overall per capita spending in social services, while Columns (2-5) consider four subgroups for social services spending, namely elderly care; services for persons with disabilities; childcare services; and other assistance such as services for families or assistance to marginalized groups. Notably, the effect of increased gender representation in local executives on social spending and relevant subgroups is not statistically different from zero for any of these outcomes. Moreover, all point estimates are quite small, meaning that they have limited economic significance: in none of the categories considered spending increases or decreases by more than 6 euros per resident.

Table 3.6 shows the results for social capital spending in the accrual account. Again, total spending on social services does not significantly change following the increase in female representation. In particular, all point estimates are close to zero and not statistically significant, and even the confidence intervals bound the estimates at no more than a 2 euro increase (or decrease) per resident.

Table 3.5: Current expenses, accrual account per year, per capita

	(1) Social	(2) Child	(3) Elderly	(4) Disabled	(5) Other
$S_{i,t} * T_{i,t}$	-5.9	-0.7	-1.2	1.4	-4.0
	(7.5)	(2.0)	(7.3)	(1.2)	(3.5)
Observations	2666	3898	2077	3190	4427
CCT pop bandwidth	348	505	275	408	575
$\bar{Y}_{pc}$	52.0	7.2	6.6	1.7	31.6

Standard errors in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$
All categories of expenditures are measured per year, per capita and in euros, and refer to current expenditures. The population bandwidth for local regression is selected via MSE optimal criterion (MSERD). Column (1) shows the 2014 RD estimates for social expenditures, Column (2) shows the 2014 RD estimates for childcare expenditures, Column (3) shows the 2014 RD estimates for elderly care expenditures, Column (4) shows the 2014 RD estimates for disabled care expenditures, and Column (5) shows the 2014 RD estimates for other expenditures.

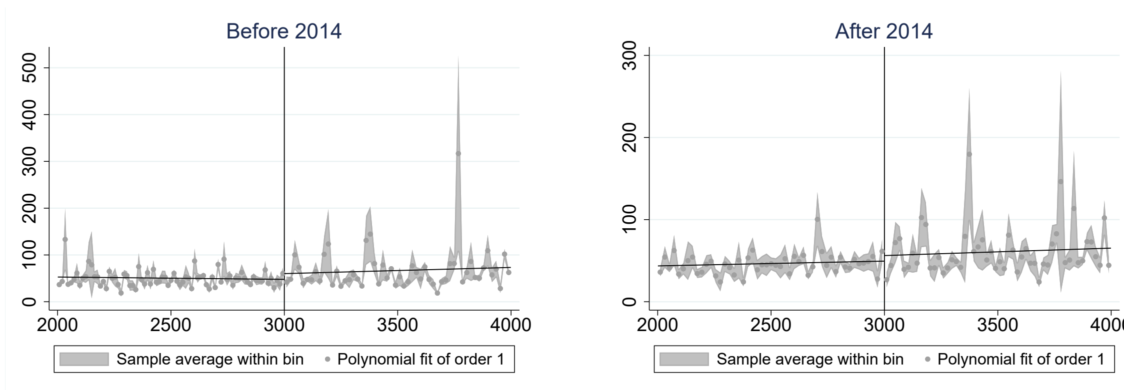
Table 3.6: **Capital expenses, accrual account per year, per capita**

	(1) Social	(2) Child	(3) Elderly	(4) Disabled	(5) Other
$S_{i,t} * T_{i,t}$	-0.6	-0.3	-0.7	-0.3	-0.7
	(1.5)	(0.3)	(1.1)	(0.2)	(0.8)
Observations	4924	3508	3538	7897	3808
CCT pop bandwidth	629	448	453	966	492
$\bar{Y}_{pc}$	3.5	0.2	0.7	0.1	0.6

Standard errors in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$
All categories of expenditures are measured per year, per capita and in euros, and refer to capital expenditures. Column (1) shows the 2014 RD estimates for social expenditures, Column (2) shows the 2014 RD estimates for childcare expenditures, Column (3) shows the 2014 RD estimates for elderly care expenditures, Column (4) shows the 2014 RD estimates for disabled care expenditures, and Column (5) shows the 2014 RD estimates for other expenditures.

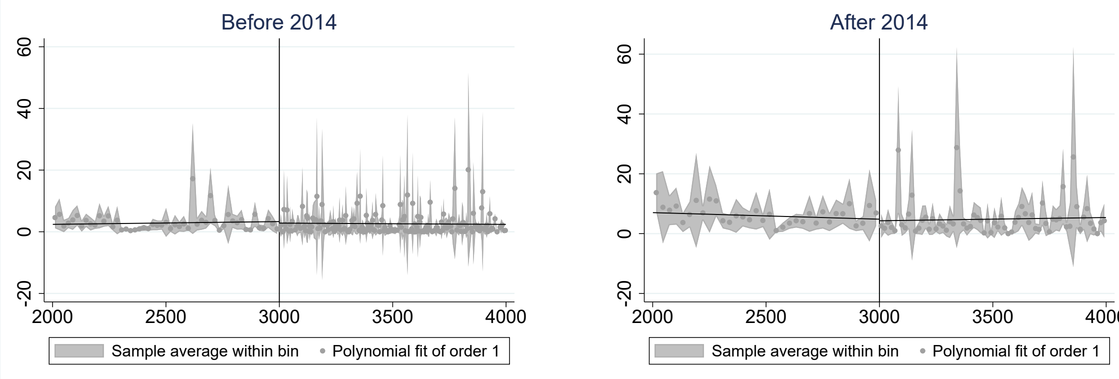
Figure 3.5 and Figure 3.6 provide some graphical evidence, showing respectively the current and capital expenses in the accrual account at the municipal level around the population threshold before and after 2014. For both types of expenses, there appears to be no substantial discontinuity around the population threshold, either before or after 2014.

Figure 3.5: Current Expenses per capita, accrual account: Before and After



Notes: Figure 3.5 shows the current expenditures in the accrual account around the population threshold before and after 2014. Expenditures are per year, per capita and in euros. The figures are produced using a polynomial of order one, epanechnikov kernel and bin selection that specifies the mimicking-variance evenly spaced method using spacings estimators.

Figure 3.6: Capital Expenses per capita, accrual account: Before and After

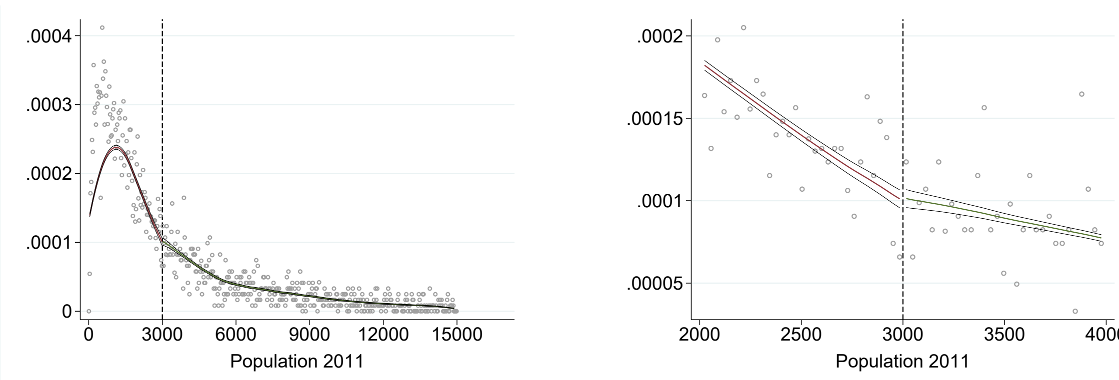


Notes: Figure 3.6 displays the capital expenses in the accrual account around the population threshold before and after 2014. Expenditures are per year, per capita and in euros. The figures are produced using a polynomial of order one, epanechnikov kernel and bin selection that specifies the mimicking-variance evenly spaced method using spacings estimators.

3.5.3 Internal Validity

Assumption A1 entails that there should be no manipulation of the running variable. If mayors could manipulate population size and sort below the threshold to avoid the gender quota rule, our estimates would suffer from selection bias. To test this, we implement the McCrary (2008) test for sorting of the population variable. Figure 3.7 shows that we find no significant discontinuity in the relative frequency of population size at the relevant threshold; moreover, the density test by Cattaneo, Jansson, and Ma (2018) suggests there is no jump at the threshold (p-value 0.158).

Figure 3.7: Manipulation test



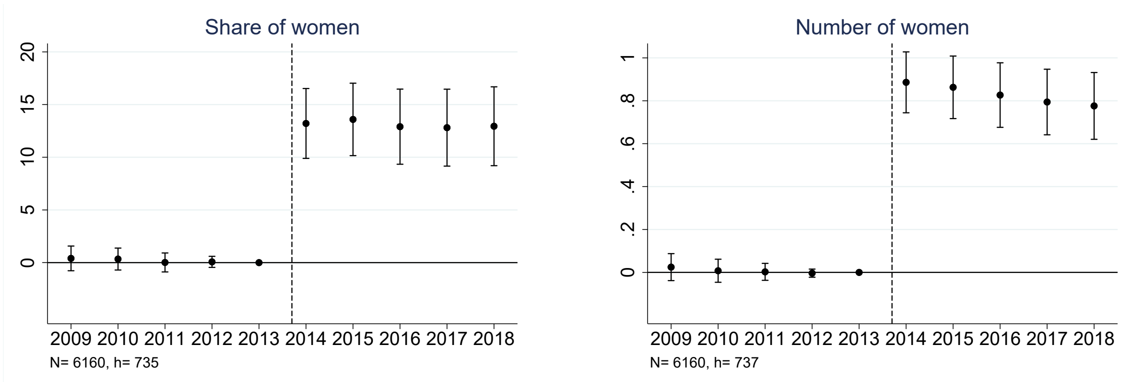
Notes: Table 3.7 shows the estimated density of population, using data from the Census 2011, for municipalities below 15000 residents.

As discussed by Eggers et al. (2018), the difference-in-discontinuities design crucially rests on the assumption of local parallel trends between treatment and control units of observation in the neighborhood of the discontinuity. We test this assumption by implementing a local event study, using the same bandwidth selection criteria employed for the results. Following Marattin, Nannicini, and Porcelli (2021), we run the following regression using municipalities with $P_{i,t} \in [P^*_{i,t} - h; P^*_{i,t} + h]$:

$$Y_{it} = \theta_0 + S_i(\theta_1 + \sum_{t=2009}^{2018} \theta_{2t}\eta_t) + \sum_{t=2009}^{2018} \theta_{3t}\eta_t + \lambda_i + u_{it} \tag{3.3}$$

where η_t and λ_i represent year and municipality dummies, respectively. This evaluates how the difference between municipalities across the population threshold has evolved over time, taking pre-treatment year 2013 as baseline. Figure 3.8 and Figure 3.9 plot the corresponding point estimates for female representation within executive committees and per capita expenditure in social services, respectively.

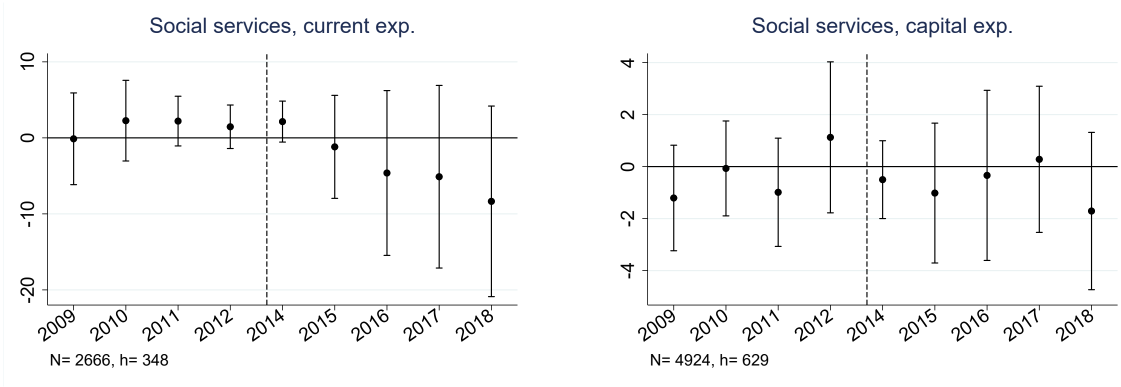
Figure 3.8: Local parallel trends, gender representation



Notes: Figure 3.8 displays gender representation over time, in terms of share and number of women in the executive councils.

The assumption of local parallel trends is satisfied in terms of both the share and number of women in the executive committee, where we observe a clear jump from 2014 that remains relatively stable over time. Regarding social expenditure, the evidence does not point to a violation of this assumption, neither for current nor capital expenditure. Notice also that there is no substantial jump in social services expenditure after the introduction of the policy. While for current expenditure we can observe a downward trend in the point estimates, the large confidence intervals suggest that the effect could be substantially positive or negative. Similarly, for capital expenditure, the point estimates remain closer to zero even after the introduction of the reform, but again with large confidence intervals.

Figure 3.9: Local parallel trends, social services



Notes: Figure 3.9 displays expenditures on social services over time, in terms of current and capital expenses. Outcomes are measured in EUR per capita.

Finally, assumption A3 is an ignorability assumption, requiring that possibility of running for a third mandate in municipalities below 3000 inhabitants from 2014 does not have a direct effect on social welfare spending. We propose two informal tests. The first relies on the mandate of the mayor in charge before the 2014 elections, denoted as $M_{i,t-1}$. A potential violation of A3 is that the possibility of a third mandate could create incentives to increase social services expenditures in municipalities below the 3000 inhabitants threshold, to attract consensus and be re-elected. The policy was introduced on April 14th, 2014, and elections were held shortly after, on May 25th. Therefore, we would expect only mayors in their first mandate to be affected, since some were campaigning to be re-elected, possibly promising and planning to increase social services expenditure upon re-election. If this was the case, then the presence of municipalities with a mayor at their first mandate prior to the 2014 elections ($M_{i,t-1}$) would imply a negative bias in our 2014 RD estimate, and consequently on the diff-in-disc estimate. This is because, if higher female representation results in a non-negative change in social services expenditure above the threshold, then the potential increase in expenditure below the threshold would make us underestimate the true effect of female representation. Note that if instead the third mandate policy pushed first-mandate mayors to decrease social services expenditures, the magnitude of our estimates, which is not distinguishable from zero, would be overstated. Assumption A3 cannot be directly tested, due to the potential outcomes involved, but we perform an informal test of its validity. The first test boils down to the following: if we restrict our sample to municipalities whose mayors were at the first mandate before the 2014 elections, do we observe a smaller discontinuity in social services expenditure after elections, relative to restricting to mayors at their second mandate? If we do not, then this suggests that the third mandate policy had no direct impact on social services expenditures. The results of this test for each dependent variable are shown in Table 3.7a and Table 3.7b: for both current and capital per capita expenditures, we estimate the RD at the population threshold in 2014 after restricting to municipalities whose mayors were at the first ($M_{i,t-1} = 1$) or second ($M_{i,t-1} = 2$) mandate before the elections. Then we take the difference between these estimates and compute their non-parametrically bootstrapped standard errors. We only find two statistically significant differences out of 10. One has a magnitude very close to zero. The other one is negative, implying the main results for this expense might be positively biased (while still indistinguishable from zero). (Figure 3.10) and (Figure 3.11) show graphical evidence for current expenditures on overall social services and childcare, respectively. Confidence intervals cross zero but they are large. Also, in this case, it should be noted that we use observations for the entire population, thus making confidence intervals no longer informative on sampling variability. As a second test, exploiting the fact that the third mandate policy remained in place for all years after 2014, we check whether year-by-year social expenditures differ at the population thresholds between mayors who are currently at their first or second mandate. In this second test, we rely on non-parametrically bootstrapped standard errors for inference. The results are summarized by Figure 3.12, where we summarize the difference in estimates between mayors at different mandates between 2014 and 2018, expressed in EUR per capita, and their 95% confidence intervals. The latter reach a difference of 30 yearly euros per capita for elderly care. To give some context on the lack of economic significance, one day in a public retirement home costs on average 112 EUR per patient (Pesaresi 2019). These results are consistent with findings from Broberg and Panizza (2021), who show that the third mandate policy does not impact overall social expenditure.

Table 3.7: Testing A3 - regression estimates

(a) Current expenditures (EUR per capita)

Type of service	RD estimate with $M_{i,t-1} = 1$	RD estimate with $M_{i,t-1} = 2$	Difference of estimates $(M_{i,t-1} = 1) - (M_{i,t-1} = 2)$
Social	11.897	10.43	1.467
	(13.723)	(9.714)	(15.389)
Child	-1.273	-3.399	2.126
	(3.872)	(3.422)	(6.507)
Elderly	12.049	7.28	4.769
	(10.697)	(5.09)	(9.241)
Disabled	0.897	0.244	0.653
	(1.216)	(0.314)	(0.94)
Other	-3.07	8.767	-11.837
	(7.775)	(6.936)	(12.614)

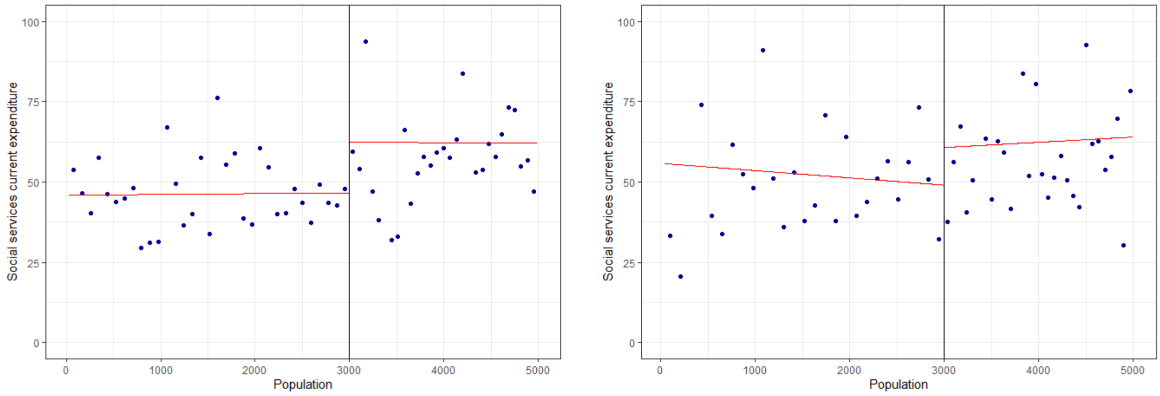
(b) Capital expenditures (EUR per capita)

Type of service	RD estimate with $M_{i,t-1} = 1$	RD estimate with $M_{i,t-1} = 2$	Difference of estimates $(M_{i,t-1} = 1) - (M_{i,t-1} = 2)$
Social	-5.989**	1.103	-7.092**
	(2.516)	(4.635)	(3.361)
Child	0.475	0.091	0.384
	(0.443)	(0.053)	(0.159)
Elderly	-0.62	3.867	-4.487
	(0.862)	(3.251)	(3.378)
Disabled	0.431	0.123	0.308***
	(0.503)	(0.133)	(0.068)
Other	-2.807	-0.246	-2.561
	(1.762)	(0.232)	(1.999)

Notes: Standard errors in parenthesis, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. The first column shows the 2014 RD estimates when restricting to municipalities whose mayor was at the first mandate before elections ($M_{i,t-1} = 1$). In the second column, the mayor was at the second mandate ($M_{i,t-1} = 2$). The last column shows the difference in estimates and its standard error, obtained by non-parametric bootstrapping. Panel a) considers current expenditure while Panel b) refers to capital expenditure, in both cases considering only the accrual account. If the third mandate policy had a direct confounding effect on expenditures, due to different incentives for re-election of first and second-mandate mayors, we would expect the two RD estimates to be different (in particular, the one for fist-mandate mayors to be larger).

Figure 3.10: Testing A3: social services current expenditure per capita (EUR)

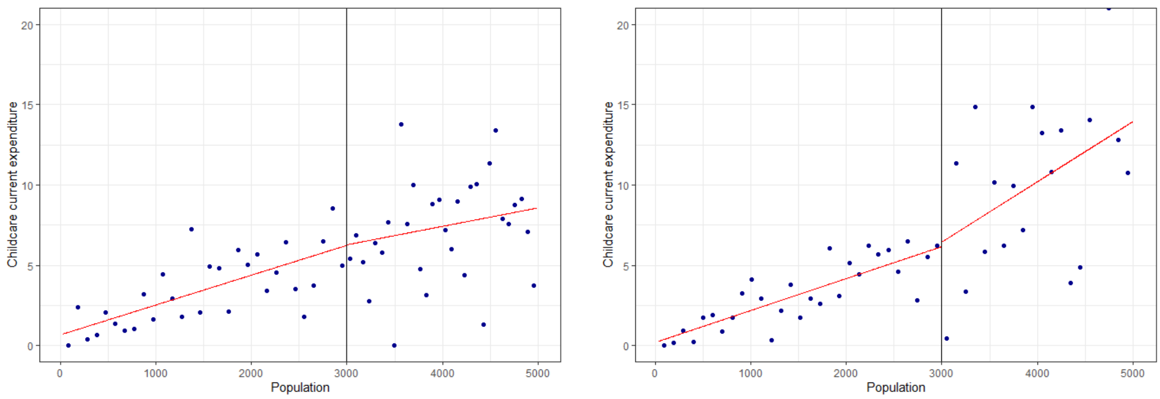
(a) Mayors at first mandate **(b) Mayors at second mandate**



Notes: the figures show how social services expenditure changed at the cutoff after the 2014 elections. In Panel (A), we restrict to municipalities whose mayor was at the first mandate before elections. In Panel (B), the mayor was at the second mandate. If the third mandate policy had a direct confounding effect on social services expenditures, we would expect the two graphs to show different RD estimates, however, they are not statistically different.

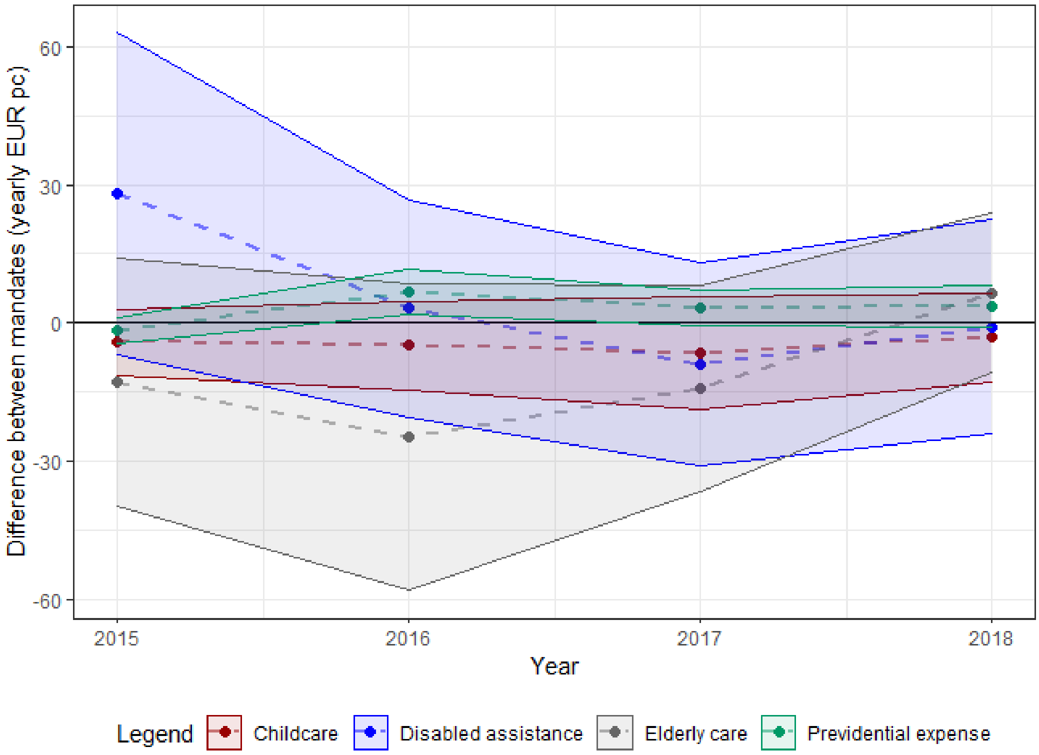
Figure 3.11: Testing A3: childcare current expenditure per capita (EUR)

(a) Mayors at first mandate **(b) Mayors at second mandate**



Notes: the figures shows how childcare expenditure change at the cutoff after 2014 elections. In Panel (A), we restrict to municipalities whose mayor was at the first mandate before elections. In Panel (B), the mayor was at the second mandate. If the third mandate policy had a direct confounding effect on childcare expenditures, we would expect the two graphs to show different RD estimates, however they are not statistically different.

Figure 3.12: Testing A3: yearly per capita current expenses (EUR), difference across mayor mandates after 2014



Notes: the figures shows the difference in expenditures between mayors at their first or second mandate, for all years after 2014, and their 95% confidence intervals. These are defined as current yearly expenditures per capita.

3.6 Robustness

In this section, we adjust and re-estimate the regression of Equation 3.2 to include for municipality fixed effects. This allows to minimize the effect of outliers in the estimation. Hence we consider the following:

$$Y_{it} = \delta_0 + \gamma_i + T_t[\alpha_0 + \alpha_1 P_i^* + S_i(\beta_0 + \beta_1 P_i^*)] + \epsilon_{it} \tag{3.4}$$

where as before S_i is a dummy for cities above 3,000 capturing treatment status, T_t an indicator for the post-treatment period, and $P_i^* = P_i - P_c$ the deviation from the population threshold, and we introduce municipality fixed effects, γ_i. The specification is adjusted accounting for the fact that S_i and P_i^* are by definition constant within municipality, since they are based on the population measured by the 2011 Census.

Table 3.8 and Table 3.9 display the results for current and capital expenses in the accrual account, respectively. All point estimates are quite similar to the main regression specification presented in Section 3.5. While standard errors are comparable

Table 3.8: Current expenses per capita, accrual account - municipality fixed effects

	(1) Social	(2) Child	(3) Elderly	(4) Disabled	(5) Other
$S_{i,t} * T_{i,t}$	-6.1	-0.9	-1.4	1.5	-3.8
	(4.0)	(1.0)	(3.5)	(1.1)	(2.1)
Observations	2666	3898	2077	3190	4427
CCT pop bandwidth	348	505	275	408	575
$\bar{Y}_{pc}$	52.0	7.2	6.6	1.7	31.6

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Standard errors in parenthesis, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. All categories of expenditures are measured per year, per capita and in euros, and refer to current expenditures. The population bandwidth for local regression is selected via MSE optimal criterion (MSERD). Column (1) shows the 2014 RD estimates for social expenditures, Column (2) shows the 2014 RD estimates for childcare expenditures, Column (3) shows the 2014 RD estimates for elderly care expenditures, Column (4) shows the 2014 RD estimates for disabled care expenditures, and Column (5) shows the 2014 RD estimates for other expenditures.

Table 3.9: Capital expenses per capita, accrual account - municipality fixed effects

	(1) Social	(2) Child	(3) Elderly	(4) Disabled	(5) Other
$S_{i,t} * T_{i,t}$	-0.7	-0.3	-0.9	-0.3	-0.7
	(1.5)	(0.3)	(1.0)	(0.2)	(0.8)
Observations	4924	3508	3538	7897	3808
CCT pop bandwidth	629	448	453	966	492
$\bar{Y}_{pc}$	3.5	0.2	0.7	0.1	0.6

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Standard errors in parenthesis, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. All categories of expenditures are measured per year, per capita and in euros, and refer to capital expenditures. The population bandwidth for local regression is selected via MSE optimal criterion (MSERD). Column (1) shows the 2014 RD estimates for social expenditures, Column (2) shows the 2014 RD estimates for childcare expenditures, Column (3) shows the 2014 RD estimates for elderly care expenditures, Column (4) shows the 2014 RD estimates for disabled care expenditures, and Column (5) shows the 2014 RD estimates for other expenditures.

when considering capital expenditures, they are significantly lower for every category of current expenses. This is in line that focusing on within-municipality variation yields more precise estimates by reducing the leverage of outliers.

Further robustness check, in future work, can check the heterogeneity of results with respect to the mayor's gender and term. We also plan to study the potential spillover effects across neighboring municipalities, which might provide social services jointly, and how the effect dynamics evolve over time.

3.7 Conclusion

Does female representation in executive government bodies increase the provision of social services? In high-income countries, does it have an effect on elderly care on top of childcare services? In this paper, we focus on the case of Italy. We use data from the Local Administrators Registry, the Population Census, and a database of municipal finances collected by the Ministry of Interior, and we exploit a 2014 reform which mandated that no gender represent less than 40% of the municipal executive committee (mayor included) in municipalities with more than 3,000 residents. The focus of the policy on executive municipal committees is interesting, since other examples of gender quotas from high-income countries typically refer to candidates' lists, resulting in the need to rely on close elections to identify the effect of single women in positions of political power.
However, the fact that the same population threshold was used in Italy for the introduction of other policies, at different points in time, poses the issue of confounding treatments. We address this by relying on Grembi, Nannicini, and Troiano (2016)'s difference-in-discontinuities estimator, combined with an extra identifying assumption that suits our specific setting, for which we provide informal testing.

Our analysis proceeds in two steps. First, we establish that in Italy, the introduction of gender quotas in municipal executive committees was indeed effective in increasing female political representation. Our results are robust to the choice of the polynomial for the estimation: both the share and the number of women in executive committees increase significantly with first, second, and third-order polynomial.

Once shown that the policy indeed affected the gender composition of executive committees, we turn to investigate the expenditures on local social services. We look at per capita and capital expenditures in overall social services, childcare, elderly care, and assistance for the disabled and miscellaneous categories. We do not find significant changes for any of them: this is a partial reassurance that the 2014 reform did not have unintended consequences, at least for local social services. These results are in line with Rigon and Tanzi (2012), who find no correlation between female representation in city councils and municipal expenses.

Despite the lack of an effect on social services expenditures, however, we cannot exclude other externalities of increased female representation, for instance on female labor force participation, attitudes towards women, women's empowerment, and gender norms, leaving scope for further research. Moreover, by relying on a policy implemented at the population cutoff of 3000 inhabitants, our analysis produces local estimates. While still relevant, the decisional dynamics of local committees can arguably change between small and large municipalities: future research could investigate heterogeneous effects by population size by employing different empirical strategies.

References

Accettura, C., and P. Profeta. 2021. “Gender Differences in Political Budget Cycles.”

Acconcia, A., C. Ronza, et al. 2021. *Women’s Representation in Politics and Government Stability.* Technical report. Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.

Ahern, K. R., and A. K. Dittmar. 2012. “The changing of the boards: The impact on firm valuation of mandated female board representation.” *The quarterly journal of economics* 127 (1): 137–197.

Alesina, A., and E. La Ferrara. 2005. “Preferences for redistribution in the land of opportunities.” *Journal of public Economics* 89 (5-6): 897–931.

Alptekin, H., D. Freire, U. Mignozzetti, and C. Roman. 2020. “The Effect of Legislature Size on Public Spending: A Meta-Analysis.”

ANCI. 2015. *La rappresentanza di genere nelle amministrazioni comunali.* http://www.anci.it/wp-content/uploads//2018/06/Contenuti/Allegati/Dossier%20Rappresentanza%20di%20genere%20nelle%20amministrazioni%20comunali%20italiane%202015.pdf. [Online; accessed 19-February-2022].

Andreoli, F., E. Manzoni, and M. Margotti. 2021. “Women at work: Gender quotas, municipal elections and local spending.” *European Journal of Political Economy,* 102175.

Bagues, M., and P. Campa. 2021. “Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design.” *Journal of Public Economics* 194:104315.

Baltrunaite, A., A. Casarico, P. Profeta, and G. Savio. 2019. “Let the voters choose women.” *Journal of Public Economics* 180:104085.

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova. 2009. “Powerful women: does exposure reduce bias?” *The Quarterly journal of economics* 124 (4): 1497–1540.

Bhalotra, S., and I. Clots-Figueras. 2014. “Health and the political agency of women.” *American Economic Journal: Economic Policy* 6 (2): 164–97.

Bhalotra, S., I. Clots-Figueras, and L. Iyer. 2018. “Pathbreakers? Women’s electoral success and future political participation.” *The Economic Journal* 128 (613): 1844–1878.

Bratti, M., E. D. Bono, and D. Vuri. 2005. “New mothers’ labour force participation in Italy: The role of job characteristics.” *Labour* 19:79–121.

Broberg, N., and P. Panizza. 2021. “Term limits and accountability, Evidence from Italy.” *mimeo.*

Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014a. “Robust data-driven inference in the regression-discontinuity design.” *The Stata Journal* 14 (4): 909–946.

———. 2014b. “Robust nonparametric confidence intervals for regression-discontinuity designs.” *Econometrica* 82 (6): 2295–2326.

Cattaneo, M. D., M. Jansson, and X. Ma. 2018. “Manipulation testing based on density discontinuity.” *The Stata Journal* 18 (1): 234–261.

Chattopadhyay, R., and E. Duflo. 2004. “Women as policy makers: Evidence from a randomized policy experiment in India.” *Econometrica* 72 (5): 1409–1443.

Clots-Figueras, I. 2011. “Women in politics: Evidence from the Indian States.” *Journal of public Economics* 95 (7-8): 664–690.

———. 2012. “Are female leaders good for education? Evidence from India.” *American Economic Journal: Applied Economics* 4 (1): 212–44.

De Benedetto, M. A., and M. De Paola. 2019. “Term limit extension and electoral participation. Evidence from a diff-in-discontinuities design at the local level in Italy.” *European Journal of Political Economy* 59:196–211.

Del Boca, D. 2002. “The effect of child care and part time opportunities on participation and fertility decisions in Italy.” *Journal of population economics* 15 (3): 549–573.

Del Boca, D., M. Locatelli, and D. Vuri. 2005. “Child-care choices by working mothers: The case of Italy.” *Review of Economics of the Household* 3 (4): 453–477.

Del Boca, D., N. Oggero, P. Profeta, and M. C. Rossi. 2021. “Household division of labor during two waves of COVID-19 in Italy.” *Covid Economics* 60.

Del Boca, D., and R. M. Sauer. 2009. “Life cycle employment and fertility across institutional environments.” *European Economic Review* 53 (3): 274–292.

Del Boca, D., and D. Vuri. 2007. “The mismatch between employment and child care in Italy: the impact of rationing.” *Journal of Population Economics* 20 (4): 805–832.

Eckbo, B. E., K. Nygaard, and K. S. Thorburn. 2016. “How costly is forced gender-balancing of corporate boards?” *European Corporate Governance Institute (ECGI)-Finance Working Paper* 463.

Eggers, A. C., R. Freier, V. Grembi, and T. Nannicini. 2018. “Regression discontinuity designs based on population thresholds: Pitfalls and solutions.” *American Journal of Political Science* 62 (1): 210–229.

Ferrari, G., V. Ferraro, P. Profeta, and C. Pronzato. 2021. “Do Board Gender Quotas Matter? Selection, Performance, and Stock Market Effects.” *Management Science.*

Ferreira, F., and J. Gyourko. 2014. “Does gender matter for political leadership? The case of US mayors.” *Journal of Public Economics* 112:24–39.

Funk, P., and C. Gathmann. 2015. “Gender gaps in policy making: Evidence from direct democracy in Switzerland.” *Economic Policy* 30 (81): 141–181.

Gagliarducci, S., and M. D. Paserman. 2012. “Gender interactions within hierarchies: evidence from the political arena.” *The Review of Economic Studies* 79 (3): 1021–1052.

Gago, A., and F. Carozzi. 2021. “Do Female Leaders Promote Gender-Sensitive Policies?” *Available at SSRN 3716566.*

Gelman, A., and G. Imbens. 2019. “Why high-order polynomials should not be used in regression discontinuity designs.” *Journal of Business & Economic Statistics* 37 (3): 447–456.

Giorgetti, I., and M. Picchio. 2021. “One billion euro programme for early childcare services in Italy.” *Metroeconomica* 72 (3): 460–492.

Grembi, V., T. Nannicini, and U. Troiano. 2016. “Do fiscal rules matter?” *American Economic Journal: Applied Economics,* 1–30.

Hessami, Z., and T. Baskaran. 2019. “Competitively elected women as policy makers.”

Hessami, Z., and M. L. da Fonseca. 2020. “Female political representation and substantive effects on policies: A literature review.” *European Journal of Political Economy* 63:101896.

Iyer, L., A. Mani, P. Mishra, and P. Topalova. 2012. "The power of political voice: women's political representation and crime in India." *American Economic Journal: Applied Economics* 4 (4): 165–93.

Lippmann, Q. 2020. "Gender and Lawmaking in Times of Quotas." *Available at SSRN 3330617.*

Marattin, L., T. Nannicini, and F. Porcelli. 2021. "Revenue vs expenditure based fiscal consolidation: the pass-through from federal cuts to local taxes." *International Tax and Public Finance,* 1–39.

Matsa, D. A., and A. R. Miller. 2013. "A female style in corporate leadership? Evidence from quotas." *American Economic Journal: Applied Economics* 5 (3): 136–69.

Mavisakalyan, A., and Y. Tarverdi. 2019. "Gender and climate change: Do female parliamentarians make difference?" *European Journal of Political Economy* 56:151–164.

McCrary, J. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142 (2): 698–714.

MIUR. 2021. "Pubblicazione del Costo Medio Studente di cui al DM 61/2021 art. 5 comma 3." *Ministero dell'Istruzione, dell'Università e della Ricerca.*

OECD, G. Ferrant, L. M. Pesando, and K. Nowacka. 2014. "Unpaid Care Work: The missing link in the analysis of gender gaps in labour outcomes." *Boulogne Billancourt: OECD Development Center.*

Pesaresi, F. 2019. "Le tariffe delle Residenze Sanitarie Assistenziali." *Network Non Autosufficienza, rivista online* 4.

Rigon, M., and G. Tanzi. 2012. "Does gender matter for public spending? Empirical evidence from Italian municipalities." *Empirical Evidence from Italian Municipalities (April 27, 2012). Bank of Italy Temi di Discussione (Working Paper) No* 862.

Vandelannoote, D., P. Vanleenhove, A. Decoster, J. Ghysels, and G. Verbist. 2015. "Maternal employment: the impact of triple rationing in childcare." *Review of Economics of the Household* 13 (3): 685–707.

Wrohlich, K. 2004. *Child care costs and mothers' labor supply: an empirical analysis for Germany.* Technical report. DIW Discussion Papers.

APPENDIX

Appendix 3.A Empirical Strategy: Diff-in-disc in Our Setting

Subsection 3.A.1 Local Parallel Trends, Expenditure Subgroups

This section reports the local-event study estimates for the expenditure subgroups considered in the main analysis, computed following Equation 3.3 and taking pre-treatment year 2013 as baseline. There is no evidence of non-parallel trends before the introduction of the policy in 2014 in all variables considered. Note however that there is also no clear development after 2014, and the confidence intervals tend to be quite wide.

Subsection 3.A.2 Results on Total Accrued Expenses

This section reports the main estimation results considering total accrued expenses, rather than expenses in the accrual account only. The results are close to the estimates reported in Tables 3.5 and 3.6, as the width of the confidence intervals does not allow to establish whether there was a change in the expenditure patterns.

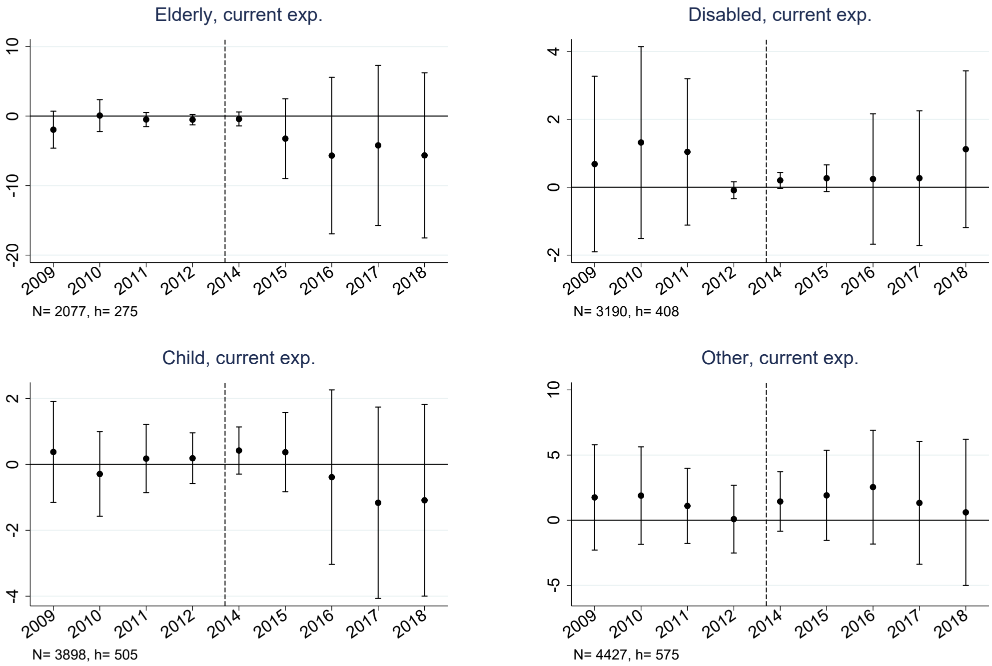
Appendix 3.B Pre-existing policies and potential confounders

Pre-existing regulations varying at the relevant population threshold include the financial compensation for the mayor and members of the executive council, or council size; an overview of this legislation is provided by Grembi, Nannicini, and Troiano (2016) for the period 1999-2004 and De Benedetto and De Paola (2019) for the years 2009-2014. Such time-invariant discontinuities in policies are accounted for by the diff-in-disc strategy, so long as Assumption A2 is satisfied.

With respect to other legislation regulating gender representation within institutions, Law 215/2012 introduced gender quotas for candidates in municipal elections for municipalities with 5,000 residents or more and modified previous legislation

Figure 3.13: Local parallel trends, expenditure in social services - subgroups

(a) Current expenditure



(b) Capital expenditure

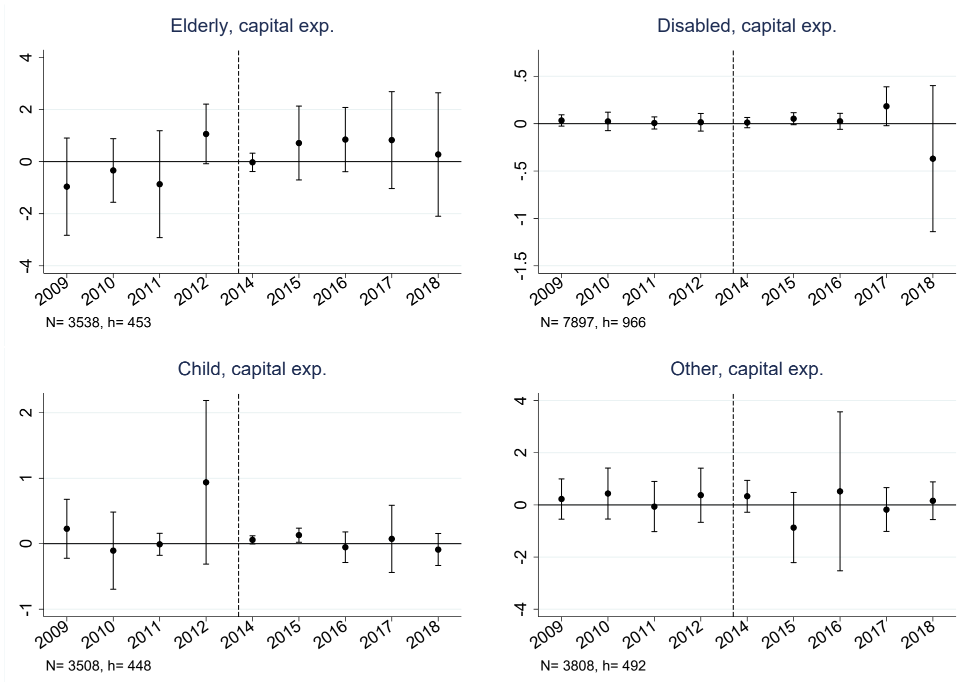


Table 3.10: Social expenses, total accrued per year, per capita

	Current expenses					Capital expenses				
	(1) Social	(2) Child	(3) Elderly	(4) Disabled	(5) Other	(6) Social	(7) Child	(8) Elderly	(9) Disabled	(10) Other
$S_{i,t} * T_{i,t}$	-4.6	-2.1	1.2	-3.0	-2.4	-5.4	-3.9	0.5	0.4	-0.8
	(10.0)	(8.6)	(2.2)	(7.6)	(2.2)	(5.8)	(4.4)	(1.1)	(1.9)	(1.2)
Observations	2676	2087	3110	3269	5665	6540	3998	3898	4644	3249
CCT pop bandwidth	351	277	401	419	712	813	523	504	598	416
$\bar{Y}_{pc}$	74.9	8.5	2.8	49.7	9.9	16.0	3.0	0.4	2.1	0.8

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(DL n. 267/2000) to guarantee the presence of both genders in executive committees, regardless of population size. The introduction of this candidate gender quota increased by 13.8 percentage points the share of female councilors and by 7.4 percentage points the share of women in municipality executives (Andreoli, Manzoni, and Margotti 2021). This will not affect our analysis, for the first part because our population of municipalities is below the relevant threshold for candidate gender quotas, and for the second part because we focus on municipalities with elections in 2009 and 2014. Thus the application of the modified regulation overlaps, and is a subset of, the introduction of the executive gender quota rule. In addition, municipalities can have specific provisions regarding gender representation within institutions in their charters, which we do not observe. However, the available data does not show any evidence of systematically different behavior in this sense across the 3,000 population threshold nor over time, as the share of women in municipal executive committees is continuous across the cutoff before 2014 (Figure 3.3).[10]

Subsection 3.B.1 Changes in Council and Executive Size

A potential concern of our empirical design is that the reform under study also modified council and executive size, which change sharply at the 3,000 threshold. Changes in the maximum number of executive members can introduce measurement error in our policy-relevant variable, the female executive share. Therefore, throughout our discussion, we also consider how the absolute number of women changes, showing that the observed change in the executive female share is not driven only by changes in the denominator (as confirmed by Figure 3.4). Changes in the size of the municipal council might also be relevant. However, previous literature has not reached an agreement regarding the relationship between legislature size and government spending at the local level, finding both positive and negative effects (Alptekin et al. 2020).

The number of maximum seats in the municipal council and the municipal executive changed several times between 2011 and 2014, with varying intensity depending on population size. seats summarizes these regulatory changes across various population thresholds, and what number of women would correspond to 40% of the executive committee for each population bin.

10. Note that municipal charters have to adjust to the provisions of L.56/2014, therefore their role can only be to anticipate the introduction of a gender quota in the executive committee or to increase the share of seats reserved for women, not to set it below 40%.

Table 3.11: Evolution of municipal councils composition laws in Italy, 2009-2014

	(1) Before 2011	(2) 2011	(3) 2012	(4) From 2014
Residents	*a) Max. council members (excl. mayor)*			
0-1,000	12	9	6	10
1,000-3,000	12	9	6	10
3,000-5,000	16	12	7	12
5,000-10,000	16	12	7	12
	b) Max. executive council members (excl. mayor)			
0-1,000	4	0	0	2
1,000-3,000	4	2	2	2
3,000-5,000	4	3	3	4
5,000-10,000	4	4	4	4
	c) Number of women corresponding to 40% of executive (incl. mayor)			
0-1,000	2	0	0	1
1,000-3,000	2	1	1	1
3,000-5,000	2	2	2	2
5,000-10,000	2	2	2	2

By restricting the sample to municipalities with elections in 2014, we fix pre-treatment council and executive size to the 2009 regulation and thus filter out changes in council size occurring in 2011 and 2012.[11] Therefore, we only need to compare columns (1) and (4). The new regulation applied from 2014 prescribes that the maximum number of executive seats decreases from 4 to 2 for control municipalities, i.e. with less than 3,000 residents, while treated municipalities see no change in the number of executive positions. This implies that for control municipalities even a smaller number of women than before would suffice for them to align with the gender quota, potentially increasing the number of compliers in the control group and attenuating our estimates. However, the evidence in Figure 3.3 shows that there is no take up from control municipalities; on the contrary, Figure 3.4 informs us that in control municipalities the number of women in the municipal executive marginally decreases after 2014, from an average of 1 to 0.8. This suggests Moreover, from 2014 both control and treated municipalities will see a reduction in the size of their municipal council which jumps from 12 to 10 (-16%) and from 16 to 12 (-25%), respectively.

11. The regulatory changes introduced in 2012 affect municipalities that have elections in 2012, 2013 and thus 2017, 2018. Note that for this other sample municipalities, the change of executive size introduced in 2014 has no bite on the absolute number of women needed to comply with the 40% gender share representation, because of rounding (see Table 3.11).

Subsection 3.B.2 Joint Provision of Childcare

Municipalities can be reasonably expected to offer childcare or elderly services jointly with their neighbors, particularly when they have small population size. This introduces two potential threats to identification: on the one hand, the number of authorized places might be underestimated; on the other hand, there may be a spillover effect on control municipalities. Either way, this would imply a downward bias in our estimates.

As municipalities jointly provide childcare or elderly care, we might observe a mass of municipalities authorizing zero childcare or elderly care. However, these municipalities would still have a positive expenditure for such services, even if they are not provided in loco. Between 2009 and 2018, 64% of municipalities with a resident population between 1000 and 5000 authorized zero childcare places, while 50% had zero expenditure for childcare services. Nevertheless, these issues should not pose a threat to identification if a) they are not exacerbated (or attenuated) by the reform, and will thus be captured by the diff-in-diff; b) they are evenly distributed across the 3000 residents threshold. Figure 3.14 shows that the probability of having zero childcare places per 100 children decreases sharply with population size, but is continuous across the 3000 residents threshold.

Figure 3.14: Probability of zero childcare by population size