

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

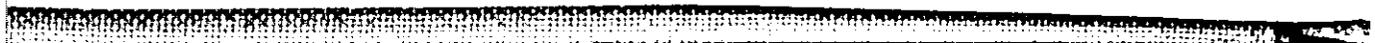
Mergers and Information

Inês Cabral

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

Florence
October 2001

B/c →



European University Institute



3 0001 0034 7950 0



EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

14
Rocco
LIB
338.83
CAB



Mergers and Information

Inês Cabral

The Thesis Committee consists of:

- Prof. Pedro Barros, Universidade Nova de Lisboa
- " Vincenzo Denicolò, Università di Bologna
- " Massimo Motta, EUI, Supervisor
- " Karl Schlag, EUI

Vertical line on the left side of the page.

Vertical line on the right side of the page.

Contents

1 Mergers, Information and Profitability	4
1.1 The Chapters	5
1.2 Merger Waves	10
1.3 Post-merger Performance	11
1.4 Conclusion	13
2 Information Exchange: Mergers and Joint Ventures	15
2.1 Introduction	15
2.2 Literature	16
2.3 The model	18
2.3.1 Benchmark case: no exchange of information	20
2.3.2 Merger	22
2.3.3 Joint Venture	25
2.4 Nash Equilibria	26
2.5 Expected Profits	28
2.5.1 Impact on participant firms' profits	30
2.5.2 Impact on rivals' profits	33
2.6 Expected consumer surplus	34
2.6.1 Impact on consumer surplus	35
2.7 Expected total surplus	36
2.8 Bertrand Competition	37
2.9 Future Research	39
2.10 Conclusion	40
2.11 Bibliography	41
2.12 Appendix	42

3	A Herding Approach to Merger Waves	44
3.1	Introduction	44
3.1.1	Mergers' performance	46
3.1.2	Herding	47
3.2	The model	48
3.3	Solving the model	51
3.3.1	Pessimistic firms	51
3.3.2	Optimistic firms	55
3.4	Merger cascades	58
3.4.1	Complete range of herding histories	61
3.4.2	Waves	63
3.5	Expected Utility	64
3.5.1	Third firm	65
3.6	Magnitude of potential losses	66
3.7	A Contagion Model	68
3.8	Concluding remarks	69
3.8.1	Future work	71
3.9	Bibliography	72
3.10	Appendix	74
4	Strategic Effects in Merger Waves	80
4.1	Introduction	80
4.2	Literature	81
4.3	The model	83
4.4	Equilibrium	86
4.5	Myopic firms	94
4.5.1	Sequentiality	94
4.5.2	Initial beliefs	95
4.6	Conclusion	96
4.7	Bibliography	98
4.8	Appendix	100

Acknowledgments

I am very grateful to my supervisors, Massimo Motta and Karl Schlag, for their help and guidance of my research. I thank the Portuguese Ministry of Foreign Affairs, the Portuguese Ministry of Science and Technology, and the European University Institute for financial support. I also thank the University of Bonn for its hospitality in the last semester of my studies.

Vertical line on the left side of the page.

Vertical line on the right side of the page.

Chapter 1

Mergers, Information and Profitability

Despite the fact that mergers are a ubiquitous phenomenon of economic life, the reasons why firms do merge are yet only poorly understood. This thesis attempts to improve our theoretical understanding of mergers. A merger consists of a legal transformation through which two or more formerly independent firms come under common control. Mergers are particularly interesting and pose a wide variety of research questions due to some stylised facts. World wide merger activity was outstanding during the last decade (peaking in 2000 with a volume of more than \$3 trillion US dollars¹), showing that the merger strategy has been preferred to alternative corporate strategies such as re-structuring or expansion. Despite these intense merger movements, there is no clear evidence that firms improve performance after merging. On the contrary, indicators tend to point to a poorer performance relative to the industry. Finally, and perhaps most interestingly, mergers appear to occur in 'waves'. Waves in the sense that the number of past mergers affects positively the number of current mergers up to a point where the number of mergers starts smoothly decreasing. The current merger movement that initiated the downturn in 2001 corresponds to the fifth historical merger wave, being the first one back in the nineteenth century. The findings on poor post merger performance apply to most of the previous waves.

The thesis is composed of three research papers on mergers and information. This introductory chapter is organised as follows. It starts by clarifying why the focus is on information. It then moves on to outline the relevance of the three studies, briefly explaining the analysis developed in each of them and the main conclusions. Finally, it provides evidence for the stylised facts claimed in the previous paragraph.

As the literature of the last sixty years has shown, the analysis of mergers is a very broad area of research covering incentives, potential effects and policy issues raised by merger activity. This work uses results from the existing literature to motivate, illustrate and to support the theoretical analysis.

¹Thomson Financial Securities Data.

The thesis makes a timely and valid contribution to the existing literature by exploring the issue of information which has not previously been considered. The element information is present in the three chapters and, as results will show, it can be a strong determinant of merger deals. Information is explicitly modelled in each of the three following essays. Profitability, as the driving force of firms' decisions, necessarily plays an important role throughout the chapters, too. Profits are modelled using standard oligopoly theory and existing empirical evidence on merger profitability is brought forward, in order to motivate the questions. The second chapter, entitled "Information Exchange: Mergers and Joint Ventures" may stand on its own. Chapters three and four, "A Herding Approach to Merger Waves" and "Strategic Effects in Merger Waves", respectively, refer closely to each other since they both centre on the phenomenon of waves.

Throughout the entire analysis, no distinction is made between the notion of merger and take-over (i.e. acquisition). Both deals bring formerly independent firms under common management. In the case of a take-over, there is only a transfer of control, whereas in the case of a merger, there is a complete integration of two or more independently run firms. Clearly, such a distinction plays no role in standard oligopoly theory since no attention is paid to the entity which actually has the control. Additional comfort for such an abuse can be drawn from the telling fact that recent empirical studies also tend to aggregate the two concepts. Due to the increasing complexity of the actual agreements, it is argued that it has become increasingly difficult to define a border between them anyway.

1.1 The Chapters

Chapter two builds on a well-known debate in Industrial Economics. It is generally held by oligopoly theory that, when firms compete in quantities, mergers typically reduce the joint profits of the participating firms. In other words, mergers are not profitable unless a very large percentage of the industry participates in the deal (Salant, Switzer and Reynolds, 1980). A merger does increase price, but mainly to the benefit of other firms. In response to a merger, it is optimal for competing firms to expand output and, as a result, the merger produces less than the merging parties, if they had remained independent. It is easy to show that industry profits do increase but the merging parties are actually worse off than without merging. This is not surprising given that, using the standard Cournot setting, (where firms are identical and compete in quantities), firms are symmetric both before and after the merger takes place. A merger of m firms in an industry of n firms ($m < n$) corresponds to a withdraw of $m-1$ firms from the market. There is an increase in concentration, since the market has now only $n-m+1$ firms, and therefore, individual profit is unambiguously higher. However, the profit of a single firm is now smaller than that of the m firms in the pre-merger equilibrium, unless m corresponds to more than 80% of the industry. Perry and Porter, (1985) suggest a competitive

advantage arising from the decrease in costs after merging. By introducing this asymmetry, merger profitability can be achieved. Deneckere and Davidson (1985) investigate the incentives to merge in a more general framework and conclude that mergers are disadvantageous when firms compete in quantities and beneficial when firms compete in prices. Such difference follows from the fact that reaction functions are typically downward sloping in quantity setting games but upward sloping in price setting games. As a result, if competition is in prices, the response of the rival firms actually reinforces the initial price increase due to the merger.

The second chapter investigates the incentives for information exchange between two firms and analyses the case of a merger and a joint venture in a quantity setting game. Market demand is random and therefore firms tend to overproduce in periods of low demand and under-produce in periods of high demand. Private information regarding demand can be seen as an asset of each firm and a merger or a joint venture allows firms to pool this information. As a result, firms may have a more accurate expectation of the demand and therefore, may perform better than competitors in the Cournot market. One of the questions asked in the chapter is whether such exchange of information between firms can act as an efficiency gain and generate profitable mergers.

More precisely, the basic setting consists of an oligopoly of many identical firms facing a stochastic demand. Before setting the quantity to produce, each firm receives a private signal, concerning the state of demand, that will influence the choice of the quantity. A merger of two firms generates a single firm with pooled information regarding the state of demand. A joint venture is defined as a deal in which the two firms share their private information about market demand but set quantities independently.

The chapter conveys the following conclusions. The informational advantage that the two merging firms obtain is not enough to achieve merger profitability. Once again, the outsider firms are benefiting the most from the concentration in the industry. By studying the case of a joint venture, the information advantage can be disentangled from the benefits of concentration. The model is solved for four sets of parameters that encompass the interesting scenarios. It is found that a joint venture can be profitable if correlation between signals is neither too high nor too low and that a joint venture is never welfare decreasing. Intuitively, on the one hand, for low signal correlation, firms can learn very little from each other since signals are almost independent, on the other hand, for high signal correlation, firms do not need to form a joint venture to learn from each other. The main contributions of this work to the literature on information exchange are mainly two. First of all, the introduction of a merger that sets-up an asymmetric industry structure and second of all, the analysis of partial and not industry-wide agreements between firms. Michael Raith (1996) presents a general model of information sharing in oligopoly that encompasses virtually all models in the current literature as special cases. Such models vary along several aspects: choice variable being price or quantity, uncer-

tainty about a common demand function or uncertainty about rivals firms' costs, degree of precision of the reported signals, among others. Through an appropriate specification of the parameters, Raith is able to account for these dimensions. However, the number of firms participating in the information exchange is not one of the dimensions explored since the focus is only on industry-wide agreements.

The third and fourth chapters of the thesis tackle two features of merger activity that - at first sight - seem to be unrelated. The first one is the phenomenon of the merger and the second one is merger profitability. As commented in section 1.2, empirical studies display very puzzling findings about mergers' performance. Furthermore, there is evidence of waves that definitely did not turn out well financially and where there were many failures (like the third and the fourth movements). Given that firms generally have the option of staying independent, why, then, are so many 'unprofitable' deals carried out? It is not the single unfortunate transaction but the stream of bad deals, which constitutes the puzzle. In these chapters, it is suggested that due to a herding mechanism, firms may engage in merger deals that then turn out to be unprofitable. Hence, the claim is that an informational cascade can be more responsible for the merger waves than relevant economic fundamentals.

Chapter three proposes a herding mechanism that attempts to solve (at least part of) the merger puzzle: such intense merger activity with such weak evidence of positive results from it. The herding literature is commented on with some detail in the chapter and hence, the current section is brief on this issue. Basically, herding is said to occur when everyone is doing what everyone else is doing, even when private information suggests doing something different. The herding mechanism works as follows. The usage of information conveyed in decisions taken by others makes each person's decision less responsive to her own information, and as a result less informative to future decision makers. The reduction of informativeness may be so severe that ex ante, the group of people (or even society) may be better off by constraining some of the people to use only their own information. Herding has been theoretically linked and empirically tested in some economic activities such as investment recommendations (Scharfstein and Stein, 1990), price behaviour of IPO's (Welch, 1992), analysts' forecasts (Trueman, 1994), brokerage recommendations (Welch, 1996), and investment newsletters (Graham, 1999). To our knowledge, however, there are neither theoretical nor empirical studies which apply the herding concept to merger activity.

The model in chapter three closely follows the information structure of Banerjee (1992). Banerjee's herding model is developed in a symmetric setting where the space of private signals is infinite. Our model, however, considers only two signals and an asymmetric setting. It also analyses explicitly the effect of prior beliefs on agents choices. It is interesting to note that priors can be directly connected with the motives for mergers described in section 1.4. Priors can trigger or stop a merger wave but play a small role in the herding process. As it occurs regarding some of the motives for and profitability of mergers, there is no clear correlation between prior beliefs and actual outcome of the

merger wave. The example of the stock market is noticeable. The stock market is believed to foster merger activity: "When the stock market is booming, managers are more emotionally ready to take big decisions. When the market is down, deals seem much riskier"². However, the wave of the eighties coincided with a sharp decline in stock prices (see section 1.4): it was less expensive to buy already existing firms than to expand businesses or create new ones. Recall that the market price of a firm incorporates expectations of future profits. These anecdotic events show how weak the relationship between the beliefs about profitability, and the economic fundamentals behind a decision to merge, might be. Clearly, if a bullish stock market would systematically foster and bring positive results to merger activity, there would be no room for the herding argument.

In our model, in order to isolate the herding effect, firms are considered to be from different industries (i.e. of the conglomerate type). Firms are sequentially called upon to choose between 'merge' or 'not merge' and hold 'pessimistic' or 'optimistic' prior beliefs about the impact of a merger on the firm's value. There are two possible states of the world: one is favourable and the other one is unfavourable to merger activity. In the first case, all merger deals are profitable whereas in the latter, all merging firms make a loss. Many factors that strongly influence the success of a merger concern the merger deal in itself and do not depend so much on the industries involved. The cultural cost of merging is a well-known example. It consists on the obstacles arising by putting teams with different nationalities, or former rival teams, working together for the same goal. Another example is the toughness of antitrust laws. Regulation about market shares and market dominant positions are general and do not concern specific industries or deals. The strictness of regulators, that varies substantially with political regimes, also affects industries and deals in quite the same way. Firms may get a signal that enables them to infer which is the current state of the world. It is shown that it can be rational for firms to imitate other firms' decision to merge, despite holding private information that supports a non-merging strategy. Due to a reduction of information passed on to subsequent decision makers, unprofitable merger waves can arise in equilibrium. A merger cascade under 'optimistic' beliefs, which sustain that a merger is profitable, occurs with higher probability than under 'pessimistic' beliefs, which sustain that a merger is unprofitable. However, and somehow surprisingly, a merger wave starts earlier under the 'pessimistic' setting than under the 'optimistic' one. Intuition for this result is actually simple. In the pessimistic environment, the first firm who decides to merge must have had a signal to do so, whereas in the optimistic environment, the first firm who decides to merge can be acting according to the initial bias that merging is profitable and may not even have received a signal. It is also shown that in the unfavourable state of the world, signal precision reduces the probability of herding, while the probability of getting a signal has a non-monotonic impact on herding behaviour.

²The Economist, January 27th 2001, pp.67.

Finally, the last chapter underpins the importance of information in merger decisions when there is herding but where firms compete in the same product market. Instead of learning about the profitability of merging from other industries, now we assume that firms learn from other firms in the same industry.

Recall the discussion about merger profitability where firms compete in quantities. All firms would be better off by reducing the number of competitors (benefit from concentration). However, firms would benefit the most by remaining independent while other firms merge (free riding). Based on the herding model from last chapter, but allowing for strategic interaction between firms in the same industry, this paper is able to combine the effect of herding with that of free riding in merger decisions.

Competition is modelled using the Cournot framework with a small number of initially identical firms. A decision to merge is risky given that it may bring efficiency gains or efficiency losses depending on the realisation of the state of the world. In case of a negative payoff, firms are allowed to shut down and make zero profit. Once again, firms receive signals about the likelihood of states of the world and sequentially decide to merge or not. Whereas in the simple herding model of chapter three, the option not to merge delivers the same payoff to all firms, in the intra-industry setting of chapter four, the option not to merge delivers a non-negative payoff that is increasing in the total number of mergers. Therefore, even if the information available supports a decision to merge, the benefits from the already high concentration in the market may lead the firm to decide not to do so. As a result, herding may be more difficult to arise in this setting.

The main conclusion is that herding can arise in equilibrium for intermediate values of risk, defined as a moderate spread between potential gains and losses. For high risk, signals become uninformative since firms have incentives to gamble and shut down in case the bad state occurs. For low risk, to be an outsider is again more attractive, even if the favourable state of the world occurred and firms who decided to merge did receive a signal to do so. As a result, in the extreme cases (high risk, low risk), herding never takes place. For intermediate levels of risk though, despite the existence of an attractive outside option from not merging, herding can occur. Even with an already high concentration in the market due to mergers accomplished by previous decision-makers, a firm that received a signal not to merge, may herd on the action to merge. This set of levels of risk happens to be the largest over the range of possible values. Note that herding may explain both, the existence of waves of successes and waves of failures. More attention is placed on the latter because it seems more difficult to justify under rational behaviour.

The relevance of herd behaviour that emerges from these studies may legitimise the opinion of some commentators who state that "[in recent merger deals] much of the attention seems to be on

the deal, rather than the integration” and that “many deals are rushed”³. It could well be that many firms simply follow other firms’ decision to merge, without having private information sustaining that strategy.

The following paragraphs present a summary of the historical merger waves and empirical evidence on merger performance that helps understanding the questions posed by each paper, and then we will conclude.

1.2 Merger Waves

The world’s merger activity is marked by five prominent waves. The first two waves took place mainly in the United States, in 1887-1904 and in 1916-1929, respectively. The turn of the century wave, generally referred to as “The Great Merger Wave” had weaker parallels in Great Britain and Germany. Then came the 1960’s merger wave, peaking at the end of the decade, with a significant expression in Europe and Canada. The fourth wave occurred in the 1980’s and we are presently witnessing the fifth merger wave, starting in the early 1990’s with the peak in 2000. The last two merger movements were more relevant in Europe than in the U.S., both in terms of the number and the value of the deals. The first and second waves are characterised by horizontal mergers, i.e., mergers between firms belonging to the same market. The absence of antitrust defence at the beginning of the century allowed for the creation of ‘giants’. Seventy five percent of the deals of the Great Wave involved at least five firms, of which 26 percent involved more than ten firms. The second merger movement was more moderate in terms of concentration due to the first steps of antitrust concern in the US. Vertical mergers are those in which firms in the transaction belong to an upstream and downstream position with respect to the production process, (mergers between a supplier and its customer, for example). This type of merger also appears significantly in the manufacturing and mining industries of the second merger wave. George Stigler refers to the first movement as “mergers for monopoly” and the second one as “mergers to oligopoly”⁴.

The 1960’s merger wave is also known as the “Conglomerate Merger Wave”. Conglomerate mergers may aim at product line extensions, geographical market extensions, or may even lack any sort of complementarity (denominated pure conglomerates). The merger movement of the sixties was of conglomerate character, perhaps due to tougher restraints against horizontal and vertical mergers. U.S. data on manufacturing and minerals industries shows that at the peak of 1968, the percentage of the assets acquired through a pure conglomerate merger had risen from 10 to 35 percent. By including product line extension and market extension mergers in a broad conglomerate definition, this number

³The Economist, January 9th 1999, pp. 21.

⁴George Stigler, “Monopoly and Oligopoly by Merger”, American Economic Review, vol. 40, pp. 23-34.

rises to 81 percent against 19 percent of horizontal and vertical mergers⁵. Despite a still significant horizontal activity, probably due to weaker antitrust laws, Europe also witnessed a relevant rise in conglomerate mergers.

The 1980's merger wave was again characterised by deals occurring mainly among firms from different industries, and was fostered by the new financial devices able to facilitate take-over bids. It is the era of junk bonds financing to fund acquisitions. As in the present merger movement, Europe and not the U.S., had the lead. Finally, in the 1990's merger movement, broadcasting, communications, banking and finance, computer software, oil and gas were the industries showing the most intense merger activity⁶. Hence, despite competition policy concerns, there seems to be a tendency to return to horizontal mergers or at least deals where there is significant overlapping. Many financial professionals believed that it would be difficult for merger and acquisition activity in 2000 to top 1999's year. While merger activity was tempered in the U.S, it was outstanding in Europe setting a 2000 single-year record (over 3 trillion US dollars). Only in 2001 there are signs of a drop in merger activity (around 600 thousand US dollars until september 2001). The figure below illustrates the last thirty years of worldwide merger activity in terms of total number and total value of deals⁷.

1.3 Post-merger Performance

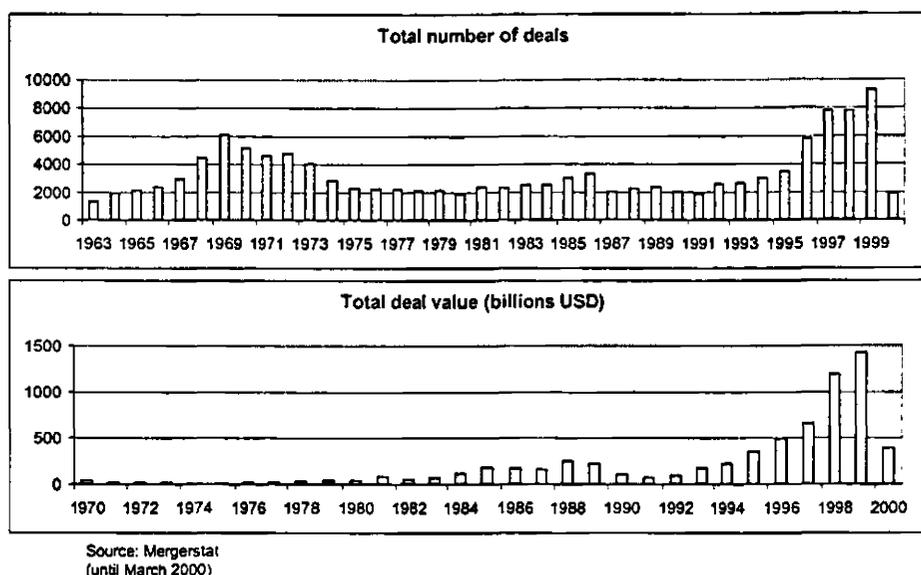
When referring to the profitability of a firm, empirical studies generally need to come up with proxies given that economic profits are not easily observable, and that accounting statistics are not considered good indicators. There are basically two approaches to the quantitative assessment of economic consequences of a merger. The so-called event studies are based on the assumption that stock markets are efficient. They examine price behaviour around the announcement of a forthcoming merger using Capital Asset Pricing Theory. If stock markets are efficient, this should constitute a good proxy for firms' expected profitability. The other approach analyses changes in firms' performance indexes such as sales, market shares, return on assets and return on equity. Divestiture or split-offs are also considered an important variable in the economic assessment of mergers. In the jargon of the literature, while the performance indexes are associated to profitability, split-offs are said to assess successful or unsuccessful mergers. This approach is typically developed over a longer time span. Comparisons between pre and post merger performance may reach ten years of company activity.

⁵Table 5.1 in Scherer and Ross, *Industrial Market Structure and Economic Performance*, 1990, 3rd Ed.

⁶Mergerstat, www.mergerstat.com, reports on M&A activity according to data compiled by Thomson Financial Securities Data on world wide M&A.

⁷Aggregate values are calculated on the 'Base Equity' price offered and ignore 'small' transactions according to Thomson Data definition. These numbers are not comparable with values for recent years that appear in brackets in the text. The latter ones, also from Thomson, refer to the 'Total Value of Deal' excluding only non completed operations.

Figure 1.1: World wide Mergers and Acquisitions



The first method is based on the assumption that stock markets at any moment reflect all available information about future economic events, and that current prices are unbiased predictors of future prices. The stock price movements of the firms in question are normalised by controlling for overall stock market movements. The standard merger event study analyses stock price changes only a few days on each side of the key announcement. Banerjee and Eckard (1998) study competitive effects of the first merger wave 1897-1903. The analysis reveals gains of about 12% to 18% on the value of merger participants in the industry and mining sectors. Given that this wave is characterised by multi-merger participants who merge to monopoly, the results are not surprising, but have little in common with the merger scene today. In a short window of time around a post 1950's merger event, target firm shareholders experienced stock price gains, while acquiring firm shareholders were neither better nor worse off. Evidence from the third and fourth merger waves is consistent with these results. Acquisitions entail a gain for the target firm's shareholders (see Jensen and Ruback (1983), Eckbo (1983), Bradley, Desai and Kim (1988)). The average return to the bidding firm's shareholders is less clear. Other studies find small but still statistically significant gains, some others find losses. Given that the bidding firm is typically larger than the target one, it is not accurate to state that a positive variation on the value of the target firm together with a close to zero variation on the value of the bidder, will yield a total positive increase in value of the resulting firm. Bradley et al. (1988) propose a weighted average gain of the deal. The authors obtain an average gain of 7.4% to both firms (target and bidder) in tender offers, yet a quarter of the combined valuations were negative, and the

bidders' shareholders gain in only 47% of the cases. As said before, the standard merger event study involves only a few months (or days) on each side of the announcement. However, when the coverage is extended to a year following the announcement using the same methodology, studies surveyed by Jensen and Ruback show an average bidding firm stock price decline of 5.5%. When the analysis is carried out for the three successive years following the announcement, the average price decline is 16% in mergers and take-overs during the seventies and the eighties⁸.

The second approach, which is closer to empirical Industrial Organisation literature, may include other indicators of performance generally over a wider sample period. The results are more impressive and can be summarised as "evidence shows...that the profits from merging companies generally decline after the mergers", Mueller (1993). In an international comparative study by Mueller (1980), it is shown for five countries that the relative value of the bidder's share rises in the year of the merger, then falls to zero or below after three years. Another comparative study, this time between pre and post merger performance of firms in the US and Great Britain from the 1950s to the 1970s, involves roughly three years before the merger and six years after it took place⁹. Results show that US acquirers lose only moderately, whereas 60% of the sample of British acquirers who outperformed the industry before the merger showed losses in subsequent years. Another possible indicator of poor performance is the number of divested acquisitions in which acquired business units are sold or re-established as independent firms. Studies on US data by Ravenscraft and Scherer (1987) show that, on average, mergers had substantial negative effects on the profitability of acquired business units. From 1974 to 1981, one fifth of the sample business units were sold off. Note, however, that the majority of these studies use data from the Conglomerate wave sample that did not turn out well financially. There are not many research studies about the current wave to allow for a global appreciation. Anecdotal evidence, though, has already counted a great deal of failures. A recent paper about the Italian banking industry (one of the active consolidating sectors of the fifth merger wave) studies separately mergers and acquisitions from 1984 to 1996. Despite being considered a very profitable sector, results are not very impressive about the benefits of mergers and acquisitions. Mergers show an increase in staff costs which makes a better post merger performance impossible, whereas acquisitions seem to show a slight improvement in performance¹⁰.

1.4 Conclusion

Given the ambiguous results about merger profitability, it is legitimate to ask what the motivations are then for such intense merger activity. The typical motivating force for horizontal mergers is monopoly

⁸ Results by Magenheim and Mueller discussed in Scherer and Ross.

⁹ This discussion involves results from various studies referred in Caves, IJIO, 1989.

¹⁰ Focarelli, Panetta and Salleo, Banca d' Italia, paper n. 361.

power. This reason definitely played a major role in the consolidations that took place around the turn of the nineteenth century. However, due to competition policy concerns, this motive for mergers has since lost importance. Speculative motives are thought to play a relevant role. During the first merger wave, while the stock market was booming, stocks of many newly consolidating firms were sold at prices far exceeding their true economic value. Despite an overall profitable wave, many deals were financial failures and collapsed shortly after formation¹¹. The wave of the eighties, instead of being stimulated by a stock market boom, coincided with a decline in stock prices. It was often less expensive to expand businesses by buying other companies rather than by improving or replacing capital equipment. The wave of the nineties is again believed to be related to a stock market boom. Hence, the typical alleged connection with the stock prices seems unfounded, given that both high stock prices and low stock prices appear to be able to trigger a merger wave. Another possible motive for mergers is the one of replacing inefficient management or saving companies that otherwise would fail. Despite its appeal, such a reason seems not to be quantitatively important, since a small percentage of firms have negative profits before the year of acquisition. Tax considerations may also affect merger decisions but are not believed to be critical enough to be a determinant reason for a deal. Synergies and efficiency gains are the preferred economic motivation for mergers. They constitute the typical defence of challenged mergers on trial. It is nevertheless very difficult to assess such gains quantitatively, and if economies of scale were significant they should have a positive impact on some measure of performance, something the empirical literature does not unambiguously illustrate. Another explanation is based on the assumption that managers and shareholders have different preferences and, as a result, managers, who take a decision to merge, may not be maximising the value of the company but undertaking personal objectives. The major weakness of such a motivation is that hostile takeovers, which drive out all the managing staff of the acquired firm, and layoffs, which frequently occur due to a merger, can be seen as devices to discipline managers. In this sense, a merger deal creates the possibility of bringing on board better management and new talents, as pointed out previously. In short, apart from the opportunity to raise market power, which was the driving force and which generated some very profitable merger deals in the first merger wave, no other motive (or combination of motives) constitutes an empirically manifested explanation for the cause behind the other waves

This introduction has highlighted the importance of information for understanding mergers. The next three chapters will explain the models described above in more detail. Chapter two shows that mergers will not arise through exchanging information. Chapter three and four speculate that merger waves do not necessarily occur because they are profitable but because the act of merging can be contagious.

¹¹Livermore (1935) in Scherer and Ross.

Chapter 2

Information Exchange: Mergers and Joint Ventures

Abstract: This chapter compares the effect of a merger and of a joint venture of two firms on the welfare of all market participants in a Cournot industry with stochastic demand. The asset pooled in a merger or a joint venture is information about an unknown future market demand. A joint venture is defined as an agreement in which firms exchange information about demand but quantity decisions remain independent. The analysis is undertaken for four sets of parameters to exemplify the interesting cases. Two main conclusions can be drawn. It is shown that a merger is neither profitable nor welfare increasing. Second of all, a joint venture can be profitable and is not welfare decreasing. Profitability is obtained when signal precision is high and the correlation between firms' signals is moderate. Under such conditions, higher efficiency is achieved: the joint venture is profitable, the firms involved perform better than the market, and consumers are better off.

2.1 Introduction

This chapter analyzes the impact of mergers and joint ventures in a market with uncertain demand. A joint venture is defined as a market research agreement in which the participating firms exchange information about unknown future demand. Firms are assumed to be identical apart from a signal of the state of demand observed privately by each firm. By merging or forming a joint venture, the participating firms share the private signals and hence enlarge their information set. In other words, by pooling their individual information about the market demand, firms can decide more accurately on the quantity to produce. When demand is volatile, it is likely that single firms underproduce in high demand periods and overproduce in low demand ones. With more information about the state of demand, firms' output is more likely to meet demand shifts and therefore productive efficiency can be

enhanced. It is unclear whether the exchange of information between firms involved in a cooperative agreement (like a merger or a joint venture) is welfare improving since, apart from efficiency gains, there is the strategic quantity choice of the cooperating parties which affects consumer surplus. Using a simple framework, this work attempts to quantify these effects and to bring some insights about the adequate policies in markets with highly volatile demand.

2.2 Literature

The literature on information exchange deals with the incentives firms' have to share their private information and whether information sharing among all firms can be welfare improving. The source of market uncertainty can be common (affecting all firms in the same way) or private (affecting firms differently). The first one is usually associated with a market demand which is stochastic and the second with uncertainty about individual cost functions. The type of competition can be in quantities (*Cournot*) or prices (*Bertrand*). This work looks at the common market uncertainty when firms compete *à la Cournot*. Some considerations are made in what concerns *Bertrand* competition. The problem of information sharing is twofold. It is reasonable to expect that more accurate information should benefit firms individually. In other words, a firm's expected profit increases with an increase of information that it gets to see alone. On the other hand, pooling information means that all firms participating on the agreement see their information sets improved. This leads to a higher correlation between the firms' outputs which reduces expected profit. Intuitively, when firm i observes a signal of low demand and reveals it to others, it reduces the likelihood that competitors overproduce and this affects positively firm i 's profit; when firm i observes a signal of high demand and reveals it to others, it reduces the likelihood that competitors underproduce and this affects negatively firm i 's profit. The overall effect of sharing information is the sum of the amounts of over- and underproduction multiplied by the respective equilibrium prices. Given that the probability of high and low demand periods is the same¹, and that prices are higher in high demand periods, the expected profit of firm i when it chooses to conceal its signal is higher than the expected profit from revealing it. Therefore, a general conclusion of this literature is that there is never a mutual incentive for industry-wide agreements if firms compete in quantities. The most important works in this area have been developed by Novshek and Sonnenschein (1982), Vives (1984 and 1990), Gal-Or (1985 and 1986), Clarke (1983) and Raith (1996). Conclusions depend heavily on the type of competition (quantities or prices) and the source of uncertainty (common or private). Raith builds a general model that encompasses all the existing ones as special cases, by appropriate specification of the parameters. Typically, the uncertain market variables are gaussian. Firms receive a signal of the state of the demand and no correlation between

¹The stochastic variable in the demand function follows a normal distribution with zero mean.

signals is assumed. A general conclusion of the referred models is that when demand is unknown, no information sharing is a dominant strategy under *Cournot* competition and sharing is a dominant strategy under *Bertrand* competition. Intuitively, the result shows that the benefits from pooling information and obtaining a larger information set in *Bertrand* are more than offset by the losses from increasing the correlation of firms' output decisions. The opposite occurs in the *Cournot* case.

Our analysis departs from this literature in two respects. First, the game is solved for asymmetric equilibria whereas the existing works in this field impose identical firms. This is because we look at partial and not at industry-wide agreements as commonly found in the literature. Second, it allows for non negative correlation between the signals whereas, in the literature, correlation is typically assumed to be zero unless the market structure is a duopoly².

By merging or forming a joint venture, firms pool their information sets and get a relative advantage comparatively to competitors. Hence, the increased correlation between firms' output is weaker and can be offset by the positive effect of information sharing, i.e., an increase in the information set of the firms involved in the agreement. The accuracy of the information about market demand unambiguously increases and information remains private for the firms involved. In this sense, both mergers and joint ventures can be a source of market efficiency since firms are more likely to meet demand and hence to lower expected price. Even though we look only at an agreement between two firms, some considerations are made in the case of more than two firms participating in the information exchange agreement.

In the absence of information about demand, price fully adjusts to clear the market. When the firm acquires information there is some adjustment in the quantity. The expected price in high demand periods with partial information sharing will be lower than the expected price with no information sharing, and the expected price in low demand periods with partial information sharing will be higher than the expected price without it. The gains from higher production in high demand periods can more than compensate consumers for the losses in low demand periods. It is shown that this is indeed the case when the information agreement occurs through a joint venture but not when it occurs through a merger. Moreover, if these agreements are profitable, they can have a non negative effect in rivals' profits³ and hence increase both the producer and the consumer surpluses.

On the regulatory side, mergers and joint ventures between firms in the same industry have long been a public policy concern. The "abuse of dominant position" or likelihood of collusive practices is the issue of major concern whereas little weight is given to possible efficiency gains. Aside from information sharing effects, the merger implies an output reduction of the merging parties that makes

²Gal-Or allows for positive or zero correlation between signals in the duopoly case.

³A merger is always beneficial for rivals' profits and a joint venture has a non positive but typically negligible effect on rivals' profits.

the welfare improvement more difficult to occur⁴. Salant Switzer and Reynolds (1983) show that, in a *Cournot* framework, there are no incentives to merge except when firms merge to monopoly. It is the increase in production by outsiders following the merger that reduces insider's profits by more than the increase in profits that would have occurred had outsider quantity remained constant. In other words, the output expansion of the outsider firms causes a reduction in profits for the merging parties. Deneckere and Davidson (1985) show that this result depends on the strategic variable being quantities or prices: mergers of any size are profitable if firms engage in price competition. Price-setting games seem to capture the traditional view rather well: mergers are always beneficial to merger parties and become more profitable as the size of the merger increases. In conformity with classical merger literature, we find that a merger is welfare decreasing when competition is in quantities but this should not be a matter of concern since there are no incentives to merge. On the other hand, we find that a joint venture can be welfare improving and increases consumer surplus. Again, this should not be a matter of concern for antitrust authorities since, if there are incentives to form a joint venture then, the joint venture is necessarily welfare improving.

In the literature on mergers, efficiency gains are typically achieved through reductions in marginal cost. Farrell and Shapiro (1990) analyze horizontal mergers in the *Cournot* oligopoly. The authors find that any merger not creating synergies raises the market price. A merger can raise output and make consumers better off only if it permits merging firms to exploit economies of scale or if merger participants learn from it (learning curve argument). In our work, efficiency may arise through information exchange between firms. More information improves decision making and can improve firms' performance in the market. We ask whether efficiency achieved through information exchange is enough to restore merger profitability in quantity setting games. We find a negative answer at least for our information technology.

This chapter is structured in 8 parts. Section 1 explains the purpose of the paper. Section 2 provides a brief discussion on the literature on mergers and information exchange. Section 3 sets out the model that is solved for merger and joint venture equilibria. Sections 4, 5, and 6 look at welfare implications to all market participants. Section 7 sets out some general considerations regarding the *Bertrand* model. Section 8 deals with future research and conclusions are drawn in the final section.

2.3 The model

Our approach follows closely the one suggested by Gal-Or. Consider an industry where firms take decisions under uncertainty due to instability of the demand function⁵. Firms receive private signals

⁴In a joint venture, information between the parties is exchanged but quantity decisions are taken independently hence, welfare improvements are more likely to occur.

⁵Uncertainty can be also due to cost volatility as long as it affects all firms equally.

about the future state of the market demand. One can interpret it as forecasts or studies performed by private research departments, or alternatively, by managers expertise in evaluating market's conditions. Signals are non negatively correlated and hence, from their received signal, firms can infer the rivals' behavior and the aggregate situation of the industry.

The oligopoly consists of n firms producing a homogeneous product and competing in quantities. For simplicity we set marginal costs to zero. Firms face a stochastic and linear demand function given by:

$$p = a - bQ + u, \quad a, b > 0$$

where p is price and Q is the aggregate quantity.

Before deciding how much to produce, each firm observes a noisy signal for u . The signal observed by firm i is x_i that is the sum of two random variables u_i and noise e_i :

$$x_i = u_i + e_i,$$

where $u_i \sim N(0, \sigma)$, and $e_i \sim N(0, m)$. The standard independence conditions apply:

$$\begin{aligned} \text{Cov}(e_i, e_j) &= 0, & i \neq j; & \quad \text{Cov}(u_i, e_j) = 0 & \quad \forall i, j; \\ \text{Cov}(u_i, u_j) &= h \leq \sigma, & i \neq j. \end{aligned}$$

The stochastic variable u is defined as:

$$u = \frac{\sum_{i=1}^n u_i}{n},$$

with unconditional expected value $E(u) = 0$ and variance of the random variable $V(u) = \frac{\sigma}{n} + \frac{(n-1)h}{n}$
(See App.1 for details)

$$u \sim N\{0, V(u)\}.$$

Parameters (σ, m, h) are known by all firms and exogenously given. Denote by x_i the information available to each firm $i = 1, \dots, n$.

Note that, by assuming a normal distribution for the error term of the demand function, prices can be negative. The fact that such an event can occur with an arbitrarily small but positive probability

must be pointed out. We believe, though, that results would not change by assuming that u follows, for instance, a uniform distribution, whereas the interpretation of parameters (σ, m, h) , and the definition of $V(u)$ would not be so appealing.

Equilibrium concept

The stochastic oligopoly model considered here is a Bayesian game. Firms know their own profit functions and the profit functions of their rivals, but are imperfectly informed about variables that specify the market state (the demand intercept). Uncertainty affects the firms in two ways. First, given that the exact environment is unknown, firms must collect information and form estimates of the market state. Second, given that the noisy information available to one firm may not be shared by other firms, firms must estimate the state estimates held by other firms. This is the notion of fulfilled expectations equilibrium that requires not only that firms maximize expected profit given their model, but further that their models not be controverted. Each firm observes its own signal and the price that results from its own chosen quantity, and by taking into account the value of its quantity, it can deduce the joint distribution of its own signal and residual demand intercept. Equilibrium requires that this observed joint distribution coincide with the firm's model. Each firm makes a conjecture about the opponent's strategy. A Bayesian Nash equilibrium is then a vector of strategies and a vector of conjectures such that: (a) each firm strategy is a best response to its conjecture about the behavior of the rivals and (b) the conjectures are right.

The output choice depends upon the information available to the firm. The strategy choice of firm i is $q_i(x_i)$ where $q_i : R \rightarrow R$. The payoff of firm i is given by:

$$\Pi^i [(q_i(x_i), Q_{-i}(x_{-i}))] = E_{x,u} \left\{ q_i(x_i) \left[a - b \sum_{k=1}^n q_k(x_k) + u \right] \right\}$$

The Nash equilibrium of the one shot game is $Q^*(x)$ that satisfies:

$$\begin{aligned} \Pi^i [(q_i^*(x_i), Q_{-i}^*(x_{-i}))] &\geq \Pi^i [(q_i(x_i), Q_{-i}^*(x_{-i}))] \quad \forall_i, \\ q_i^*(\cdot), q_i(\cdot) &: R \rightarrow R \end{aligned}$$

The timing of the game is the following: (1) firms $i = 1, \dots, n$ receive private signal x_i and choose quantity $q_i(\cdot)$ that maximizes conditioned expected profit Π^i ; (2) u is realized and profits are distributed accordingly.

2.3.1 Benchmark case: no exchange of information

When firm i chooses quantity q_i , variable u and the true values of the signals observed by all other firms X_{-i} remain unobserved. Firm i can infer them based on its information set x_i . Hence, $q_i(x_i)$ is

chosen to maximize:

$$\Pi^i = E_{X_{-i}, u} \left\{ q_i(x_i) \left[a - b \sum_{j=1}^n q_j(x_j) + u \right] \mid x_i \right\}$$

given quantities chosen by other firms. Or equivalently,

$$\Pi^i = q_i(x_i) \left[a - b q_i(x_i) - b \sum_{j \neq i} E_{x_j} (q_j(x_j) \mid x_i) + E_u(u \mid x_i) \right].$$

The normal distributions of the random variables and the linearity of demand enable us to derive explicit forms for conditional and unconditional expected values. Given the linearity of the expected value operator, and since the mean value of all random variables is zero, we restrict the solutions to linear decision rules of the form:

$$q_i(x_i) = A_0 + A_1 x_i \quad \forall i=1, \dots, n$$

and verify that they are an equilibrium for each firm i .

By maximizing Π^i taking $q_j(x_j)$ as given (see App.2 for details) we find:

$$q_i(x_i) = \frac{a - b \sum_{j \neq i} E(q_j(x_j) \mid x_i) + E(u \mid x_i)}{2b}, \quad \forall i=1, \dots, n.$$

Calculating conditional expectations,

$$E(u \mid x_i) = \frac{1}{n} \sum_{k=1}^n E(u_k \mid x_i) = \frac{1}{n} \left[E(u_i \mid x_i) + \sum_{k \neq i} E(u_k \mid x_i) \right].$$

The terms in the summation are given by,

$$E(u_i \mid x_i) = \frac{\sigma}{\sigma + m} x_i$$

$$E(u_j \mid x_i) = \frac{h}{\sigma + m} x_i$$

and therefore,

$$E(u \mid x_i) = \frac{1}{n} \left[\frac{\sigma}{\sigma + m} x_i + \frac{(n-1)h}{\sigma + m} x_i \right] = \frac{x_i}{\sigma + m} V(u),$$

(see App.3 for a complete derivation). Using the suggested decision rule,

$$E(q_j(x_j) | x_i) = A_0 + A_1 E(x_j | x_i)$$

where,

$$E(x_j | x_i) = \frac{h}{\sigma + m} x_i.$$

one finds,

$$E(q_j(x_j) | x_i) = A_0 + A_1 \frac{h}{\sigma + m} x_i.$$

Substituting in the first order conditions and solving the system for the constants A_0, A_1 , yields the unique Nash equilibrium of the game (2.1):

$$\begin{aligned} q_i(x_i) &= A_0 + A_1 x_i, & (2.1) \\ \text{where } : A_0 &= \frac{a}{b(n+1)} \\ A_1 &= \frac{V(u)}{b[2(\sigma+m) + (n-1)h]} \\ V(u) &= \frac{\sigma + (n-1)h}{n} \end{aligned}$$

Note that A_0 corresponds to the deterministic solution under quantity competition. When the market demand is known, all firms produce the same quantity which depends exclusively on demand parameters and on the number of firms. In the stochastic case, equilibrium quantities will exceed or be less than this value depending on the realization of the random variable x_i . Given that A_1 is positive, a firm will produce more (less) than in the deterministic case when it receives a high demand signal $x_i > 0$ (low demand signal $x_i < 0$). By computing first derivatives it is found that A_1 is higher, the higher the correlation among signals, h , and the lower the variance of x_i , σ . This result supports the intuition that firms give more importance to the received signal if they know it is highly correlated with other firms' signals. In this case, quantities are likely to fluctuate more with demand. On the contrary, firms follow more closely the deterministic solution when signals' variances are high, since in this case x_i is not so informative.

2.3.2 Merger

The same framework from last section is now used in an asymmetric context. Mergers are seen as a combination of assets from the merging parties. The relevant asset here is information about the

demand. By enlarging its information set, the merging parties' expectation of market demand is likely to be more accurate. One can think of increased managerial expertise and capacity of interpreting the evolution of demand or of joint efforts made by different market research departments. It is important to clarify that by merging, firms do not establish a single research department but keep the two existing ones and hence get two signals of the state of the demand. Since decision making is improved, it is more likely that quantity supplied meets the realized demand at each price.

The information pooling between the merging parties is introduced in the model in the following way. There are the same n private signals of the pre-merger firms. It is assumed that two firms⁶ decide to merge. Since information sets are the only source of asymmetry, all other $n - 2$ firms are equal even though each one receives a different signal x_j . Let firms 1 and 2 be the merging parties (without loss of generality). The information set of the merging firm includes the two private signals x_1 and x_2 . Let M be the merging firm, its information set is defined by:

$$x_M = \{x_1, x_2\}$$

For the other $n - 2$ firms the information set is given by

$$x_i = \{x_i\} \quad \forall_{i=3, \dots, n}$$

as before.

Given that all random variables are gaussian with mean zero, that the expected value is a linear operator and that private signals are perfectly uncorrelated, one can restrict solutions to the following functional form:

$$\begin{aligned} q_i(x_i) &= B_0 + B_1 x_i, \quad \forall_{i=3, \dots, n} \\ q_M(x_M) &= M_0 + M_1 \frac{x_1 + x_2}{2} \end{aligned}$$

First order conditions are given by

$$\begin{aligned} q_i(x_i) &= \frac{a - b \left[E(q_M(x_M) | x_i) + \sum_{j \neq i} E(q_j(x_j) | x_i) \right] + E(u | x_i)}{2b}, \quad \forall_{i=3, \dots, n} \\ q_M(x_M) &= \frac{a - b \sum_{j=3}^n E(q_j(x_j) | x_M) + E(u | x_M)}{2b}, \end{aligned} \quad (2.2)$$

where

⁶The number of merging parties is irrelevant for the conclusion unless there is a merger to monopoly.

$$E(u \mid x_M) = \frac{1}{n} \left[\frac{\sigma + (n-1)h}{\sigma + m + h} (x_1 + x_2) \right] = V(u) \frac{x_1 + x_2}{\sigma + m + h},$$

$$E(u \mid x_k) = \frac{\sigma + (n-1)h}{n(\sigma + m)} x_k = V(u) \frac{1}{\sigma + m} x_k, \quad \forall k=3, \dots, n \text{ as before,}$$

(see App.4 for a complete derivation). Expected values of rivals quantities are given by:

$$E(q_M(x_M) \mid x_k) = M_0 + M_1 E\left(\frac{x_1 + x_2}{2} \mid x_k\right) = M_0 + M_1 \frac{h}{\sigma + m} x_k$$

$$E(q_j(x_j) \mid x_k) = B_0 + B_1 E(x_j \mid x_k) = B_0 + B_1 \frac{h}{\sigma + m} x_k, \quad \forall k=3, \dots, n$$

$$E(q_j(x_j) \mid x_M) = B_0 + B_1 E(x_j \mid x_1, x_2) = B_0 + B_1 \frac{2h}{\sigma + m + h} \frac{x_1 + x_2}{2},$$

Substituting in (2.2) and solving the system for the constants yields a system of four equations and four unknowns. The unique Nash equilibrium of this game is given by:

$$q_i(x_i) = B_0 + B_1 x_i, \quad \forall i=3, \dots, n \tag{2.3}$$

$$q_M(x_M) = M_0 + M_1 \frac{x_1 + x_2}{2}$$

$$\text{where : } B_0 = M_0 = \frac{a}{bn}$$

$$B_1 = (\sigma + m) \Psi$$

$$M_1 = [2(\sigma + m) - h] \Psi$$

$$\Psi = \frac{V(u)}{b[(\sigma + m + h)[2(\sigma + m) + (n-3)h] - h^2(n-2)]}$$

$$V(u) = \frac{\sigma + (n-1)h}{n}$$

Note again that the constant term in q_i or q_M corresponds to the deterministic quantity competition result with $n-1$ firms. In the standard Cournot model, a merger involving m firms leads to the creation of a single new firm identical to any of the m previous ones. The information exchange agreement between the two merging parties introduces asymmetries in the industry and confers the merged firm M an informational advantage. Quantities are higher than the deterministic ones when a firm get a signal of high demand ($x_i > 0$) and lower than the deterministic ones when ($x_i < 0$). In what concerns the signal's coefficients, one finds B_1 for non merging firms and M_1 for the merging parties. Given that the covariance can not exceed the variance ($\sigma > h$) it is clear that $M_1 > B_1$. This confirms the intuition that by merging and pooling their signals, firms can trust more the observed signals. Hence, in equilibrium it is found that merging parties give more weight to their observed signals than rivals do.

2.3.3 Joint Venture

A joint venture is defined as an agreement between two firms in which the parties share their observed signal x_i but the quantity decision remains independent. One can think of two firms combining the efforts of their market research departments but working independently. It should be clear that the two research departments do not become one but do exchange forecasts and available information. Hence, each firm would observe a specific signal and then share it with the other member of the joint venture. Suppose that firms 1 and 2 reveal completely their respective signals but do not merge.

One could argue that firms might not have incentives to reveal the true signal given that signal x_i is not observed by firm j . If firm i receives a signal of high demand and reveals it to j it reduces the probability that j would underproduce, this harms firm i 's profit. If i receives a signal of low demand and reveals it to j it reduces the probability that j would overproduce and this affects firm i 's profit positively. Hence, firms would like to share truthfully only the low demand signals. However, given that firms know each others' profit functions and the relevant volatility parameters, one can assume that firms share information about future quantities. By reporting q_i (ex-ante) to the partner j , firm j can infer the underlying value of x_i and the truth revelation problem vanishes because q_i is observable.

It is easy to show that equilibrium in this case is given by the following quantities for the firms which do not take part in the information agreement ($i = 3, \dots, n$) and for the partners ($i = 1, 2$):

$$\begin{aligned} q_i(x_i) &= A_0 + \tilde{B}_1 x_i, & \forall i=3, \dots, n \\ q_i(x_M) &= A_0 + \tilde{M}_1 \frac{x_1 + x_2}{2}, & \forall i=1, 2 \end{aligned} \quad (2.4)$$

where :

$$\begin{aligned} A_0 &= \frac{a}{b(n+1)} \\ \tilde{B}_1 &= [3(\sigma + m) - h] \Phi \\ \tilde{M}_1 &= 2[2(\sigma + m) - h] \Phi \\ \Phi &= \frac{V(u)}{b} \left[\frac{1}{3(\sigma + m + h)[2(\sigma + m) + (n-3)h] - 4h^2(n-2)} \right] \end{aligned}$$

Note that the deterministic term A_0 doesn't change. Given that quantity decision remains independent among the joint venture parties, the number of operating firms in the market does not change under the *Cournot* framework.

2.4 Nash Equilibria

In order to analyze the effects of a merger or a joint venture, four relevant cases were chosen and studied in detail. Recall that h is the covariance among signals, m is a measure of precision of observed signals and σ is the volatility of the demand function. The studied cases are the following:

- $h = 0$ (non correlated signals)
- $h = \sigma$ and $m = 0$ (maximum correlation and highly informative signals)⁷
- $h = \frac{\sigma}{2}$ and $m = 0$ (moderate correlation and highly informative signals)
- $h = \frac{\sigma}{2}$ and $m = \sigma$ (moderate correlation and non informative signals)

The equilibrium coefficients of each of the above cases are given by:

⁷It is interesting to note that this second case will reproduce the pre equilibrium for the case of joint ventures. All firms and consumers remain indifferent between the symmetric equilibrium and the joint venture equilibrium.

CASE 1: Non-Correlated Signals		
	Merger	Joint Venture
$h = 0$	$V(u) = \frac{\sigma}{n},$	
	$A_1 = \frac{\sigma}{2bn(\sigma+m)}$	
	$B_1 = A_1$	$\tilde{B}_1 = A_1$
	$M_1 = 2A_1$	$\tilde{M}_1 = \frac{3}{4}A_1$
CASE 2: Maximum correlation and highly informative signals		
	Merger	Joint Venture
$h = \sigma$ $m = 0$	$V(u) = \sigma$	
	$A_1 = \frac{1}{b(n+1)}$	
	$B_1 = \frac{1}{bn}$	$\tilde{B}_1 = A_1$
	$M_1 = B_1$	$\tilde{M}_1 = A_1$
CASE 3: Moderate correlation and highly informative signals		
	Merger	Joint Venture
$h = \frac{\sigma}{2}$ $m = 0$	$V(u) = \frac{\sigma(n+1)}{2n}$	
	$A_1 = \frac{n+1}{bn(n+3)}$	
	$B_1 = \frac{2(n+1)}{bn(2n+5)}$	$\tilde{B}_1 = \frac{5(n+1)}{bn(5n+17)}$
	$M_1 = \frac{3(n+1)}{bn(2n+5)}$	$\tilde{M}_1 = \frac{6(n+1)}{bn(5n+17)}$
CASE 4: Moderate correlation and non informative signals		
	Merger	Joint Venture
$h = \frac{\sigma}{2}$ $m = \sigma$	$V(u) = \frac{\sigma(n+1)}{2n}$	
	$A_1 = \frac{n+1}{bn(n+7)}$	
	$B_1 = \frac{4(n+1)}{bn(4n+27)}$	$\tilde{B}_1 = \frac{11(n+1)}{bn(11n+83)}$
	$M_1 = \frac{6(n+1)}{bn(4n+27)}$	$\tilde{M}_1 = \frac{14(n+1)}{bn(11n+83)}$

These four tables summarize the coefficients of the equilibrium quantities both for the case of a merger and for the case of a joint venture that will be used in the following sections dealing with the welfare analysis. Note that in all cases one finds $M_1 \geq B_1$ corroborating the idea that by pooling signals about uncertain demand, firms can trust more the observed signals and hence give them more weight in the optimal quantity decision.

In the next section it will be shown that no merger or joint venture would occur when correlation between the signals is zero (case 1). To understand the no correlation case, recall the problem of

information exchange under quantity competition. A firm is always better off by increasing the information it gets to see alone. However, sharing implies that the other participant in the information agreement also sees its information set improved. The two firms will maximize profits based on the same information set and this increases output's correlation. With no correlation between signals ($h = 0$) nothing can be inferred about the state of demand and only the negative effect is relevant. By sharing a low demand signal a firm reduces the likelihood that the information agreement party overproduces which increases expected profits. By sharing a high demand signal a firm reduces the likelihood that the party underproduces which decreases expected profits. Later it is explained why the second effects dominates.

It is interesting to note that the joint venture case with maximum correlation and perfect observability (case 2) corresponds to the benchmark case were there is no information exchange: $\tilde{B}_1 = \tilde{M}_1 = A_1$. Given that the number of firms in the market does not change and that correlation between signals is the highest possible (equal to the variance), the firms taking part on the information agreement can infer about market demand as much as those who do not share information. In other words, there are no gains from pooling information since by observing its specific signal, a firm can already infer much about the rivals' behavior. For the firms, the main drawback of information sharing is the increase in output correlation that is already at its maximum level in this case 2. Hence, we will show that both firms and consumers are indifferent between a joint venture and a 'no information sharing' equilibrium.

The third and fourth cases assume moderate covariance ($h = \frac{\sigma}{2}$) that allows for some inference about the state of demand (which has a positive effect on firms' profits) without increasing too much output's correlation (which harms firms' profits). The parameter at stake is the accuracy of the signals (m). Case 3 allows for the described situation in which output correlation is not extremely high and there is room for gains by sharing signals. Case 4, by imposing a high variance in the signals due to the volatility of the noise term (e_i), reproduces the situation in Case 1 in which signals are not informative.

2.5 Expected Profits

We now check whether it is profitable to merge under *Cournot*, i.e., we verify under which conditions the following expression holds:

$$E(\Pi_M | x_M) \geq E(\Pi_1 | x_1) + E(\Pi_2 | x_2).$$

Conditional expected profit is given by

$$\begin{aligned} E(\Pi_i | x_i) &= E \left[\left(a - b \sum_{j \neq i} q_j(x_j) - bq_i(x_i) + u \right) q_i(x_i) \mid x_i \right] = \\ &= \left[E \left(a - b \sum_{j \neq i} q_j(x_j) + u \mid x_i \right) - bq_i(x_i) \right] q_i(x_i). \end{aligned}$$

Recall that from the first order conditions

$$2bq_i(x_i) = E \left[a - b \sum_{j \neq i} q_j(x_j) + u \mid x_i \right],$$

hence, conditional expected profits can be written as

$$E(\Pi_i | x_i) = (2bq_i(x_i) - q_i(x_i)) q_i(x_i) = bq_i(x_i)^2.$$

Therefore, unconditional expected profits are given by

$$E(\Pi_i) = bE(q_i(x_i)^2),$$

where $q_i(x_i)$ is defined in (2.1). After computing the cross products and operating unconditional expected value on the random variables the following expression is obtained:

$$E(\Pi_i) = b [A_0^2 + A_1^2(\sigma + m)] \quad \forall_{i=1, \dots, n}$$

for the symmetric equilibrium stage. The same exercise within the asymmetric equilibria (2.3) and (2.4) yields the expected profits of the merger and the 'outsiders' after the merger took place:

$$\begin{aligned} E(\Pi_M) &= b \left[B_0^2 + M_1^2 \frac{(\sigma + m + h)}{2} \right] \\ E(\Pi_i)_{post}^M &= b [B_0^2 + B_1^2(\sigma + m)], \end{aligned}$$

and for the two joint venture parties and the $n-2$ non participants after the joint venture is established:

$$\begin{aligned} E(\Pi_{JV}) &= b \left[A_0^2 + \tilde{M}_1^2 \frac{(\sigma + m + h)}{2} \right] \\ E(\Pi_i)_{post}^{JV} &= b [A_0^2 + \tilde{B}_1^2(\sigma + m)]. \end{aligned}$$

After determining expected profits of all market participants we now compare firms' performance before and after the information agreements.

2.5.1 Impact on participant firms' profits

The analysis of the impacts of a merger or joint venture between two arbitrary firms enables us to conclude about the incentives to merge or share information and which equilibria might occur. It is shown that a merger is off the equilibrium path and a joint venture is likely to occur for specific values of the parameters.

Merger

A merger is profitable if the following condition holds,

$$E(\Pi_M) > 2E(\Pi_i) \Leftrightarrow \quad (2.5)$$

$$B_0^2 + M_1^2 \frac{(\sigma + m + h)}{2} > 2[A_0^2 + A_1^2(\sigma + m)].$$

Recall that in the deterministic case, a merger is never profitable unless firms merge to a monopoly⁸, (see discussion on Salant Switzer and Reynolds in section 2). In the non deterministic case, what is computed is not firms' profits but firms' expected profits conditioned on their respective information sets. Analogously, the term $B_0^2 - 2A_0^2$ is always negative (except for $n = 2$) and is independent of the volatility parameters.

Computing the difference for the four relevant cases, it is easy to find that, like in the deterministic case, there are no incentives to merge unless firms merge to monopoly. The information gain, given by the term $M_1^2 \frac{(\sigma + m + h)}{2} - 2A_1^2(\sigma + m)$ is as well negative (except for case 3) and hence reinforces the standard result of the *Cournot* model. In case (3) for $n = 3$ and very high demand volatility σ this informational term is positive but not strong enough to compensate the deterministic (negative) part. A merger to monopoly is always profitable since both effects are positive.

Joint Venture

A joint venture is profitable if the following condition holds,

$$\widetilde{M}_1^2 \frac{(\sigma + m + h)}{2} - 2A_1^2(\sigma + m) > 0.$$

Recall that contrary to the merger case, quantity decision remains independent and consequently the number of firms does not change. Hence, the deterministic term in the expression above is equal to zero. Despite the fact that the deterministic term is zero, only in the third case ($h = \frac{\sigma}{2}, m = 0$) do we conclude that there are incentives to form a joint venture. In the other situations the parties are either indifferent or made worse off.

Results for $n > 2$ can be summarized in the table below:

⁸In the deterministic case with n symmetric firms and zero production costs, a merger between two firms leads to the following pre and post merger profits: $\Pi_{(n)} = \left(\frac{\alpha}{n+1}\right)^2 \frac{1}{b}$ and $\Pi_{(n-1)} = \left(\frac{\alpha}{n}\right)^2 \frac{1}{b}$. Now, $2\Pi_{(n)} > \Pi_{(n-1)}$ for $n \neq 2$ and thus, the joint profits of the merging firms is smaller than the sum of their profits prior to merger.

	Merger	Joint Venture
$h = 0$	-	-
$h = \sigma, m = 0$	-	0
$h = \frac{\sigma}{2}, m = 0$	-	+ (1)
$h = \frac{\sigma}{2}, m = \sigma$	-	- (2)

(1) the effect is positive the higher the number of firms and for $n \geq 7$.

(2) the effect is positive for an unreasonably high number of firms ($n \geq 81$).

Comments

The results report the effects of a merger or joint venture in insider firms' expected profits, when the random variables follow a normal distribution.

Recall the intuition for the industry-wide information sharing agreement. When a firm reveals a signal of low demand ($x_i < 0$), it firm reduces the likelihood that competitors overproduce which affects positively expected profits. When a firm reveals a signal of high demand ($x_i > 0$), it reduces the likelihood that competitors underproduce which harms expected profits. Because x is a gaussian random variable, the two situations occur with the same probability. The fact that prices are higher in high demand periods than in low demand ones implies that losses from sharing high demand signals are not compensated by gains from sharing low demand signals.

In the asymmetric case, when information exchanges occur only between two firms through a merger or a joint venture, rivals' information set remains unchanged. Hence, this negative effect due to an increase in market correlation, is likely to be weaker.

It has been widely criticized the use of the *Cournot* settings in modelling mergers (see section 2). This is because when firms compete *à la Cournot* and there is a merger between m firms, this corresponds to an exit of $m - 1$ firms of the market. Given that there are no capital, capacities or other assets to accumulate by merging, the m merging firms are replaced by one with the same size of each one of the merging participants. This effect that characterizes the deterministic case is emphasized in the stochastic one. Hence, not only the deterministic term⁹ in expression (2.5) is always negative (for $n > 2$) but also the one dependent on the volatility values. This is due to the higher market correlation effect explained above but also due to the fact that one must compare two *pre* merger coefficients (A_1) with one *post* merger coefficient (M_1). Moreover, expected profits depend positively on demand volatility and by pooling information, firms are able to reduce expected demand variance.

The joint venture case is more interesting to look at since it isolates the information effect. There are no changes in the number of firms, therefore, the difference between expected profits in the *pre*

⁹By deterministic term it is meant the one that is independent from the volatility parameters σ , m , and h .

and *post* joint venture situation is a function of the volatility parameters exclusively. The joint venture parties face the following trade-off. By sharing private signals they increase independently the number of observations they get to see alone relatively to rival firms, this affects positively expected profits. However, by sharing those signals the joint venture parties increase their output correlation, and this affects negatively expected profits. Notice that in the first two cases (minimum and maximum covariance h) there are no incentives to form a joint venture. The first case is the one of no correlation between signals. Here, little can be inferred about the demand behavior from the information exchange. Hence, the negative effect of increasing output correlation of the two firms forming a joint venture dominates the positive effect of enlarging the information set of each of them. The second case is the one of maximum correlation with perfect observability. Here, correlation is already so high that the relative advantage of the joint venture parties relative to rivals is negligible. By pooling the signals, the negative effect dominates again. In the last (fourth) case, signals are not highly informative, even though correlation is moderate, due to the high variance of the observed signal. This reproduces the situation generated by case one since it is as if signals were uncorrelated. With such a lack of precision, little can be inferred about the stochastic demand. Only in the third case of moderate correlation and highly informative signals can the joint venture be profitable. Correlation is small enough to allow for a relative advantage of the information sharing firms but not 'too' small so that some valuable inference about the market demand is possible. Perfect observability stresses this value of the increase of the information that the joint venture parties get to see alone. In this case, benefits more than compensate losses from sharing privately observed signals and therefore, there are incentives to form a joint venture.

Industry-wide information sharing

In this section we analyse the case of a industry-wide exchange of information. The classical papers on information exchange focus on agreements involving all firms in the market whereas our work focuses only partial agreements through a merger or a joint venture. In the literature, (see Gal-Or, 1985 and Vives, 1984), firms receive a signal from nature and report them to an agency. This agency, that can be seen as a trade association, redistributes the signals throughout the industry. Firms have two decision variables: quantity or price and the accuracy of the signal reported to the agency (variance s). A main result of the literature is that there are no incentives to exchange information in quantity setting games.

To show that our model concurs with such a result, we compute the equilibrium under information sharing between all firms. We assume that firms truthfully reveal their signals to the agency that then redistributes them to all firms (i.e., maximal accuracy of the reported signals, $s = 0$).

When there is no information sharing equilibrium coefficients and expected profit are given by:

$$\begin{aligned} A_0 &= \frac{a}{b(n+1)} \\ A_1 &= \frac{\sigma}{2bn(\sigma+m)} \\ E(\Pi_i)_{pre} &= b[A_0^2 + A_1^2(\sigma+m)] \end{aligned}$$

If there is information sharing among n firms, it is easy to check that,

$$\begin{aligned} E(u \mid x_i) &= \frac{\sigma}{n(\sigma+m)} \sum_{i=1}^n x_i, \\ E(q_i \mid x_i) &= A_0 + A_1 \sum_{i=1}^n x_i. \end{aligned}$$

Hence, the industry wide information sharing coefficients and expected profit are given by:

$$\begin{aligned} A_0 &= \frac{a}{b(n+1)} \\ A'_1 &= \frac{\sigma}{bn(n+1)(\sigma+m)} \\ E(\Pi_i)_{post} &= b[A_0^2 + A_1'^2 n(\sigma+m)] \end{aligned}$$

Comparing the two expected profits one finds that

$$E(\Pi_i)_{pre} > E(\Pi_i)_{post} \Leftrightarrow (n-1)^2 > 0$$

for all values of the parameters. Therefore, the model presented here confirms the literature common result that there are no incentives for market information sharing in a *Cournot* environment.

2.5.2 Impact on rivals' profits

It can be shown that a merger makes rivals better off while the joint venture makes rivals either indifferent or worse off. The following expressions refer to rivals' profits before and after the information agreements take place.

Merger

The condition under which the merger doesn't harm rivals is given by,

$$\begin{aligned} E(\Pi_i)_{post}^M - E(\Pi_i)_{pre} &\geq 0 \Leftrightarrow \\ &\Leftrightarrow (B_0^2 - A_0^2) + (\sigma+m)(B_1^2 - A_1^2) \end{aligned}$$

The difference $(B_0^2 - A_0^2)$ is always positive and corresponds to the deterministic case analyzed by Salant, Switzer and Reynolds¹⁰. Since $B_1 \geq A_1$ for any values of the parameters, one can conclude

¹⁰ $B_0^2 - A_0^2 = \left(\frac{a}{bn(n+1)}\right)(2n+1) > 0$

that non merging firms increase expected profits with the merger (zero correlation corresponds to the case $B_1 = A_1$).

Hence, the deterministic result is reinforced: firms that do not participate in the merger (outsiders) are always better off and for even more than in the deterministic case. The higher the volatility of demand and the higher the correlation coefficient, the higher the expected profits of an outsider.

Joint Venture

The condition under which the joint venture doesn't harm rivals is given by,

$$E(\Pi_i)_{post}^{JV} - E(\Pi_i)_{pre} \geq 0 \Leftrightarrow (\sigma + m) (\tilde{B}_1^2 - A_1^2)$$

The results can be summarized in the following table,

	Merger	Joint Venture
$h = 0$	+	0
$h = \sigma, m = 0$	+	0
$h = \frac{\sigma}{2}, m = 0$	+	-
$h = \frac{\sigma}{2}, m = \sigma$	+	-

2.6 Expected consumer surplus

Given this simple setting, impact of information sharing agreements on the consumer surplus can be assessed by checking the impact on total output. Since the demand function is linear, if total output increases/decreases, prices are necessarily lower/higher. Consumer surplus increases with lower prices and decreases with higher prices. Impact on the expected consumer surplus is evaluated comparing the expected value of squared total quantities produced in the *pre* and *post* merger situation.

$$E(Q^2)_{post} \geq E(Q^2)_{pre}$$

where

$$\begin{aligned} E(Q^2)_{pre} &= n^2 A_0^2 + n A_1^2 [\sigma + m + (n-1)h] \\ E(Q^2)_{post} &= (n-1)^2 B_0^2 + (n-2) B_1^2 [\sigma + m + (n-3)h] + \\ &\quad + 2B_1 M_1 h (n-2) + M_1^2 \left[\frac{\sigma + m + h}{2} \right]. \end{aligned}$$

2.6.1 Impact on consumer surplus

By computing the difference between the expected squared quantities in the *pre* and *post* situations, one finds opposite results. If a merger occurs, it leaves consumers always worse off. On the contrary, if a joint venture takes place, consumers can never be worse off, they either remain indifferent or are made better off. The findings can be summarized in the table below:

	Merger	Joint Venture
$h = 0$	-	+
$h = \sigma, m = 0$	-	0
$h = \frac{\sigma}{2}, m = 0$	-	+
$h = \frac{\sigma}{2}, m = \sigma$	-	+

A merger can never increase consumer surplus since it reduces the number of firms in the market and therefore the aggregate output. The information sharing effect reinforces this collusive effect. A larger information set of the merging firms is not only insufficient to make aggregate output fluctuate with demand but yet it has the opposite effect. Since the merging parties act now as a single unit, they use information strategically and reduce quantity more in low demand periods than increase it in high demand ones.

A joint venture is always beneficial to consumers because it increases general information about the market demand. Given that the joint venture parties choose quantities independently, no collusive effect takes place. The information exchange implies that in expected value, total output matches better the market demand. This means that in expectations, total output increases in high demand periods leading to a low expected price and total output decreases in low demand periods leading to a higher price. In other words, output fluctuates more with the uncertain demand. Figure 1 illustrates this intuition. Let D , Q , and P denote Demand, aggregate Quantity and Price respectively. Subscripts L, H and E stand for low, high and expected respectively. If realized demand is high, and quantity adjusted to Q_H there is a gain in consumer surplus. Without adjustment the price charged would be P_H and due to the increase in supply it is lower than this value. If realized demand is low and quantity adjusted to Q_L there is a loss in consumer surplus since the price charged will be higher than P_L . For symmetric shifts (recall that random variables are gaussian), given that gains in consumer surplus (marked with +) are larger than losses (marked with -) in expected value, the overall effect is positive¹¹.

¹¹This graphical analysis is performed for the case of a monopoly with stochastic demand by Khun and Vives

Figure 1- Impact on Consumer Surplus (intuition)

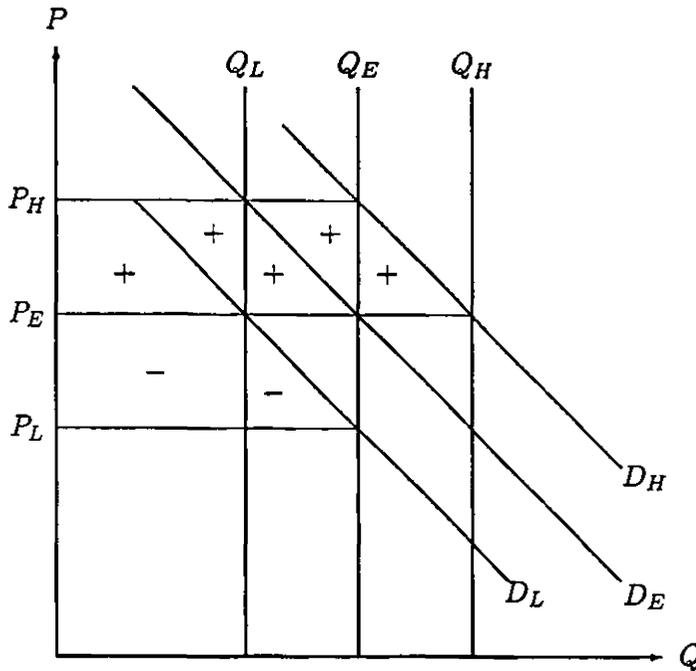


Figure 1: With a joint venture, total output tends to fluctuate more with market demand. (*L* stands for low, *H* stands for high and *E* stands for the expected level with no extra information). Consumer surplus increases in expected value.

2.7 Expected total surplus

The ambiguous case to evaluate in terms of total welfare effects is the one of a joint venture with moderate correlation and highly informative signals, (case number 3: $h = \frac{\sigma}{2}$, $m = 0$). In this setting, there is an incentive to form a joint venture, the expected consumer surplus increases, but expected rivals' profits decrease. One can go back to the original utility function to build a measure of expected total surplus.

Assume that demand function $p = a - bQ + u$ is the result of consumers maximizing in q_i the following utility function:

$$U(Q) = (a + u)Q - \frac{b}{2}Q^2, \quad Q = \sum_{i=1}^n q_i$$

subject to the budget constraint $p \sum_{i=1}^n q_i = M$. Consumer surplus corresponds to the value of the utility function in equilibrium net of expenses $p \sum_{i=1}^n q_i$. Total surplus is defined as the sum of

consumer surplus with total profits. The assumption that firms produce at no cost enables one to look only at $U(Q)$ given that the other two terms cancel out. In an uncertainty context, one needs to take expectations on $U(Q)$. We look at the two components of $U(Q)$ separately:

$$E \left[(a + u) \sum_{i=1}^n q_i \right] = anA_0 + nA_1V(u)$$

(recall that a is deterministic but u is stochastic), and

$$E \left[\left(\sum_{i=1}^n q_i \right)^2 \right] = n^2A_0^2 + nA_1^2[(\sigma + m) + (n - 1)h].$$

Hence, expected total surplus before information agreements ETS_{pre} , can be expressed in the following way:

$$ETS_{pre} = anA_0 + nA_1V(u) - \frac{b}{2} [n^2A_0^2 + nA_1^2(\sigma + m + (n - 1)h)].$$

Analogously, in the post joint venture case, this expression is given by:

$$ETS_{post}^{JV} = anA_0 + (n - 2)\tilde{B}_1V(u) + 2\tilde{M}_1V(u) - \frac{b}{2} [n^2A_0^2 + (n - 2)\tilde{B}_1^2(\sigma + m + (n - 3)h) + 4(n - 2)\tilde{B}_1\tilde{M}_1h + 2\tilde{M}_1^2(\sigma + m + h)].$$

It is straightforward to check that,

$$ETS_{post}^{JV} - ETS_{pre} > 0 \quad \text{for } h = \frac{\sigma}{2}, m = 0, \forall \sigma$$

and the difference is higher, the higher the volatility of the demand function, σ . Recall that for a sufficiently high number of participating firms (in the limit all), the joint venture is no longer profitable¹². Therefore we can conclude that, if there are incentives to form a joint venture, then it will necessarily increase expected total surplus.

2.8 Bertrand Competition

One must be aware of the possible lack of robustness of results when the market structure changes from *Cournot* to *Bertrand*, especially when results in the existing literature depend heavily on the strategic variable. In the works on information exchange, (see Vives, Gal-Or, Novshek and Sonnenschein, Clark), the focus is on the incentives to share information. Firms can report their signals to an agency

¹²This is the common result of the literature on information exchange: an industry-wide agreement is not profitable.

and choose the accuracy (s) of this transmission, i.e., the parameter s is a decision variable. The authors conclude that in *Cournot*, firms' optimal strategy is not to share information ($s = \infty$). This means that firms report a non informative signal. In the *Bertrand* case, the referred studies cover only a market with two firms. The conclusion is exactly the opposite: the optimal strategy for both firms is to reveal the true signal and hence to choose $s = 0$. Analogously, with unknown common demand, Vives (1984) demonstrates that sharing is a dominant strategy with *Bertrand* competition and concealing is a dominant strategy with *Cournot* competition.

Deneckere and Davidson (1985) analyze the incentives to merge in the *Bertrand* model. They find the (now standard) result that there are always incentives to merge. This is due to the fact that the merging parties push the price upwards and this effect is reinforced by all rivals since prices are strategic complements. The overall effect increases all firms profits. On the contrary, in the *Cournot* setting (see Salant, Switzer and Reynolds), the merging parties reduce their quantity and this effect is opposed by the rivals' response that expand their output (quantities are strategic substitutes). As a result, the merging parties are worse off than in the pre merger situation and therefore there are no incentives to merge.

Given these two results, one would expect that, in a *Bertrand* environment with uncertainty, merger profitability would be restored. However, given the strategic complementarity and the upward sloping best response functions, prices are not likely to decrease. If in the *Cournot* setting, the rivals' output expansion is not enough to enhance consumers' welfare, it is not likely that this would happen when the market compete *à la Bertrand*. Vives (1990) shows that information sharing is good from the social point of view under *Cournot* but bad under *Bertrand*.

Hence, one can conclude that in the presence of *Bertrand* competition, efficiency gains through pooling information would reinforce the profitability of a merger but would produce a larger reduction in the consumer surplus than in the presence of *Cournot* competition. In other words, the desired effect of merger profitability and price reduction due to information sharing are less likely to occur under *Bertrand* than under *Cournot*. In what concerns joint ventures one can conclude directly from the existing literature. This is because the way information sharing is described in the literature can be mapped to the way joint venture on market research departments is defined in this work. The departure from the existing literature consists of modelling a partial agreement between two firms and not an industry-wide agreement. We can say that for the case of a joint venture under price competition, firms would have indeed incentives to share private information but this would reduce consumer surplus and as a result, would not be socially desirable.

2.9 Future Research

The main drawback from our analysis stems from the fact that results are not derived for all range of parameters but only for four chosen subsets of parameters (m, h) . Generalisation is particularly difficult for the case of the joint venture. It is our intention to generalise the results, possibly by simplifying the information structure. The fact that negative prices can occur in equilibrium, even though with an arbitrarily small probability, is also troublesome. This inconvenient event can be easily corrected by dropping normality and assuming a uniform distribution for the random variables. Surprisingly, papers cited throughout this chapter typically assume normal distributions on the demand or cost functions and show no concern for this problem.

It would be very interesting to endogenise two variables that would improve the generality of results: the number of firms involved in the information exchange agreement and the number of information agreements in the whole industry. This paper focuses only on a single partial agreement (merger or joint venture) between two firms. Nothing is said about agreements between more than two firms and nothing is said about pair-wise agreements involving a high number or even all firms in the industry.

In what concerns the merger case, these points are irrelevant. It is shown that a merger is never profitable unless firms merge to monopoly. Hence, the number of merging firms is not restricting the generality of the conclusions. More merging parties would lead to a more significant reduction of the number of firms in the market but would not restore the profitability of the merger unless all industry merges to one firm. On the other hand, the number of mergers is as well not restricting the generality of the conclusions. Note that by allowing for many pairs of merging firms the same effect arises: a reduction of the number of firms in the market to $(\frac{n}{2})$. This does not affect conclusions since the market does not become a monopoly.

In what concerns the joint venture case one has to be more careful. From the literature on information sharing we can infer that the higher the number of participants in the information agreement, the less likely that the joint venture is profitable. This means that if the model allowed for an increasing number of joint venture participants, at a certain point, the joint venture would no longer be profitable. The other crucial point consists of considering many pair-wise information agreements among firms in the industry. Given the findings of the existing literature, we would tend to say that again, the increasing correlation between firms' outputs would make these agreements less and less profitable. However, it should be checked whether there is a range of agreements under which joint ventures are profitable but total welfare decreases. There may arise socially undesirable joint ventures whereas in our model, this information agreement done by only two firms, always increases expected total surplus.

It is as well interesting to assess robustness of the results with respect to the information technology.

We intend to perform the analysis assuming other specifications for the information problem and check whether merger profitability can be achieved under quantity competition. If merger profitability is obtained, information sharing can then be interpreted as an efficiency gain that improves relative performance of the merging parties and can even benefit consumers through lower equilibrium prices.

2.10 Conclusion

This chapter studies the effect of an information agreement between two firms, in a *Cournot* market with uncertain demand, on the welfare of all market participants. Firms receive a signal about the state of the market demand and may share it by forming a merger or a joint venture. Both cases are studied independently and compared to the no information sharing case. In a joint venture, quantity decisions remain independent for the participating firms whereas in the merger case, the merging parties act as a single firm. The analysis is carried out for four sets of parameters that include the interesting scenarios. The main findings of this work are the following: on the one hand, a merger is never profitable nor welfare increasing; on the other hand, a joint venture can be profitable and is not welfare decreasing. Profitability is achieved with high signal precision and moderate correlation between firms' signals. When a joint venture is profitable, it is also welfare increasing: consumers are better off and competitors are worse off but the overall effect is positive. Our work makes a contribution to the literature on information exchange in two directions. First, it introduces market asymmetries by allowing for a merger or a joint venture between two firms. Second, it allows for non negative correlation among signals whereas the referred literature typically assumes this correlation to be zero. The existing literature demonstrates that industry-wide exchange information agreements under *Cournot* competition with identical firms are unprofitable. This paper shows that partial agreements can be profitable. Moreover, if these agreements take place, they are necessarily welfare increasing: the pooling of information not only increases parties profits but also reduces the expected price. This is due to an improvement in production efficiency, i.e., aggregate output fluctuates more according to the uncertain demand. Given that the joint venture parties share private information, they have, in expected value, a more accurate output decision.

Results might bring some insights regarding competition policy. If *Cournot* competition is a good description of an industry behavior and demand is highly volatile, antitrust authorities should be tolerant about a single market research agreement (as joint ventures are defined in this paper) since it is likely to benefit firms and consumers. In what concerns a merger, even though it is not profitable and hence not likely to occur in this context, authorities should be aware that it can not be welfare increasing.

2.11 Bibliography

Clarke, Richard, 1983, "Collusion and the incentives for information sharing", *Bell Journal of Economics*, 14.

Deneckere, Raymond and Carl Davidson, 1985, "Incentives to form coalitions with Bertrand Competition", *Rand Journal of Economics*, 16.

Farrell, Joseph and Carl Shapiro, 1990, "Horizontal Mergers: an equilibrium analysis", *American Economic Review*, 80, No.1.

Gal-Or, Esther, 1985, "Information sharing in oligopoly", *Econometrica*, 53, No.2.

Gal-Or, Esther, 1986, "Information transmission-Cournot and Bertrand equilibria", *Review of Economics Studies* LIII.

Kühn, Kai-Uwe and Xavier Vives, 1994, "Information Exchanges Among Firms and their Impact on Competition", *Institut d'Anàlisi Econòmica (CSIC), Barcelona*.

Novshek, William and Hugo Sonnenschein, 1982, "Fulfilled expectations Cournot duopoly with information acquisition and release", *Bell Journal of Economics*, 13.

Raith, Michael, 1996, "A General Model of Information Sharing in Oligopoly", *Journal of Economic Theory*, 71.

Salant, Stephen, Sheldon Switzer and Robert Reynolds, 1983, "Losses from Horizontal Mergers: the effects of an exogenous change in industry structure on *Cournot-Nash* equilibrium", *Quarterly Journal of Economics*, 98.

Vives, Xavier, 1984, "Duopoly Information Equilibrium: Cournot and Bertrand", *Journal of Economic Theory*, 34.

Vives, Xavier, 1990, "Trade association disclosure rules, incentives to share information, and welfare", *Rand Journal Of Economics*, 21.

2.12 Appendix

A. 1 - Moments of random variable u :

The stochastic variable u is defined as:

$$u = \frac{\sum_{i=1}^n u_i}{n},$$

unconditional expected value and variance of the random variable u can be easily computed as follows,

$$E(u) = \frac{1}{n} \sum_{i=1}^n E(u_i) = 0$$

$$V(u) = E[u^2] - [E(u)]^2 = E\left[\left(\frac{u_1 + u_2 + \dots + u_n}{n}\right)^2\right] = \frac{\sigma}{n} + \frac{(n-1)h}{n}.$$

A. 2 - Maximization Problem (benchmark case):

Firm i maximizes Π^i taking $q_j(x_j)$ as given:

$$\begin{aligned} \frac{\partial \Pi^i}{\partial q_i} &= a - 2bq_i(x_i) - b \sum_{j \neq i} E(q_j(x_j) | x_i) + E(u | x_i) = 0 \\ q_i(x_i) &= \frac{a - b \sum_{j \neq i} E(q_j(x_j) | x_i) + E(u | x_i)}{2b}, \quad \forall i=1, \dots, n. \end{aligned}$$

The second order condition is satisfied: $\frac{\partial^2 \Pi^i}{\partial q_i^2} = -2b < 0$.

A. 3 - Conditional expected value of u in the symmetric equilibrium:

$$\begin{aligned} E(u_i | x_i) &= E(u_i) + \frac{\text{Cov}(u_i, x_i)}{\text{Var}(x_i)} (x_i - E(x_i)) \\ E(u_i) &= E(x_i) = 0 \\ \text{Cov}(u_i, x_i) &= E(u_i \times (u_i + e_i)) = \sigma \\ \text{Var}(x_i) &= E((u_i + e_i)^2) = \sigma + m \\ E(u_i | x_i) &= \frac{\sigma}{\sigma + m} x_i \end{aligned}$$

Analogously, for $\forall j \neq i$

$$\begin{aligned} E(u_j | x_i) &= E(u_j) + \frac{\text{Cov}(u_j, x_i)}{\text{Var}(x_i)} (x_i - E(x_i)) \\ E(u_j) &= 0 \\ \text{Cov}(u_j, x_i) &= E(u_j \times (u_i + e_i)) = h \\ E(u_j | x_i) &= \frac{h}{\sigma + m} x_i \end{aligned}$$

Hence,

$$E(u | x_i) = \frac{1}{n} \left[\frac{\sigma}{\sigma + m} x_i + \frac{(n-1)h}{\sigma + m} x_i \right] = \frac{x_i}{\sigma + m} V(u).$$

Analogously for x_j ,

$$\begin{aligned} E(x_j | x_i) &= E(x_j) + \frac{\text{Cov}(x_j, x_i)}{\text{Var}(x_i)} (x_i - E(x_i)) \\ \text{Cov}(x_j, x_i) &= E((u_j + e_j) \times (u_i + e_i)) = h \\ E(x_j | x_i) &= \frac{h}{\sigma + m} x_i. \end{aligned}$$

A. 4 - Conditional expected value of u in the asymmetric equilibrium:

$$E(u_i | x_1, x_2) = E(u_i) + \left[\text{cov}(x_1, u_i) \quad \text{cov}(x_2, u_i) \right] \Omega_{xx}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \forall i=1, \dots, n$$

$$\text{where } \Omega_{xx}^{-1} = \frac{1}{(\sigma + m)^2 - h^2} \begin{bmatrix} \sigma + m & -h \\ -h & \sigma + m \end{bmatrix}$$

hence,

$$\begin{aligned} E(u_i | x_1, x_2) &= \frac{1}{(\sigma + m)^2 - h^2} [(\sigma(\sigma + m) - h^2)x_i + (h(\sigma + m) - h\sigma)x_j], \\ \text{for } i, j &= 1, 2 \text{ and } i \neq j \\ E(u_k | x_1, x_2) &= \frac{h}{\sigma + m + h} (x_1 + x_2). \end{aligned}$$

Chapter 3

A Herding Approach to Merger Waves

Abstract: This chapter uses a variant of Banerjee's herding model to explain two features of merger activity that appear unrelated: (i) mergers tend to occur in waves and (ii) profits often decline after merging takes place. Herding occurs when a firm follows the behavior of the preceding decision maker ignoring its own information. The model explains why it can be rational for firms to imitate other firms' decision to merge, despite holding private information that supports a non-merging strategy. Surprisingly, although occurring with lower probability, a merger wave 'starts earlier' under 'pessimistic' priors than under 'optimistic' beliefs concerning the post-merger value of the firms.

3.1 Introduction

One remarkable stylized fact about mergers is that they seem to occur in waves and despite many previous studies this phenomenon is still not well understood¹. The merger wave which started in the beginning of the nineties is called the fifth merger wave². The first and second merger waves took place in the United States between 1887-1904 and 1916-1929 respectively. A third post World War II merger wave is noticeable also in Europe and Canada and covers roughly the period from 1948 to the mid seventies. The fourth wave started in the early eighties and like the current one, was more tempered in the US but outstanding in Europe³. Given such merger activity, one might expect that the merging strategy would significantly increase the value of the firms and improve their market performance.

¹In this paper, when referring to mergers it is meant both mergers and takeovers. Both transactions consist of a legal transformation through which two or more formerly independent firms come under common control. While in the case of takeovers there is only a transfer in control, in the case of mergers there is a complete integration of independently run firms.

²The Economist, January 27th 2001, pp. 67.

³See Scherer for a complete description of historical merger waves.

Surprisingly, specially the third and fourth waves provide evidence that ex-post merger performance can be very poor and that many mergers deals do reduce the value of the firms (see section 1.1).

This chapter attempts to explain these two features of merger activity. It suggests a reason for both why mergers occur in waves and why it may well be rational for firms to choose (ex-post) non profitable merging deals. The basis of the whole argument is herding behavior among firms. Herding occurs when agents choose to ignore their private information and mimic actions of agents who acted previously. For simplicity, in the model there are two states of the world that determine merger profitability. The states of the world can be favorable or adverse to merger activity. This hypothesis tries to capture merger specific factors like the so called "cultural cost of merging". Specially with cross-border mergers, combining efforts of two teams with different culture (and frequently former rivals) is not always successful. A German-English merger deal in the Communications sector can be a good sign for a potential German-English merger in the Banking sector. Another possible merger specific factor is related to the role of Antitrust authorities. The new Bush administration is expected to adopt a more permissive behavior in Antitrust issues than some of its predecessors (namely the Clinton administration). There is also some evidence that when stock markets are booming merger activity is intensified.

In our model, firms are considered to be from different industries so that results do not depend on strategic effects. The third chapter illustrates what would happen if strategic effects are taken into account. The main conclusions are that herding behavior can generate a merger cascade even if the state of the world is adverse to merger activity and if firms have the prior belief that a merger reduces their value. There is uncertainty about merger profitability but firms may receive a signal indicating which state of the world is more likely to occur. Even if firms expect a priori that a merger is not profitable, a positive signal received by the first firms may influence the actions of all other firms. By observing the mergers of the first firms, the other firms decide to ignore their own signals and herd. This event occurs with substantial positive probability and is due to rational behavior. Surprisingly, although occurring with lower probability, a merger wave is likely to start earlier under these 'pessimistic' priors than when firms have 'optimistic' beliefs which sustain that a merger increases their value. Precision of information, represented by parameter (β), has a clear interpretation in the model. It has a positive impact on herding behavior in the merger activity if the state of the world is favorable and a negative impact on herding if the state of the world is adverse to mergers. On the other hand, the probability of receiving information, represented by parameter (α) has a non monotonic impact on herding behavior on the merging activity when the state of the world is adverse: intermediate values of (α) may foster herding behavior while values at the extremes lead to less herding. By computing the expected utility of the game and comparing the marginal impact of the two parameters, we show that firms prefer to invest in information precision (β) unless the

probability of getting information (α) is very low.

The Contagion Model of industry population dynamics suggests that movements in populations are driven by information cascades which lead to an overreaction to surprise shocks. We find that such model fits data on the last thirty years of merging activity reasonably well.

This chapter is organized as follows. The first section provides evidence for the two features of merger activity claimed above and a brief survey of the herding literature. Section 2 describes the model while section 3 solves it and presents the main results. Section 4 discusses the probabilities of the events and externalities. Section 5 computes expected utility of the game and provides some guidance for firms' investment in information. Section 7 performs a simple econometric exercise based on the Contagion Model of industry population dynamics and section 8 concludes.

3.1.1 Mergers' performance

Evidence from historical merger waves is consistent with the following results. Acquisitions entail a gain for the target firm's shareholders (see Jensen and Ruback (1983), Eckbo (1983), Bradley, Desai and Kim (1988)) while the average return to the bidding firm's shareholders is less clear and typically around zero. The empirical IO literature on mergers' performance involves a larger time span (three to ten years between pre and post merger period). The results are more dramatic and can be summarized as "evidence shows...that the profits from merging companies generally decline after the mergers", Mueller (1993). Many case studies and anecdotal examples illustrate this point. About the promising DaimlerChrysler merger one can now read "impending announcements of huge Chrysler losses, and a share-price collapse so serious that the combined company is worth less than Daimler-Benz was before its takeover of Chrysler was finalized in November 1998"⁴.

To what concerns theoretical literature, there are several explanations for unprofitable mergers. Roll's (1986) hubris hypothesis suggests that managers make valuation mistakes because they overestimate own abilities and the value of the target firm, as result they overpay. Shleifer and Vishny (1988) sustain that managers do not necessarily maximize the value of the firm. Managers pay for the benefits of the acquisition that they care about but that have no value to shareholders. Fauli-Oller and Motta (1996) show that unprofitable takeovers may arise if managerial incentives depart from profit maximization to include considerations of size. Fridolfsson and Stennek (2000) argue that a defensive merger mechanism may explain why mergers reduce profits: an unprofitable merger may occur if being an outsider is even more disadvantageous than being a merging party.

⁴The Economist, October 14th-20th 2000, pp. 86.

3.1.2 Herding

Theoretical herding literature can be divided in three categories which are not mutually exclusive: informational cascades, reputational herding, and investigative herding. Applications and examples include typically portfolio or investment decisions. The first two types of herding occur when individuals choose to ignore their private information and mimic actions of individuals who acted previously. Informational cascades occur when the existing aggregate information becomes so overwhelming that an individual's single piece of private information is not strong enough to reverse the decision of the crowd. Hence, the individual chooses to mimic the crowd rather than act on his private information. If this holds for an agent, it is likely to also hold for anyone acting after this person. This domino effect is the so called cascade: a train of individuals acting irrespective of the content of their signal. The models of Banerjee (1992), Bikhandani *et al.* (1992), and lately, Neeman (1999), Taylor (1999) fall into this category. Our model follows the approach of Banerjee.

Reputational herding concept and basic model were developed by Scharfstein and Stein (1990). Herd behavior arises as a consequence of rational attempts by agents to enhance their reputation as decision makers. Later contributions include Graham (1999) and Avery and Chevalier (1999).

Investigative herding occurs when an agent chooses to investigate a piece of information he believes others will also examine. The agent would like to be the first to discover the information but can only profit from an investment if other investors follow suit and push the price of the asset in the direction anticipated by the first analyst. Otherwise, the first agent may be stuck holding an asset that he cannot profitably sell. Papers by Brennan (1990), Froot, Scharfstein and Stein (1992), fall into this group.

Empirical herding literature is inspired by a different type of reasoning. Given that it might be difficult to distinguish whether observed herding occurs for the one or another reason pointed above, this literature investigates empirical clustering without directly testing the implications of the herding models. Herding has been observed among pension funds, mutual funds, and institutional investors (for example, when a disproportionate share of investors engage in buying, or at other times selling, the same stock).

Our approach is placed in the first category of the theoretical works even though the idea comes from observed clustering behavior: merger waves. Two main contributions to the existing literature can be pointed out. The first one is that the herding concept was never applied to merger waves, even though it has been extensively applied to portfolio management. The second one is the adjustment of Banerjee's herding model to the case of a finite set of actions and the analysis of the influence of priors in agents' behavior.

3.2 The model

This chapter performs an application of the concept of herding to merger waves. The sequentiality of merger decisions, the extensive available information about merger activity industry-wide, and the relatively long time lag required to assess merger's profitability, match the essential settings of the herding models and make the herding story likely to be a reasonable explanation. This work points out a plausible reason for merger unprofitability that encompasses many of the explanations listed above: herding behavior. This would explain why according to various commentators, many mergers are carried out precipitately, with no explicit future plan and hence more likely to lose value. The fact that many deals do not create value tends to support the herding idea.

Consider N risk neutral firms that maximize profits. Assume that these firms are not part of the same industry. The purpose of this assumption is to isolate the results of firms' strategic interaction when they are part of the same industry. Industry concentration has no relation to the realized payoffs. Hence, the findings of sections 2 and 3 are obtained without any kind of oligopolistic strategic interaction between firms and are entirely due to herding behavior. Section 4 extends the model to consider intra-industry merger waves.

There are two possible states of the world, *Good* and *Bad*, that occur with probabilities $\lambda, (1 - \lambda)$ respectively. In the *Good* state of the world, the expected variation in a single firm's profits from merging is positive (+1). In the *Bad* state of the world the expected variation in a single firm's profits from merging is negative (-1). In both states of the world, the profit variation from remaining independent and not merging is assumed to be zero meaning that there is no variation from the current payoff. This assumption allows one to exclude all intra-market strategic behavior referred to above. Profits are announced after all firms made their choice. This hypothesis is not completely unrealistic if one takes into account that mergers' performance can only be evaluated in the medium-long run. These two states of the world capture the correlation of merging activity in different industries. Merger waves are not confined to a specific industry even though a particular sector may take the lead. Hence, the correlation addressed here is merger specific and not industry specific. Examples are the "cultural costs of merging", the strictness of behavior of Antitrust authorities and the state of the stock market. A merger deal between firms from countries A and B in sector W is likely to be a positive sign for a planned merger combining also firms from countries A and B in sector Y (in the sense that other firms also believe that those cultural costs are not that high). Analogously, the fact authorities are adopting a more permissive behavior towards merger regulation is a good sign for potential deals. Finally, it is a generalized opinion of commentators that "when the stock market is booming, managers are more emotionally ready to take big decisions. When the market is down, deals seem much riskier"⁵.

⁵The Economist, January 27th 2001, pp. 67.

The decision making is sequential, one firm chosen at random takes its decision first. The next firm, also chosen at random, takes its decision next but it is allowed to observe the choice made by the previous firm. However, the performance (payoff) of the previous merger cannot be observed before taking one's decision. The action space is composed by only two actions: to merge, denoted by (M) or not to merge, denoted by (X). Let a_j , with $j = 1, \dots, N$, be the action of firm j . Hence, firms would like to play (M) if the state is *Good* and to play (X) if the state is *Bad*. The matching problem of the merging parties is totally left aside of this modelization⁶. Potential partners (or takeover targets) are assumed to belong to a set of firms P . Denote by F the set of N firms called upon to choose and action (M or X). The sets are such that $P \cap F = \emptyset$.

No special assumptions are imposed on the underlying distributions of profit variation in each state of the world apart from the fact that they overlap⁷. Hence there will be firms that would always increase value with a merger, there will be firms that always lose value with a merger, and there will be firms that increase value with a merger if the state of the world is *Good* but lose value if the state of the world is *Bad*. More intuitively, in the *Good* state, to merge is good on average and in the *Bad* state, to merge is bad on average, even though there are firms that would always be better off by merging and others that always lose by merging. Note that the *Good* and *Bad* states of the world have merger specific properties that do not have to be correlated with states of the economy (expansions and recessions).

Each firm may receive a signal, with probability α ,⁸ or not receive a signal. Denote s_j , with $j = 1, \dots, N$, the signal obtained by firm j . The signal enables firms to infer which is the current state of the world. The signal can be one of two: "a merger increases the value to the firm" denoted by $\{m\}$, and "a merger reduces the value of the firm" denoted by $\{x\}$. The timing of signal distribution is irrelevant. It can be right at the beginning to all firms, or short before each firm is called upon to choose an action, individually. The only crucial aspect is that the received signal is not observed by other firms. The parameter α represents the probability of having ideas about firm's future plans. These might consist of the merging strategy or the investment strategy to make the firm grow internally. Typically the latter requires more time to deliver profits and therefore I assume that it brings no variation in the short-run profit level. The driving force of the ideas might be the press, the industry performance, the aggregate state of the economy and signs that the firm's business is declining. Managers learn what is happening in other industries by reading newspapers or leading business magazines like *The Financial*

⁶Note that the merging parties do not combine received signals. The model illustrates the situation where one merger party (or the bidder in the case of an acquisition) is stronger and takes the decision whereas the other party (or target) acts passively and does not contribute significantly to the deal process.

⁷The same results are obtained without this assumption. However, it enriches the model with respect to economic interpretation.

⁸We keep the notation of the herding model of Abijit Banerjee (1992).

Times, The Wall Street Journal, Business Week, Fortune or Forbes. Managers learn what is the state of the world through the level of observed demand and official data on relevant indicators. Basically, the signal $\{m\}$ means that the state *Good* is the most likely state of the world and the signal $\{x\}$ means that the state *Bad* is the most likely state of the world. *Good* and *Bad* may mean expansion and recession, high demand and low demand, or any other more specific indicator relevant to the firm. Clearly, the incentives to change plans might also come from internal problems, from market share reductions, from decreasing profits, and other possible signs of decay. Hence, the signal $\{m\}$ might also arise due to declining performances of firms⁹. On the contrary, many other firms may consider that nothing major is changing in the economy and specially in their industries, may be not so well informed, and therefore have a passive behavior. These are less likely to get ideas about future plans and hence less likely to have a signal.

The signal is not always true and the probability that it is false is $(1 - \beta)$. It is assumed that $\beta > \frac{1}{2}$ to avoid the uninteresting case in which firms randomize among the two actions when they received a signal. The parameter β represents the precision of the signal. Consider a firm who gets the signal $\{m\}$. The accuracy of such signal might depend on the opinion of consultant groups, on the time spent preparing the deal, on the time spent preparing future plans, on strengths and similarities between the firm and the chosen partner¹⁰. Consider now a firm who gets the signal $\{x\}$. The accuracy of such signal might depend on the time spent preparing an investment strategy to grow internally, on the quality of the firm's market research department, on the cohesion among the different departments within the firm, and again, the ratio of passed failures over the total number of deals. Note that these are only possible interpretations of the variables α, β . Throughout the model α, β are not choice variables but are taken by firms as given.

Firms might also receive no signal. In this case they rely on prior beliefs and on passed observed actions to make their decision. Two cases will be studied with respect to priors: 'pessimistic' and 'optimistic' firms. While 'pessimistic' firms have the common prior belief that "a merger reduces the value of the firm", 'optimistic' firms have the opposite common prior belief that "a merger increases the value of the firm". Each new decision maker supports his decision on the basis of his priors, of the history of the past decisions, and his own signal (if it exists). The structure of the game and Bayesian rationality are common knowledge. To avoid trivialities, λ is considered to be close to $\frac{1}{2}$ otherwise signals would play a minor role in taking decisions.

⁹This would illustrate the case of mergers as a product of weaknesses referred to in section 1.

¹⁰Choi and Philippatos (1983,1984) found negative changes in the post-merger value of the bidder to occur in acquisitions that were unrelated (i.e. has no obvious basis for synergy). Some other investigators who analyzed market valuations of mergers in cross-section found that 'related' mergers are valued more highly than those without any apparent synergistic potential (see You at al., 1986). The Economist, 9th January 1999, pp 13, 'Similarity of outlook: Sandoz and Ciba Geigy'.

More formally, the signal's space S is defined as $S = \{\emptyset, m, x\}$ and the action's space is defined as $A = \{M, X\}$. Denote H_k the history of actions taken by firms 1 to $k - 1$, i.e., $H_k = \{a_1, a_2, \dots, a_{k-1}\}$. The information set of player k is composed by H_k , the signal s_k (if it exists) and the initial priors.

3.3 Solving the model

3.3.1 Pessimistic firms

As described before, firms are randomly chosen to take an action which can be M or X . Firms make their decision based on their information set. From the history of past actions the decision making firms can infer past signals and hence evaluate the likelihood of each state of the world. To avoid uninteresting cases we impose $\beta > 1 - \lambda$. If the probability that the state is *Bad* is higher than the probability that signals are right, firms will be more skeptical in following own signal. In the limit, if the state is *Bad* with probability 1, no firm will take the decision to merge. Pessimistic firms hold prior the belief that the *Bad* state of the world is more likely to occur, therefore, $1 - \lambda > \lambda$.

- First firm

The first firm will decide upon only signals and priors since H_1 is an empty set. If a firm has a signal, the probability that it is right is $\beta > \frac{1}{2}$. In the next line we will show that if firm 1 receives a signal it will always follow it¹¹. In other words, if $s_1 = \{m\}$, firm 1 will play $a_1 = \{M\}$, if $s_1 = \{x\}$, firm 1 will play $a_1 = \{X\}$, and if firm 1 has no signal it will choose according to priors and play $a_1 = \{X\}$. More formally, the decision is made as follows. Suppose firm 1 got signal $s_1 = \{m\}$. The expected increase in profit from playing $\{M\}$, denoted $E(M | m)$ is given by:

$$E(M | m) = \frac{\lambda(\alpha\beta)}{P[m]} - \frac{(1-\lambda)\alpha(1-\beta)}{P[m]} > 0 \text{ for } \beta > 1 - \lambda$$

$$P[m] = \alpha(\lambda\beta + (1-\lambda)(1-\beta)).$$

The expected increase in profit from playing $\{X\}$ is zero by construction, no matter received signals or stage of the game,

$$E(X | m, x, \emptyset) = 0.$$

Hence, firm 1 will play $a_1 = \{M\}$. If $s_1 = \{x\}$, the expected increase in profit from playing $\{M\}$ is now:

$$E(M | x) = \frac{\lambda\alpha(1-\beta)}{P[x]} - \frac{(1-\lambda)\alpha\beta}{P[x]} < 0 \text{ for } \beta > \lambda$$

$$P[x] = \alpha(\lambda(1-\beta) + (1-\lambda)\beta).$$

¹¹Given that the potential gains/losses from merging are symmetric in expected value.

Therefore, firm 1 will play $a_1 = \{X\}$. Finally, if firm 1 has no signal, the expected increase in profit from playing $\{M\}$ is trivially given by:

$$E(M | \emptyset) = \frac{\lambda(1-\alpha)}{(1-\alpha)} - \frac{(1-\lambda)(1-\alpha)}{(1-\alpha)} < 0 \text{ for } 1-\lambda > \lambda.$$

The pessimistic firm will again decide not to merge.

If players knew the current state of the world, they would like to play $\{M\}$ if the state is *Good* and $\{X\}$ if the state is *Bad*. In later stages of the game, players analyze the probability that the state is *Good* given their information set and play $\{M\}$ if it is higher than $\frac{1}{2}$, otherwise play $\{X\}$. Given the symmetry of payoffs it is as if firms for each decision would compare the probability that they are right with the probability that they are wrong and choose $\{M\}$ or $\{X\}$ accordingly. We will use this interpretation to simplify the exposition from now onwards.

- Second firm

The second firm can observe two possible histories, $H_2 = \{M\}$ or $H'_2 = \{X\}$. Take the first case, $H_2 = \{M\}$, firm 2 is sure that firm 1 got a signal and that it was $\{m\}$. If firm 2 got $s_2 = \{m\}$, it will follow own signal that is reinforced by firm 1's action. If firm 2 got no signal, it will follow firm 1 given that firm 1 is right will probability β which is higher than $\frac{1}{2}$. The interesting case is when firm 2 has the opposite signal, $s'_2 = \{x\}$. In this case, firm 2 knows that firm 1 also had a signal but one of the two is wrong. The probability that firm 1 is right and the probability that 2 is right are given by:

$$P[1 \text{ right} | H_2, s'_2] = \frac{\lambda\alpha\beta\alpha(1-\beta)}{\alpha\beta\alpha(1-\beta)} = \lambda$$

$$P[2 \text{ right} | H_2, s'_2] = \frac{(1-\lambda)\alpha\beta\alpha(1-\beta)}{\alpha\beta\alpha(1-\beta)} = 1-\lambda.$$

Given that signals are of equal quality, firm 2 will follow the priors, will play $a_2 = \{x\}$ and no herding occurs.¹²

Consider now the other possible history $H'_2 = \{X\}$. Firm 2 learns that firm 1 either had no signal (and the choice was made based on priors) or had $s_1 = \{x\}$. If firm 1 has no signal, it follows priors and hence it is right with probability $(1-\lambda)$. Both if firm 2 has no signal or if it has $s_2 = \{x\}$, firm 2's action will be $a_2 = \{x\}$. The interesting case arises when $s_2 = \{m\}$. More formally,

$$P[1 \text{ right} | H'_2, s_2] = \frac{(1-\lambda)[\alpha\beta + (1-\alpha)]\alpha(1-\beta)}{P[H'_2, s_2]},$$

$$P[2 \text{ right} | H'_2, s_2] = \frac{\lambda[\alpha(1-\beta) + (1-\alpha)]\alpha\beta}{P[H'_2, s_2]}.$$

$$P[2 \text{ right} | H'_2, s_2] > P[1 \text{ right} | H'_2, s_2] \text{ for } \frac{1}{2} > \lambda > f(\alpha, \beta).$$

¹²Other derivations can be found in the appendix.

In this case firm 2 will follow own signal and play $a_2 = \{M\}$ and no herding occurs. Even though it might be possible that both firms 1 and 2 had opposite signals (in which case priors would decide for $\{X\}$), it is also possible that firm 1 simply had no signal. Therefore, firm 2 will follow its own signal as long as λ is close to $\frac{1}{2}$. (See appendix for details).

• Third firm

The third firm can observe four different histories given that all combination of the two possible actions might occur. Consider the case in which $H_3 = \{M, M\}$. Firm 3 knows that firm 1 had signal $\{m\}$ and that firm 2 had no signal or the same signal. If $s_3 = \{m\}$ or $s_3 = \{\emptyset\}$, firm 3 will play $a_3 = \{M\}$ and no herding occurs. If $s'_3 = \{x\}$ then, as shown below, firm 3 ignores own signal and follows firm 1.

$$P[1 \text{ right} | H_3, s'_3] = \frac{\lambda \alpha \beta [\alpha \beta + (1 - \alpha)] \alpha (1 - \beta)}{P[H_3, s'_3]},$$

$$P[3 \text{ right} | H_3, s'_3] = \frac{(1 - \lambda) \alpha (1 - \beta) [\alpha (1 - \beta) + (1 - \alpha)] \alpha \beta}{P[H_3, s'_3]},$$

$$P[3 \text{ right} | H_3, s'_3] < P[1 \text{ right} | H_3, s'_3] \text{ for } \lambda > \frac{1 - \alpha \beta}{2 - \alpha}.$$

The condition on λ imposes that the Good state has to occur with a sufficiently high probability even though firms are pessimistic. In the extreme case where there are no signals ($\alpha = 0$) it has to occur as often as the Bad state. The condition corresponds to $\lambda > 1 - \beta$ in the case of $\alpha = 1$. Hence, for $\alpha > 0$ it allows the coexistence of herding under pessimistic beliefs. Under such conditions, the merging cascade is already created. From the third firm onwards, all firms will herd on the action $\{M\}$ as will be shown by induction. First, I should clarify the concept of 'cascade' in the herding literature.

Definition: *A cascade occurs when it is optimal for a decision maker, having observed the actions of those ahead of him, to follow the behavior of the preceding decision maker without regard to his own information.*

Once this stage is reached, his decision is uninformative to others. Hence, the next decision maker draws the same inference from the history of past decisions provided that signals are independent.

Proposition 3.1 *If the first two firms decide to merge, all N firms will merge, independently of their signals.*

Proof: Suppose that the fourth firm who is called upon to play receives signal $s'_4 = \{x\}$ and faces history $H_4 = \{M, M, M\}$, (the solution is trivial for in the case of other signals). Firm 4 can infer that the first firm received a signal to merge, the second firm might have had no signal or the also a signal

to merge. Once this stage is reached, his decision is uninformative to others. No information can be inferred from third firm's behavior since it might have had no signal, a signal to merge and a signal not to merge. In any of the cases firm 3 would follow firm 1 and ignore own signal. More formally,

$$\begin{aligned}
 P[1 \text{ right} | H_4, s'_4] &= \frac{\lambda\alpha\beta[\alpha\beta + (1-\alpha)][\alpha\beta + \alpha(1-\beta) + (1-\alpha)]\alpha(1-\beta)}{P[H_4, s'_4]}, \\
 P[4 \text{ right} | H_4, s'_4] &= \frac{(1-\lambda)\alpha(1-\beta)[\alpha(1-\beta) + (1-\alpha)][\alpha\beta + \alpha(1-\beta) + (1-\alpha)]\alpha\beta}{P[H_4, s'_4]} \\
 P[3 \text{ right} | H_3, s'_3] &= P[4 \text{ right} | H_4, s'_4].
 \end{aligned}$$

Note that the expression $[\alpha\beta + \alpha(1-\beta) + (1-\alpha)]$ is equal to 1 and corresponds to the action of firm 3. Everything might have happened given that firm 3 herds. As a result, the fourth firm is exactly in the same situation as a third firm who receives signal $s'_3 = \{x\}$. Again, given that firm 1 has a higher probability of being right, the fourth firm will follow the herd disregarding its own signal. The k^{th} firm will then have $k-3$ terms with value 1 on the numerator of the probabilities of being right or wrong. Given that firms after 3 draw the same inference from the history of past decisions, they will also ignore their private information and take the same action as the previous decision maker. And so do all later decision makers. .

More generally, if firm j is in a cascade, then its action conveys no information and firm $j+1$ draws the same inference from all previous actions. Since the signal s_{j+1} , if it exists, has the same quality as signal s_j , firm $j+1$ is also in a cascade. Thus, by induction, all firms after j are in a cascade. As a result, a cascade once started will last forever. ■

Consider now the case in which $H'_3 = \{X, X\}$. Again, if $s'_3 = \{x\}$ or $s_3 = \{\emptyset\}$, firm 3 will play $a_3 = \{x\}$ and no herding occurs. The interesting case arises when $s_3 = \{m\}$.¹³ After simple calculations it can be showed that firm 3 follows own signal $\{m\}$ if

$$(1-\alpha)^2\alpha[\beta - (1-\lambda)] - 2\alpha\beta\alpha(1-\beta)(1-2\lambda)(1-\alpha) - \alpha\beta\alpha(1-\beta)\alpha(\beta-\lambda) \geq 0.$$

The case of equality corresponds to the situation in which both firms (1 and 3) have the same probability of being right. As was discussed before, firm 3 will act according to priors and will also play $\{X\}$. Little intuition is provided by this expression. However, considering the case of equal probability of the two states ($\lambda = \frac{1}{2}$) it simplifies to the following one,

$$\begin{aligned}
 (1-\alpha)^2 - \alpha^2\beta(1-\beta) &\geq 0 \Leftrightarrow \\
 \alpha &\leq \frac{1 - \sqrt{\beta(1-\beta)}}{1 - \beta(1-\beta)}.
 \end{aligned}$$

¹³In terms of information structure, this case corresponds to the one of optimistic firms, $H_3 = \{M, M\}$ and $s_3 = \{x\}$ studied in the next section. More details will be given there.

Therefore, firm 3 will herd for levels of α above this threshold and will follow its own signal otherwise. The expression depends positively on β . One can interpret the result as follows, herding is likely to occur for high values of α (probability of having a signal), given that with high probability both firms 1 and 2 chose according to a signal (which is the same) and did not simply followed priors. The relation with β , the signals' precision, is also clear. The higher the value of β , the more firm 3 trusts its own signal and therefore, the less likely the herding behavior is to start. As β approaches 1, the narrow the range of α that can sustain the herding solution, in the limit, the condition to be satisfied would be $\alpha > 1$, which is impossible.

This result shows that the herding process in both actions is not symmetric. Surprisingly, firms start earlier herding on the action that is contrary to their beliefs than on the action that supports their beliefs. This point will be clarified in the next section.

3.3.2 Optimistic firms

This section illustrates the influence of priors in herding probability. Assume now that firms are 'optimistic' i.e., have the common prior belief that "a merger increases the value of the firm". Clearly, the first firm will play $\{M\}$ both if it has no signal and if it received signal $\{m\}$, and will play $\{x\}$ if it received a signal not to merge $\{x\}$. When observing history $H_2 = \{M\}$, firm 2 will play also $\{M\}$, both if it has no signal or if it received signal $\{m\}$. However, it will play $\{x\}$ if it received signal $\{x\}$, as long as $\beta > \lambda$. Analogously, if $H'_2 = \{x\}$, firm 2 will play also $\{x\}$, both if it has no signal or if it received signal $\{x\}$. However, it will play $\{M\}$ if it received signal $\{m\}$ given that for signals of equal quality, firm 2 will act according to priors. Consider now a firm 3 that observes history $H_3 = \{M, M\}$, again, its action will be $\{M\}$ in the case of no signal or signal $\{m\}$. The interesting case arises when firm 3 receives a signal not to merge $s'_3 = \{x\}$.

$$P[1 \text{ right} | H_3, s'_3] = \frac{\lambda[\alpha\beta + (1-\alpha)][\alpha\beta + (1-\alpha)]\alpha(1-\beta)}{P[H_3, s'_3]},$$

$$P[3 \text{ right} | H_3, s'_3] = \frac{(1-\lambda)[\alpha(1-\beta) + (1-\alpha)][\alpha(1-\beta) + (1-\alpha)]\alpha\beta}{P[H_3, s'_3]},$$

By equating the expressions it can be shown that for $\lambda = \frac{1}{2}$,

$$\alpha \leq \bar{\alpha} \Rightarrow \text{No herding occurs, } a_3 = \{X\}$$

$$\alpha > \bar{\alpha} \Rightarrow \text{Herding occurs, } a_3 = \{M\}$$

$$\text{with, } \bar{\alpha} = \frac{1 - \sqrt{\beta(1-\beta)}}{1 - \beta(1-\beta)}.$$

The function $\bar{\alpha}(\beta)$ is increasing in $\beta \in (\frac{1}{2}, 1)$ and ranges from $(0.66, 1)$. The higher the β , the higher the upper bound of $\bar{\alpha}(\beta)$, and hence, the less likely that herding takes place. Hence, for high

values of α firm 3 ignores own signal and starts herding whereas for low values of α firm 3 will follow own signal and decide not to merge. Contrasting this result with the one obtained in the previous section the following conclusion can be derived.

Proposition 3.2 *'Pessimistic' firms are likely to start herding earlier on the action to merge than 'Optimistic' ones.*

Proof: From proposition 1 and the result above. See appendix A.1 for a discussion of the general case $\lambda \in (0, 1)$ ■

The result seems paradoxical since the skeptical firms who believe that a merger destroys value is the one that starts earlier a merger cascade. Note that the main difference between the two situations is the action of the first firm who is called play. In a pessimistic environment, if firm 1 decided to merge, all subsequent firms are sure that firm 1 had a signal $\{m\}$. On the contrary, an optimistic environment who sees the first firm deciding to merge cannot distinguish whether firm 1 really had a signal or is acting according to priors. Therefore, for low probability of getting a signal, (low α , or more precisely, $\alpha \leq \bar{\alpha}$) the third firm of an optimistic society who receives a signal not to merge, thinks that both firms 1 and 2 are acting according to beliefs. The result is not that surprising in the economic literature, it is deeply related to the argument of a paper by Cukierman and Tommasi (1997)¹⁴. As one would expect, the higher the value of β , the higher the signal's credibility and therefore, the less likely that firm 3 ignores own signal to follow the crowd. Note that as β approaches 1, ($\beta \rightarrow 1$), the range of α 's for which the result holds also increases steadily until the point in which all values of α sustain the no herding result. Now, the fact that herding in the merging action occurs earlier under the pessimistic society than under the optimistic one, does not mean that it occurs there with a higher probability. This point will be discussed in the next section.

The impact of λ is somehow intuitive. Recall that for optimistic firms $\lambda > \frac{1}{2}$. Again, we are particularly interested in λ close to $\frac{1}{2}$ otherwise results are trivial. If α is low enough, firms 1 and 2 are more likely to be choosing based on the prior than based on a signal. Therefore, firm 3 has incentives to follow its signal and not herd on the action to merge. However, the higher parameter λ , the more likely a firm whose behavior is based on priors is right. Thus, the fact of having received a signal becomes 'less important' and the condition on $\bar{\alpha}$ is relaxed, i.e., for each β , threshold $\bar{\alpha}$ decreases. As a result, the herding condition is satisfied by a larger range of parameters (α, β) .

¹⁴In a paper with the suggestive title "When does it take a Nixon to go to China?", about policy reversals, the authors show that extreme, but rarely proposed, policies are more likely to be implemented by "unlikely" actors. The intuition behind is that a leftwing policymaker has more credibility when he proposes a significant policy shift to the right than when he proposes a significant policy shift to the left. The opposite would occur with a rightwing policymaker.

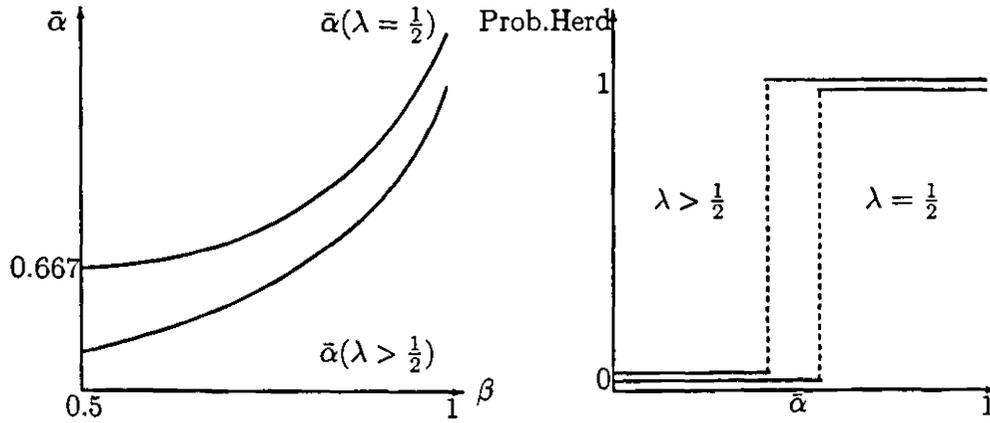


Figure 1 - Effect of $\lambda > \frac{1}{2}$

Recall that for α high, i.e. $\alpha > \bar{\alpha}$, herding starts with the third firm. As a result, its decision is uninformative to others. When the fourth firm is called upon to play, it draws the same inference from the history of past decisions. Consider again the situation in which firms 1 to 3 took the decision to merge and the fourth firm gets a signal $\{x\}$. Firm 4 it will also decide to ignore own signal and 'follow the crowd'. Now, for α low, herding might not yet start with the fourth firm, even if all previous firms decided to merge. However it will start at some point in a later stage.

Proposition 3.3 *Herding behavior will eventually start among 'Optimistic' firms when all previous firms decided to merge.*

Proof: Consider the case of low enough α and history $H_l = \{M, M, \dots, M\}$. If $l - 1$ firms have followed the first firm and decided to merge, and the l th firm has a different signal $s'_l = \{x\}$, the ratios of probabilities are given by

$$P [1 \text{ right} | H_l, s'_l] = \frac{\lambda [\alpha\beta + (1 - \alpha)] [\alpha\beta + (1 - \alpha)]^{l-2} \alpha (1 - \beta)}{P [H_l, s'_l]}$$

$$P [l \text{ right} | H_l, s'_l] = \frac{(1 - \lambda) [\alpha (1 - \beta) + (1 - \alpha)] [\alpha (1 - \beta) + (1 - \alpha)]^{l-2} \alpha\beta}{P [H_l, s'_l]}$$

Note that from the terms to the power of $l - 2$,

$$\alpha\beta + (1 - \alpha) > \alpha (1 - \beta) + (1 - \alpha).$$

Therefore, as l increases, the ratio of the first probability to the second also increases (no matter the values of the parameters). Hence, for large enough l , the decision maker will decide to ignore own signal and also follow firm 1. ■

This result is derived for history $H_l = \{M, M, \dots, M\}$. The $l - 2$ firms who played before l had either no signal or the same signal as firm 1. Hence, the higher the number of these ‘herding’ firms, the more the l th decision maker tends to think: ‘they can not be all wrong’¹⁵. It can be shown that for uniform priors, herding behavior starts with the fourth firm. Therefore, one is tempted to claim that parameter l does not need to increase ‘too much’.

3.4 Merger cascades

In this section, in order to assess the relevance of herding phenomena, it is shown that the probabilities associated to cascades are definitely not negligible. It was proved in previous sections that history $H_3 = \{M, M\}$ under the pessimistic environment is enough to generate a merger cascade starting in firm 3. Let the state be *Good*. The probability of this event is given by:

$$P[\{M, M\} | \text{Good}] = \alpha\beta(\alpha\beta + 1 - \alpha).$$

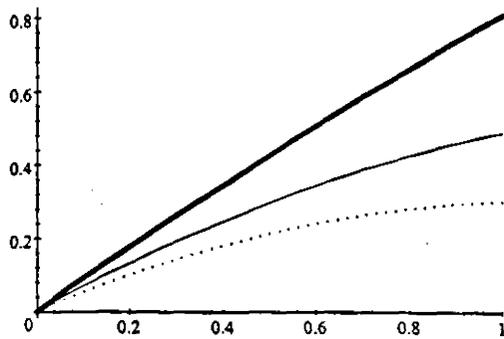
Recall that

α	probability of receiving a signal
β	probability that the signal is right

It is easy to check that $P[\{M, M\} | \text{Good}]$ depends positively on both parameters α and β , and can take values in the interval $[0, 1]$. Figure 1 below plots this probability for given values of β (see table 1 in appendix 2 for the actual values of α, β).

¹⁵See Banerjee, 1992, for a more general discussion.

Figure 1 - Probability of N merger decisions in the *Good* state of the world.



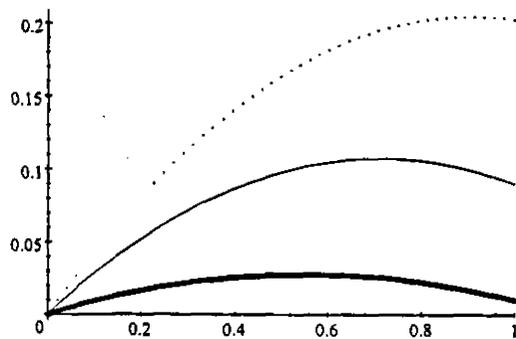
$\beta = 0.55$ — dotted line, $\beta = 0.7$ — thin, $\beta = 0.9$ — thick

Now let the state be *Bad*. As one would expect, the probability of a merger cascade is smaller, even though not irrelevant for a wide range of the parameters:

$$P\{\{M, M\} | Bad\} = \alpha (1 - \beta) [\alpha (1 - \beta) + 1 - \alpha].$$

Figure 2 below plots this probability for the same given values of β . (Find the corresponding table in appendix 2).

Figure 2 - Probability of N merger decisions in the *Bad* state of the world.



$\beta = 0.55$ — dotted line, $\beta = 0.7$ — thin, $\beta = 0.9$ — thick

In this case, the probability depends negatively on β and depends non monotonically on α , i.e., the derivative with respect to α is positive for $0 < \alpha \leq \frac{1}{2\beta}$, and negative otherwise. Hence, the probability of a merger cascade in the wrong state decreases in the interval $\frac{1}{2\beta} < \alpha < 1$. Intuitively what happens is that if α is not too high, in a history of two past mergers, firm two is very likely to be acting upon no signal. This definitely increases the probability that a cascade is started based on a wrong signal

of the first firm. On the other hand, for high values of α , i.e., when almost all firms get a signal, the probability that both firms one and two are acting upon a wrong signal is clearly smaller given that signals are more likely to be right than wrong. Hence, for high levels of α and β it is not likely that the two first firms both receive a signal which is wrong. For the same reason, the probability of a merger cascade in the bad state increases in α for $0 < \alpha \leq \frac{1}{2\beta}$. Firm three knows that firm one must have had a signal $\{m\}$ to play $\{M\}$. The higher the parameter α , the more likely that firm 2 also had a signal. But if firm 2 had a signal it must have been $\{m\}$ otherwise firm 2 would not follow firm 1. This enhances credibility of the first two players decision and makes it possible to generate herding behavior leading to the merger cascade.

Proposition 3.4 *Consider firms who received no signal or a signal not to merge under the 'pessimistic environment'. A merger cascade imposes a positive externality on those firms if the state of the world is Good and imposes a negative externality on those firms if the state of the world is Bad.*

Proof: The assumption about the distribution of payoffs according to the states of the world implies this result. If firms could not observe other firms actions there would be $\alpha\beta N$ firms merging in the *Good* state of the world. However, if firms can observe previous players choices, with the probabilities of the first table all N firms will merge. $N - \alpha\beta N$ firms are on average better off, given that get payoff (+1) on average, whereas by following own signal or priors the payoff would be zero.

On the contrary, if the *Bad* state of the world occurs, there are $N - \alpha(1 - \beta)N$ firms that decide to merge without having a signal suggesting that strategy and acting against own priors. These firms are on average worse off, given that they on average get payoff (-1), whereas by following own signal or priors the payoff obtained would be zero. ■

Note that the probability with which the negative externality is imposed is not negligible and can reach 20% for low values of signal precision β . This result may shed some light on the merger unprofitability puzzle described in section 1. If 'pessimistic' priors are correct and the *Bad* state occurs indeed more often in a repeated scenario, the overall losses are even higher.

Corollary: *Consider firms who received no signal or a signal to merge under the 'optimistic environment'. A cascade on the action not to merge imposes a negative externality on those firms if the state of the world is Good and imposes a positive externality on those firms if the state of the world is Bad.*

Proof: The result follows from the fact that the herding process on the action to merge among 'pessimistic' firms is equal to the herding process on the action not to merge among 'optimistic firms'. Hence, the tables can "translated" to the 'optimistic environment' in the following way. Probability

of herding on the action not to merge when the state of the world is *Bad* (first table) and probability of herding on the action not to merge when the state of the world is *Good* (second table).■

3.4.1 Complete range of herding histories

In this section the complete set of histories that are able to generate a merger cascade and their respective probabilities are computed. Again, the focus is on the 'pessimistic' environment. Note that in the previous tables only history $H = \{M, M\}$ is taken into account. By covering all possible histories the results of the previous section can only be reinforced. In Table 1 those histories are enumerated. Since signals are of equal quality, a sequence of alternating signals in any history provides no information for the current decision maker. Parameter k represents the number of times the pair is repeated in a particular history. Given that the focus is on the pessimistic setting, the herding histories of the optimistic side are not exhausted.

Table 1 - Histories generating merger cascades

	Herding histories
$0 \leq \lambda \leq \frac{1-\alpha\beta}{2-\alpha}$	\emptyset
$\frac{1-\alpha\beta}{2-\alpha} < \lambda < \frac{1}{2}$	$\{MM\}, \left\{ \begin{array}{l} (MX)^k MM \\ (XM)^k MM \end{array} \right\}$
$\lambda = \frac{1}{2}$	$\{XMM\}, \{XXMM\}^*$
$\frac{1}{2} < \lambda \leq \frac{\alpha\beta+(1-\alpha)}{2-\alpha}$	$\{MM\dots M\}$
$\frac{\alpha\beta+(1-\alpha)}{2-\alpha} < \lambda \leq \frac{\alpha\beta(1-\alpha\beta)}{2\alpha\beta\alpha(1-\beta)+\alpha(1-\alpha)}$	$\{MM\}$
$\frac{\alpha\beta(1-\alpha\beta)}{2\alpha\beta\alpha(1-\beta)+\alpha(1-\alpha)} < \lambda \leq \beta$	$\{M\}, \{XM\}$
$\beta < \lambda \leq 1$	\forall

* indicates that it is also valid for $\{(MX)^k XXMM\}, \{(XM)^k XXMM\}$
(k stands for the number of times the pair is repeated).

Note that the probability of a cascade is a step function and not a continuous function in the parameter λ . This is due to the fact that herding decisions are discrete for any given set of parameters. For λ close to zero, we are in the world of pessimistic firms and the *Bad* state of the world occurs too often. As a result, no herding history is sustainable. Clearly, there may be a history where all firms decide to merge (no one gets a signal not to merge) but the train of mergers is not due to herd behavior but to firms following its own signal. The threshold corresponds to condition on the behavior of the third firm that faces a history of two past mergers.

The second range of values is the interesting one. Firms are still pessimistic ($\lambda < \frac{1}{2}$) but the *Good* state occurs with probability high enough to induce herding. Herding will be triggered if the history

starts with two consecutive mergers or if firms alternate actions M, X until two consecutive merger decisions take place. Superscript k stands for the number of times the pair is repeated. A cascade can start with the third firm or with a firm in the position $2k + 3$. Alternate actions require that firms had opposite signals (except for the first firm that may be acting according to priors), as a result, because signals have equal quality, this conveys little information to the current decision maker. It is as if these actions cancel out and play no role in the information game.

When both states are equally likely to occur, not only the previous history $\{MM\}$ but histories $\{XMM\}$, $\{XXMM\}$ are able to generate a cascade. Again, alternated actions M, X or X, M preserve the herding mechanism.

The fourth range of values of λ belongs already to the optimistic environment. No histories are indicated because one can refine this range indefinitely to find herding mechanisms on the action to merge. All histories can not contain two consecutive 'no mergers' in order to generate a merger cascade. As it was shown theoretically, two consecutive mergers cannot trigger a cascade however, one can refine the interval of λ under which histories $\{MMM\}$, $\{MMMM\}$, ... are able to generate herding.

The fifth interval corresponds to the one in which two consecutive mergers are enough to trigger a merger cascade. The higher the λ , the lower the number of consecutive mergers needed to start the herding process. As λ approaches one, the probability that the state is Good is so high that firms will play X in less and less situations. The sixth range of values of λ , requires only one merger to start a cascade. Recall that action M played by the first firm may occur both if the firm had a signal or had no signal. Nevertheless, a second firm that receives a signal not to merge will ignore it and play M . Note that because λ is still smaller than β , if firm one gets a signal not to merge it will follow the signal and play X .

The last interval corresponds to the extreme case: the Good state occurs with a probability that is higher than the precision of the signals. As expected, all firms will ignore signals (in particular a signal not to merge) and play M .

Tables 2 and 3 compute the probabilities of a merger cascade for the parametrization $\beta = 0.75$, $\alpha = 0.6$ and 500 firms (recall that in the model firms belong to all existing industries). The first table computes the probability that the herding mechanism starts with the first firm. For these values of the parameters, a history of three consecutive mergers is enough to generate a cascade under optimistic beliefs. Table 3 computes the overall probability of a cascade starting at any point in the queue¹⁶. Results are not very sensitive to the total number of firms given that histories with many actions occur with insignificant probabilities. As expected, the probability of a merger cascade depends positively

¹⁶A herding cascade on the merging action among 'optimistic' firms may be generated by the following histories $H = \{M, M, M\}$, $H = \{(X, M)_k, M, M, M\}$ or $H = \{(M, X)_k, M, M, M\}$ for $\alpha < \bar{\alpha}$.

on parameter λ , i.e., the probability that the state is *Good*. However, it can occur with positive probability for low levels of λ .

Table 2 - 'Simple' probabilities of a merger cascade (for $\beta = 0.75, \alpha = 0.6$)

Beliefs	History	Good state	Bad state
Pessimistic: $\frac{1-\alpha\beta}{2-\alpha} < \lambda < \frac{1}{2}$	{MM}	.3825	.0825
Optimistic: $\frac{1}{2} < \lambda \leq \frac{\alpha\beta+(1-\alpha)}{2-\alpha}$	{MMM}	.6141	.1664

Table 3 - Complete probabilities of a merger cascade (for $\beta = 0.75, \alpha = 0.6, N = 500$)

Beliefs	Good state	Bad state
Pessimistic: $\frac{1-\alpha\beta}{2-\alpha} < \lambda < \frac{1}{2}$.5117	.0997
Optimistic: $\frac{1}{2} < \lambda \leq \frac{\alpha\beta+(1-\alpha)}{2-\alpha}$.7425	.2226

Hence, even though herding may start later than under 'optimistic' beliefs, it starts with a higher probability. Clearly, a merger cascade in the Bad state of the world occurs with lower probability than one in the Good one given that it would involve many firms with wrong signals. However, note that Optimistic firms are more vulnerable to 'Bad' cascades. Due to their positive beliefs about mergers, optimistic firms are less skeptical to merge and therefore run a higher risk of unprofitable deals than pessimistic firms.

Proposition 3.5 *The probability of a merger cascade is higher under optimistic beliefs than under pessimistic ones.*

Proof: See appendix A.3.

3.4.2 Waves

In the previous section it is shown that a merger wave occurs with high probability due to herding behavior among firms. Herding, can occur in both actions: to merge or not to merge. Once started, herding is irreversible as long as there are no changes on the priors $(\lambda, 1 - \lambda)$. The wave shape is precisely generated by such changes in priors, i.e., changes in the parameter λ . Hence, there is no need to introduce memory loss or other mechanisms to generate the merger waves. The idea is that firms believe in a certain value of λ that generates a specific behavior, and, after a time lag (some years or a decade), it changes. Underlying these changes of prior beliefs can be new anti-trust concerns (or lack of them), changes in interest rates expectations, changes in expected behavior of stock markets. The important fact is that these are not fundamentals of the economy but beliefs.

Stock prices seem to play an important role in merger activity, however, they do not appear to have a great influence in merger profitability. Moreover, the same movement in stock prices can both intensify and refrain merger activity which shows that it can not be considered as an economic fundamental.

About the most recent wave, peaking in 2000 and with signs of decay already at the beginning of 2001 one can read: "When the stock market is booming, managers are more emotionally ready to take big decisions. When the market is down, deals seem much riskier."¹⁷ Similarly, speculative motives are also considered possible triggers the first merger wave 1887-1904: "By exciting false expectations, the promoters were able to sell the stocks of newly consolidated firms at prices far exceeding their true economic value". On the contrary, the merger wave of the eighties was more related to a slump than to a boom in the stock market. Stock prices were so low that it was often less expensive to buy other companies than to build new plants. Another relevant belief was related to expectations: "If we don't make that XYZ acquisition soon, it will cost us more next year because of rising stock prices".¹⁸ To sum up, priors have basically to do with the expectations of decision makers, and not so much with fundamentals of the economy. The example above shows how exactly the same issue can have totally opposite effects in merger activity in different points in time.

It is important to note that the model developed in section two suggests a contagion mechanism based on herding to explain the phenomenon of merger waves. Nevertheless, it does not elaborate much on the way a cascade is started or ends. This is the role of the beliefs that are assumed to be exogenous and change over years or decades.

3.5 Expected Utility

By calculating the expected utility of the game for each player one is able to assess the marginal impact of the parameters in the expected revenue. Obviously, the higher the expected utility the less frequent merger cascades in the 'wrong' state do occur. Therefore, it is of interest to select the informational parameters that are able to reduce both individual potential losses and negative externalities. Note however, that this exercise is a "thought experiment" given that in the model parameters are exogenously given. It indicates though, how a firm would like to change the informational parameters if it could actually influence their values. Recall that λ is the probability of a *Good* state selected by Nature, α is the probability of receiving a signal, and β is the probability that the signal is right. In principle firms cannot affect parameter λ . Investing in parameter α means increase importance given to the press, word of mouth among managers, informal contacts. Investing in parameter β means increasing the accuracy of received informations. This can be achieved by improving research departments, improving pre-merger talks and future plans. More generally, α is related to obtaining more information while β is related to the precision of that information. Little attention will be given to the first two players given that they play no role in the herding process. The results obtained are the

¹⁷The Economist, January 27th 2001, pp.67.

¹⁸Scherer and Ross, Industrial Market Structure and Economic Performance, 3rd Ed., pp.158.

expected ones.

Expected utility of the first player is given by:

$$EU_1 = \lambda(\alpha\beta) - (1 - \lambda)\alpha(1 - \beta) = \alpha(\beta + \lambda - 1)$$

The expression $EU_1(\alpha, \beta, \lambda)$ depends positively on the three components. Recall that the condition for 'pessimistic firms is $\beta > 1 - \lambda$. For $\beta + \lambda - 1 > \alpha$ firm 1 should invest in α , otherwise firm 1 should invest in β .

Expected utility of the second player is given by:

$$EU_2 = \alpha(2 - \alpha)(\beta - (1 - \lambda))$$

Again, the expression $EU_2(\alpha, \beta, \lambda)$ depends positively on the three components. For β sufficiently high, i.e.,

$$\beta > (1 - \lambda) + \frac{\alpha(2 - \alpha)}{2(1 - \alpha)}$$

firm 2 should invest in α , otherwise firm 2 should invest in β . Note that the second term is increasing in $\alpha \in (0, 1)$ takes value 1 for $\alpha = .58579$. Therefore, it is more likely that β is under such threshold and hence the investment should be done in β .

3.5.1 Third firm

Expected utility of the third firm is more difficult to compute since some histories depend on values of the parameters (see appendix A.3 for details). For simplicity consider the cases of $\lambda = \frac{1}{2}$ where states are equally likely to occur. The case of $\lambda = \frac{1}{4}$ where the *Bad* state occurs with higher probability can be found in appendix.

Expected utility of firm 3 is given by:

$$2EU_3 = \alpha(2\beta - 1)[3(1 - \alpha) + 2\alpha\beta\alpha(1 - \beta)] + \alpha^3[\beta^3 - (1 - \beta)^3] \text{ if } \alpha \leq \bar{\alpha}$$

$$2EU_3 = \alpha(2\beta - 1)(1 - \alpha)(2 + \alpha) + 3\alpha\beta\alpha(1 - \beta)(2\beta - 1) + \alpha^3[\beta^3 - (1 - \beta)^3] \text{ if } \alpha > \bar{\alpha}$$

$$\text{where } \bar{\alpha} = \frac{1 - \sqrt{\beta(1 - \beta)}}{1 - \beta(1 - \beta)}$$

When α is high firm three herds on history $H = \{X, X\}$, when α is low it will follow its own signal. Recall that when firm three observes a history with no mergers it does not know whether the previous firms had signals or simply followed their priors. For a low probability of getting a signal, α it is more likely that in fact previous firms had no signal.

The third player would prefer to invest in increasing β when marginal impact of β on the expected utility is higher than the one of α . Non surprisingly, the marginal impact of β is decreasing in β . Investment recommendations would be as follows:

Table 3 - Investment Recommendations when $\lambda = \frac{1}{2}$

β , Investment decision:	α = information/ideas	β = precision
0.55	$0 < \alpha \leq .04760$	$.04760 < \alpha < 1$
0.6	$0 < \alpha \leq .09080$	$.09080 < \alpha < 1$
0.7	$0 < \alpha \leq .1661$	$.1661 < \alpha < 1$
0.8	$0 < \alpha \leq .22936$	$.22936 < \alpha < 1$
0.9	$0 < \alpha \leq .2828$	$.2828 < \alpha < 1$

A firm would rather invest in signal precision β unless the probability of having ideas is very low. In this case, a firm would benefit more by increasing parameter α . The thresholds corresponds to the pairs (α, β) such that $\frac{d}{d\beta} EU_3 = \frac{d}{d\alpha} EU_3$. It can be shown that the higher the probability that the state is adverse to merger activity, the more relevant the investment in α becomes (see appendix)..

3.6 Magnitude of potential losses

Intuitively, a high potential loss might refrain some herding process on the risky action even for risk neutral firms. Consider the case of pessimistic priors. Take a first player who received a signal to merge. The decision to follow own signal will depend on the degree of signal precision. If the magnitude of the potential loss is equal to the magnitude of potential gains, then the decision maker will follow own signal for $\beta > 1 - \lambda \geq \frac{1}{2}$.

To isolate the importance of lack of symmetry in the payoffs, assume that λ is sufficiently close to $\frac{1}{2}$. For an expected potential gain (+1) and an expected potential loss ($-z$) with $z > 1$, the decision maker clearly will require a higher precision of the signal in order to follow its own signal. In the limit, when losses approach infinite ($z \rightarrow \infty$), he will require total precision (zero probability that the signal is wrong) to follow own signal, i.e., $\beta \rightarrow 1$. On the contrary, for an expected potential loss (-1) and an expected potential gain ($+z$) with, the decision maker will be not so demanding and he will play according to signal even for values of precision less than 0.5. Trivially, in the limit, when potential gains approach infinite ($k \rightarrow \infty$), the decision maker will require no precision (total uninformative signal) to follow the signal with risky action.

	Potential loss (+1, $-z$)	Potential gain($+z, -1$)
$\lambda = \frac{1}{2}$	$\beta > \frac{z}{z+1}$	$\beta > \frac{1}{z+1}$
$1 - \lambda > \lambda$	$\beta > \frac{z(1-\lambda)}{\lambda+z(1-\lambda)}$	$\beta > \frac{1-\lambda}{1+\lambda(z-1)}$

Consider the case of increasing potential loss, player 1 will follow own signal provided that such losses are not too high, (according to the table above). Suppose that the second firm also has a signal

to merge. Decision to follow the signal will be taken if

$$E[M | m] = P[1 \text{ right} | H_2](+1) + (1 - P[1 \text{ right} | H_2])(-z) > 0.$$

Again, that simplifies to a maximum level of multiplier z that happens to be the squared previous level.

Table 5 - Tolerated potential losses for firms 1 and 2 (if both firms have signal $\{m\}$).

	Firm 1	Firm 2
$\lambda = \frac{1}{2}$	$z < \frac{\beta}{1-\beta}$	$z < \left(\frac{\beta}{1-\beta}\right)^2$
$1 - \lambda > \lambda$	$z < \frac{\lambda\beta}{(1-\lambda)(1-\beta)}$	$z < \frac{\lambda\beta^2}{(1-\lambda)(1-\beta)^2}$
$\beta, \lambda = \frac{1}{2}$	$\frac{\beta}{1-\beta}$	$\left(\frac{\beta}{1-\beta}\right)^2$
0.6	1.5	2.25
0.7	2.33	5.44
0.8	4.0	16.0
0.9	9.0	81.0

Note that as before, the higher the signal precision the more likely firm 2 will follow its own signal. However, due to the existence of two signals, firm two has now a higher tolerance towards potential losses than firm one. If firm two had no signal, it would be in the same situation as firm one above and would therefore be more sceptical in carrying on a decision to merge. Recall that to generate the merger cascade it would be enough that the third firm faced the history of two past mergers. If gains and losses were not symmetric, the firm would now equate

$$E[M | x] = P[1 \text{ right} | H_3](+1) + P[3 \text{ right} | H_3](-z)$$

to take the decision. It is easy to check that firm three will not ignore its own signal to mimic previous players' actions as long as,

$$z > \frac{\lambda[\alpha\beta + (1-\alpha)]}{(1-\lambda)[\alpha(1-\beta) + (1-\alpha)]} > 1$$

Recall that higher the probability of the Good state of the world, the more likely a merger cascade is to occur. The following table shows how potential losses can refrain the cascade among pessimistic firms, even when their beliefs about the Good state are at its maximum ($\lambda = \frac{1}{2}$).

Table 6 - Tolerated potential losses for firm three.

β, α	0.1	0.3	0.5	0.7	0.9
0.6	1.0213	1.0732	1.1429	1.2414	1.3913
0.7	1.043	1.1519	1.3077	1.549	1.973
0.8	1.0652	1.2368	1.5	1.9545	2.9286
0.9	1.0879	1.3288	1.7273	2.5135	4.7895

Proposition 3.6 *Slightly higher potential losses refrain herding behavior in the action to merge.*

Proof: Follows from above. ■

Note that the maximum tolerated level of potential losses sharply decreases with the third decision maker even to levels lower than the first player's who has no past history. This is due to the uncertainty that the second player did really have a signal. Hence, as one would expect, the fear for large potential losses refrains significantly the herding behavior in the risky action. By relaxing the symmetry assumption, the higher the potential losses, the higher signal precision β has to be in order to firms to follow a signal to merge. It is remarkable that for 70% of accuracy, it is enough that the value of the loss is slightly higher than the gain (4% or 15% more for the lower levels of α), to stop the herding mechanism described in the basic model. Similar results would be obtained by keeping equal the magnitude of gains and losses but assuming risk averse firms.

3.7 A Contagion Model

One well known model used to account for industry population dynamics is the so called Contagion Model¹⁹. This model suggests that movements in populations are driven by fads or information cascades which lead to an overreaction to surprise shocks. Although agents try to anchor their expectations in "the fundamentals", information cascades have a strong effect on agents actions. 'Contagion' is said to happen when the occurrence of an event affects the rate of subsequent occurrence²⁰. In other words, when the probability of an event depends on the previous number of events. A simple model of contagion (as in Geroski and Mazzucatto) can be written as follows:

$$\Delta N(t) = \gamma_0 + \gamma_1 N(t-1) + \gamma_2 \Delta N(t-1) + \varepsilon(t).$$

The constant (γ_0) accounts for the "fundamentals" and should be positive. Coefficient γ_2 should be positive and is the driving force of the 'contagion mechanism': an unexpected rise (fall) in N in period $t-1$ will be magnified over time since agents alter their expectations and therefore their behavior. On the contrary, coefficient γ_1 should be negative so that the higher the level of variable N the 'more

¹⁹Geroski and Mazzucatto use it to study entry and exit of firms in the US population of car producers.

²⁰This phenomena can generate the so called 'overdispersion'.

difficult' to increase next period. Coefficient γ_1 is responsible for inverting the growth path of the variable over time.

Out of curiosity, this section checks whether the Contagion Model fits the data on the last thirty years of merger activity presented in the Introduction. Given the simplicity of the analysis and the small size of the sample it is far fetched to claim that the model explains the dynamics of merger activity. However, one can say that it is not in conflict or even that it supports the theoretical argument. Let $N(t)$ be the number of world-wide mergers in year t .²¹ As it is found often in empirical literature on models of industry evolution both a trend and a quadratic trend are needed to perform the regression. The results are summarized in the table below:

Table 5 - Results from the Contagion Model

	γ_0	γ_1	γ_2	t	t^2
coefficient	1649.9	-0.17005	0.20184	-162.64	4.6229
t - value	(2.564)	(-2.065)	(1.141)	(-2.656)	(2.995)
$R^2 = 0.3685$	$n = 36$		$DW = 1.97$		

Despite the fact that γ_2 is not significant and that not much intuition can be taken from the trends, all variables have the correct signs. The number of new merging firms in period t depends positively on the number of new merging firms in the last period $t - 1$. However, it depends negatively on the absolute number of mergers that occurred in that period. Intuitively, the higher the level of merging firms in $t - 1$, the more difficult it is to top it in the next year. Parameter γ_1 drives the reversion mechanism that creates to the downward part of the wave. On the other hand, if γ_2 is large, then any kind of unexpected rise in the number of merging firms in period t will be magnified over time. Nevertheless, an increase in the number of merging firms will inflate merging rates for subsequent years by gradually diminishing amounts. Hence, much of the interesting dynamics in the data seem to be the consequence of bandwagon responses to 'surprise' movements on Merger & Acquisition activity and not of changes in the 'fundamentals' as the theoretical model of this paper predicts.

3.8 Concluding remarks

This paper attempts to make sense out of two facts concerning merger activity. The fact that mergers tend to occur in waves and the fact that profits of merging companies often decline after the merger is completed. Herding is said to occur when decision makers ignore private information and follow decisions of previous decision makers. A herding model is able to connect these two empirical findings showing that herding behavior can generate merger waves in which all firms lose value. No strategic

²¹The number of world wide mergers between 1963 - 2000 were taken from Mergerstat: www.mergerstat.com.

interaction between firms is considered, therefore, the model is more applicable to the third and fourth merger waves where conglomerates and inter-industry mergers constitute the most common type of consolidation. Specially the third merger wave led to significant losses to the participating firms. Given that a firm has always the option of remaining independent, it is puzzling why would firms engage in such unprofitable deals. This paper suggests that a herding mechanism might be responsible for such loss making waves. Despite being applied to so many economic activities like investment recommendations or forecasts, the herding concept, to our knowledge, has not been used to study merger activity.

Banerjee's herding model is adjusted for a finite set of actions. The effect of priors on the herding mechanism is studied in detail. Firms are sequentially called upon to choose between 'merge' or 'not merge' and hold 'pessimistic' or 'optimistic' beliefs about the impact of a merger on the firm's value. States of the world can be favorable or unfavorable to merger activity. A favorable state of the world may be one in which antitrust policy is more tolerant, cultural cost of merging is not so relevant and hence there is higher probability of success. Despite holding private information contrary to specific action, due to inference based on other players' actions, it can be rational to ignore it and to follow the herd. A merger cascade under 'optimistic' beliefs occurs with higher probability than when firms hold pessimistic beliefs. However, and somehow surprisingly, a merger wave starts earlier under pessimistic beliefs than under optimistic ones. Intuition for this phenomenon is actually simple. In the optimistic environment, when firms observe the first decision maker deciding to merge, they do not know whether such chosen action is based on priors or on a signal to merge. On the contrary, among 'pessimistic' firms, the first firm who decides to merge must have had a signal to do so. As a result, given that a merger decision by a pessimistic firm is more trustful, two firms merging sequentially are enough to generate the cascade.

Comparative statics are centered on the two crucial parameters of the model: the probability of receiving information, α , and the precision of that information, β . Whereas information precision reduces the probability of herding in the adverse state of the world, the probability of getting information has a non monotonic impact on herding behavior when the state of the world is adverse to merger activity. Intuitively, what happens is that if α is not too high, in a history of two past mergers, there is a high probability that a cascade is started based on a wrong signal of the first firm. For high values of α , (almost all firms get a signal), the probability that both firms (one and two) are acting upon a wrong signal is clearly smaller. By computing the marginal impact of these informational parameters on the expected utility of the game it is found that, for equally likely states of the world, a firm would be better off by increasing information precision β as long as the probability of having a signal α is not too low. On the contrary, the higher the probability that the state is adverse to merger activity, the more relevant the investment in collecting more information, α , becomes.

As one would expect, the herding mechanism that was found depends on the symmetry between gains and losses. The increase of potential losses relatively to the gains hinders the herding process in the action to merge. Analogously, if potential gains are higher than losses herding will occur earlier and with a higher probability. Finally, a simple exercise is performed based on one model of industry population dynamics: the Contagion Model. It is shown that such model fits the data from the last thirty years of merging activity reasonably well.

3.8.1 Future work

A strong simplifying assumption of the model is the one that the distribution of payoffs is the same across industries. Still considering an inter-industry framework, the introduction of a more general payoff distribution, with a stochastic term in the definition of the payoffs, could be a valuable improvement to this setting. It would then be possible to allow for differences across industries.

The model is able to illustrate the herding mechanism but has no word about what triggers and stops a cascade. A change in the underlying beliefs is responsible for the reversion of the mechanism and the descending part of the wave. It would be interesting though, to understand better such behavior of the underlying common beliefs and eventually to endogenise them in the model.

One prediction of the model is that firms with higher confidence in own signals are more likely to merge. Hence, more mergers should be observed in a *Good* state when the confidence indicator of managers is higher. By finding a proxy for managerial confidence (or an index of market confidence like debt ratings if it were suitable for some industries) one could test empirically such implication of the model.

This work claims that firms may take the decision to merge just because they observe other firms deciding to merge. As a result, many deals can be pursued precipitately and without a clear future plan. These may constitute enough reasons for failure. A test of the model would be to check the profitability of mergers that carried outside the historical waves. An isolated merger would suggest more preparation, planning and discussion before the completion of the deal. It would definitely not be the result of herding and should then have more chances to succeed.

3.9 Bibliography

Avery, C. and Chevalier, J., "Herding over the Career", *Economics Letters*, 1999, vol. 63, pp. 327-333.

Banerjee, A. V., "A Simple Model of Herd behavior", *Quarterly Journal of Economics*, 1992, vol. 107, pp. 797-817.

Banerjee, A. V., "Herd behavior and the Reward for Originality", mimeo, 1989.

Banerjee, A., and Eckard, E.W., "Are Mega-mergers Anticompetitive? Evidence from the first Great Merger Wave", *Rand Journal of Economics*, 1998, vol.29, No 4, pp. 803-827.

Bikhchandani, S., Hirshleifer, D., and Welch, I., "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *Journal of Political Economy*, 1992, vol. 100, pp. 992-1026.

Bradley, M., Desai, A. and Kim, E., "Synergistic Gains from Corporate Acquisitions and their Division between the Stockholders of Target and Acquiring Firms", *Journal of Financial Economics*, 1988, vol. 21, pp. 3-40.

Brennan, M., "Latent Assets", *Journal of Finance*, 1990, vol.45, pp.709-730.

Bulow, J. and Klemperer, P., "Rational Frenzies and Crashes", *Journal of Political Economy*, 1994, vol.102, pp. 1-23.

Carrol, G. and Hannan, M., *Dynamics of Organizational Populations*, 1992, Oxford University Press, Oxford.

Caves, R. E., "Mergers, Takeovers and Economic Efficiency", *International Journal of Industrial Organization*, 1989, vol. 7, pp. 151-174.

Choi, D. and Philippatos, G., "An Examination of Merger Synergism", *Journal of Financial Research*, 1983, vol. 6, pp 239-256.

Choi, D. and Philippatos, G., "Post Merger Performance among Homogeneous Firm Samples", *Financial Review*, 1984, vol. 19, pp 173-194.

Cukierman, A., and Tommasi, M., "When does it take a Nixon to go to China?", *Tilburg University Center for Economic Research*, 1997, working paper No. 9791.

Deneckere, R., and Davidson, C., "Incentives to form Coalitions with Bertrand Competition", *RAND Journal of Economics*, 1985, vol.16, pp 473-486.

Devenow, A. and Welch, I., "Rational Herding in Financial Economics", *European Economic Review*, 1996, vol.40, pp. 603-615.

Eckbo, B. Espen, "Horizontal Mergers, Collusion and Stockholder Wealth", *Journal of Finance Economics*, 1983, vol. 11, pp.

Farrell J., and Shapiro, C., "Horizontal Mergers: An equilibrium analysis", *American Economic Review*, 1990, vol. 80, pp 107-126.

Fauli-Oller, R. and Motta, M., "Managerial Incentives for Takeovers", *Journal of Economics and*

Management Strategy, 1996, vol. 5, pp. 497-514.

Fridolfsson, S. and Stennek, J., "Why Mergers reduce Profits and raise Share Prices: a Theory of Pre-emptive Mergers", 2000, CEPR discussion paper No. 2357.

Froot, K., Scharfstein, D. and Stein, J., "Herd on the Street", 1992, Journal of Finance, vol.47, pp 1461-1484.

Geroski, P., and Mazzucato, M., "Modelling the Dynamics on Industry Populations", 2000, mimeo.

Graham, J., "Herding among Investment Newsletters: Theory and Evidence", Journal of Finance, 1999, vol. 54, pp. 237-268.

Jensen, M. and Ruback, R., "The Market for Corporate Control: The Scientific Evidence", 1983, Journal of Financial Economics, vol. 11, pp. 5-50.

Kirman, A., "Ants, Rationality and Recruitment" Quarterly Journal of Economics, 1993, vol. 108, pp. 137-156.

Mueller, D. C., The Determinants and Effects of Mergers: An International Comparison, Ed. 1980, Oelgeschlager, Gunn & Hain, Cambridge, MA.

Mueller, D. C., "U.S. Merger Policy and the 1992 Merger Guidelines", Review of Industrial Organization, 1993, vol. 8, pp. 151-162.

Neeman, Z., "Herding and the Winner's Curse in Markets with Sequential Bids", Journal of Economic Theory, 1999, vol. 85, pp. 91-121.

Ravenscraft, D. and Scherer, F.M., "Mergers, Sell-offs and Economic Efficiency", 1987, Brookings Institution, Washington, DC.

Roll, R., "The Hubris Hypothesis of Corporate Takeovers", Journal of Business, 1986, vol.59, no.2, pt.1, pp. 197-216.

Scharfstein, D. and Stein, J., "Herd behavior and Investment", American Economic Review, 1990, vol. 80, pp. 467-479.

Scherer, F.M., Industrial Market Structure and Economic Performance, 1980, 2nd Edition, Rand McNally College Publishing Company, U.S.A.

Shleifer, A. and Vishny, R., "Value Maximization and the Acquisition Process", Journal of Economic Perspectives, 1998, vol. 2, No 1, pp 7-20.

Taylor, C. R., "Time-on-the-Market as a Sign of Quality", Review of Economic Studies, 1999, vol. 66, pp. 555-578.

3.10 Appendix

A.1. - Pessimistic Firms

Recall that $\beta > 1 - \lambda > \lambda$. Expressions are not simplified to allow the reader to easily follow the computations.

- Firm 2

$$H_2 = \{M\}, s_2 = \{m\}$$

$$P[1, 2 \text{ right} | H_2, s_2] = \frac{\lambda(\alpha\beta)^2}{\lambda(\alpha\beta)^2 + (1-\lambda)(\alpha(1-\beta))^2}$$

$$P[1, 2 \text{ wrong} | H_2, s_2] = \frac{(1-\lambda)(\alpha(1-\beta))^2}{\lambda(\alpha\beta)^2 + (1-\lambda)(\alpha(1-\beta))^2}$$

$$P[1, 2 \text{ right} | H_2, s_2] > P[1, 2 \text{ wrong} | H_2, s_2]$$

Play $\{M\}$.

$$H_2 = \{M\}, s_2 = \{\emptyset\}$$

$$P[1 \text{ right} | H_2, s_2] = \frac{\lambda\alpha\beta(1-\alpha)}{(1-\alpha)\alpha[\lambda\beta + (1-\lambda)(1-\beta)]}$$

$$P[1 \text{ wrong} | H_2, s_2] = \frac{(1-\lambda)\alpha(1-\beta)(1-\alpha)}{(1-\alpha)\alpha[\lambda\beta + (1-\lambda)(1-\beta)]}$$

This case is equivalent to the one of firm 1 with $s_1 = \{m\}$

$$P[1 \text{ right} | H_2, s_2] > P[1 \text{ wrong} | H_2, s_2]$$

Play $\{M\}$.

$$H'_2 = \{X\}, s_2 = \{m\}$$

Expressions are in the text. Firm 2 will follow own signal if $\lambda > f(\alpha, \beta)$ where

$$f(\alpha, \beta) = \frac{\alpha\beta\alpha(1-\beta) - (1-\alpha)\alpha\beta + \alpha(1-\alpha)}{2\alpha\beta\alpha(1-\beta) + \alpha(1-\alpha)}$$

Note that $f(\alpha, \beta) < \frac{1-\alpha\beta}{2-\alpha}$ for $(1-\alpha)^2 - \alpha\beta\alpha(1-\beta) \geq 0$. The ratio $\left(\frac{1-\alpha\beta}{2-\alpha}\right)$ is the lower bound of λ for the case $H_3 = \{M, M\}$. The inequality $(1-\alpha)^2 - \alpha\beta\alpha(1-\beta) \geq 0$ defines $\bar{\alpha}$ of the Optimistic firms section.

$$H'_2 = \{X\}, s'_2 = \{x\}$$

$$P[1,2 \text{ right} | H'_2, s'_2] = \frac{(1-\lambda)[\alpha\beta + (1-\alpha)]\alpha\beta}{(1-\lambda)[\alpha\beta + (1-\alpha)]\alpha\beta + \lambda[\alpha(1-\beta) + (1-\alpha)]\alpha(1-\beta)}$$

$$P[1,2 \text{ wrong} | H'_2, s'_2] = \frac{\lambda[\alpha(1-\beta) + (1-\alpha)]\alpha(1-\beta)}{(1-\lambda)[\alpha\beta + (1-\alpha)]\alpha\beta + \lambda[\alpha(1-\beta) + (1-\alpha)]\alpha(1-\beta)}$$

$$P[1,2 \text{ right} | H'_2, s'_2] > P[1,2 \text{ wrong} | H'_2, s'_2]$$

Play $\{X\}$.

$$H'_2 = \{X\}, s_2 = \{\emptyset\}$$

$$P[1 \text{ right} | H'_2, s_2] = \frac{(1-\lambda)[\alpha\beta + (1-\alpha)](1-\alpha)}{(1-\lambda)[\alpha\beta + (1-\alpha)](1-\alpha) + \lambda[\alpha(1-\beta) + (1-\alpha)](1-\alpha)}$$

$$P[1 \text{ wrong} | H'_2, s_2] = \frac{\lambda[\alpha(1-\beta) + (1-\alpha)](1-\alpha)}{(1-\lambda)[\alpha\beta + (1-\alpha)](1-\alpha) + \lambda[\alpha(1-\beta) + (1-\alpha)](1-\alpha)}$$

$$P[1 \text{ right} | H'_2, s_2] > P[1 \text{ wrong} | H'_2, s_2]$$

Play $\{X\}$.

• Firm 3

$$H_3 = \{M, M\}$$

In the text it is shown that for $\lambda > \frac{1-\alpha\beta}{2-\alpha}$ firm 3 will herd and play $\{M\}$ if $s_3 = \{x\}$. It is then obvious that firm 3 will play also $\{M\}$ after signals $s_3 = \{m\}$, $s_3 = \{\emptyset\}$.

Condition $\lambda > \frac{1-\alpha\beta}{2-\alpha}$ is trivially satisfied for $\lambda = \frac{1}{2}$ since $\beta > \frac{1}{2}$.

$$H_3 = \{M, X\}, s_3 = \{m\}$$

$$P[3 \text{ right} | H_3, s_3] = \frac{\lambda\alpha\beta\alpha(1-\beta)\alpha\beta}{\alpha(1-\beta)\alpha\beta[\lambda\alpha\beta + (1-\lambda)\alpha(1-\beta)]}$$

$$P[3 \text{ wrong} | H_3, s_3] = \frac{(1-\lambda)\alpha(1-\beta)\alpha\beta\alpha(1-\beta)}{\alpha(1-\beta)\alpha\beta[\lambda\alpha\beta + (1-\lambda)\alpha(1-\beta)]}$$

Equivalent to case of firm 1 with $s_1 = \{m\}$.

$$P[3 \text{ right} | H_3, s_3] > P[3 \text{ wrong} | H_3, s_3]$$

Play $\{M\}$.

$$H_3 = \{M, X\}, s_3 = \{x\}$$

$$P[3 \text{ right} | H_3, s_3] = \frac{(1-\lambda)\alpha(1-\beta)(\alpha\beta)^2}{\alpha(1-\beta)\alpha\beta[(1-\lambda)\alpha\beta + \lambda\alpha(1-\beta)]}$$

$$P[3 \text{ wrong} | H_3, s_3] = \frac{\lambda \alpha \beta [\alpha (1 - \beta)]^2}{\alpha (1 - \beta) \alpha \beta [(1 - \lambda) \alpha \beta + \lambda \alpha (1 - \beta)]}$$

Equivalent to case of firm 1 with $s_1 = \{x\}$.

$$P[3 \text{ right} | H_3, s_3] > P[3 \text{ wrong} | H_3, s_3]$$

Play $\{X\}$.

$$H_3 = \{M, X\}, s_3 = \{\emptyset\}$$

$$P[1 \text{ right} | H_3, s_3] = \frac{\lambda \alpha \beta \alpha (1 - \beta) (1 - \alpha)}{\alpha (1 - \beta) \alpha \beta [\lambda (1 - \alpha) + (1 - \lambda) (1 - \alpha)]} = \lambda$$

$$P[1 \text{ wrong} | H_3, s_3] = \frac{(1 - \lambda) \alpha (1 - \beta) \alpha \beta (1 - \alpha)}{\alpha (1 - \beta) \alpha \beta [\lambda (1 - \alpha) + (1 - \lambda) (1 - \alpha)]} = 1 - \lambda$$

Equivalent to case of firm 1 with $s_1 = \{\emptyset\}$.

$$P[1 \text{ wrong} | H_3, s_3] > P[1 \text{ right} | H_3, s_3]$$

Play $\{X\}$.

$$H_3 = \{X, M\}, s_3 = \{m\}$$

$$P[3 \text{ right} | H_3, s_3] = \frac{\lambda [\alpha (1 - \beta) + (1 - \alpha)] (\alpha \beta)^2}{\lambda [\alpha (1 - \beta) + (1 - \alpha)] (\alpha \beta)^2 + (1 - \lambda) [\alpha \beta + (1 - \alpha)] [\alpha (1 - \beta)]^2}$$

$$P[3 \text{ wrong} | H_3, s_3] = \frac{(1 - \lambda) [\alpha \beta + (1 - \alpha)] [\alpha (1 - \beta)]^2}{\lambda [\alpha (1 - \beta) + (1 - \alpha)] (\alpha \beta)^2 + (1 - \lambda) [\alpha \beta + (1 - \alpha)] [\alpha (1 - \beta)]^2}$$

$$P[3 \text{ right} | H_3, s_3] > P[3 \text{ wrong} | H_3, s_3]$$

Play $\{M\}$.

$$H_3 = \{X, M\}, s_3 = \{x\}$$

$$P[3 \text{ right} | H_3, s_3] = \frac{(1 - \lambda) [\alpha \beta + (1 - \alpha)] \alpha (1 - \beta) \alpha \beta}{\alpha (1 - \beta) \alpha \beta [(1 - \lambda) [\alpha \beta + (1 - \alpha)] + \lambda [\alpha (1 - \beta) + (1 - \alpha)]]}$$

$$P[3 \text{ wrong} | H_3, s_3] = \frac{\lambda [\alpha (1 - \beta) + (1 - \alpha)] \alpha \beta \alpha (1 - \beta)}{\alpha (1 - \beta) \alpha \beta [(1 - \lambda) [\alpha \beta + (1 - \alpha)] + \lambda [\alpha (1 - \beta) + (1 - \alpha)]]}$$

$$P[3 \text{ right} | H_3, s_3] > P[3 \text{ wrong} | H_3, s_3]$$

Play $\{X\}$.

$$H_3 = \{X, M\}, s_3 = \{\emptyset\}$$

$$P[1 \text{ right} | H_3, s_3] = \frac{(1 - \lambda) [\alpha \beta + (1 - \alpha)] \alpha (1 - \beta) (1 - \alpha)}{(1 - \lambda) [\alpha \beta + (1 - \alpha)] \alpha (1 - \beta) (1 - \alpha) + \lambda [\alpha (1 - \beta) + (1 - \alpha)] \alpha \beta (1 - \alpha)}$$

$$P[1 \text{ wrong} | H_3, s_3] = \frac{\lambda [\alpha(1-\beta) + (1-\alpha)] \alpha \beta (1-\alpha)}{(1-\lambda) [\alpha \beta + (1-\alpha)] \alpha (1-\beta) (1-\alpha) + \lambda [\alpha(1-\beta) + (1-\alpha)] \alpha \beta (1-\alpha)}$$

$$P[1 \text{ wrong} | H_3, s_3] > P[1 \text{ right} | H_3, s_3]$$

Equivalent to the case of firm 2, $H_2 = \{X\}$, $s_2 = \{m\}$

Play $\{M\}$.

$$H_3 = \{X, X\}, s_3 = \{m\}$$

$$P[3 \text{ right} | H_3, s_3] = \frac{\lambda [\alpha(1-\beta) + (1-\alpha)]^2 \alpha \beta}{\lambda [\alpha(1-\beta) + (1-\alpha)]^2 \alpha \beta + (1-\lambda) [\alpha \beta + (1-\alpha)]^2 \alpha (1-\beta)}$$

$$P[3 \text{ wrong} | H_3, s_3] = \frac{(1-\lambda) [\alpha \beta + (1-\alpha)]^2 \alpha (1-\beta)}{\lambda [\alpha(1-\beta) + (1-\alpha)]^2 \alpha \beta + (1-\lambda) [\alpha \beta + (1-\alpha)]^2 \alpha (1-\beta)}$$

$$P[3 \text{ right} | H_3, s_3] > P[3 \text{ wrong} | H_3, s_3] \Leftrightarrow$$

$$(1-\alpha)^2 \alpha [\beta - (1-\lambda)] - 2\alpha \beta \alpha (1-\beta) (1-2\lambda) (1-\alpha) - \alpha \beta \alpha (1-\beta) \alpha (\beta - \lambda) \geq 0.$$

For $\lambda = \frac{1}{2}$, it simplifies to:

$$(1-\alpha)^2 - \alpha \beta \alpha (1-\beta) \geq 0.$$

A.1 - Optimistic firms

- Firm 3

$$H_3 = \{M, M\}, s_3 = \{x\}$$

Same as case before. Now $\lambda > \frac{1}{2}$.

A.2 Merger cascades

Table 1 - Probability of N merger decisions in the *Good* state of the world.

β, α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.55	0.0525	.1001	.1427	.1804	.2131	.2409	.2637	.2816	.2945	.3025
0.7	0.0679	0.1316	0.1911	0.2464	0.2975	0.3444	0.3871	0.4256	0.4599	0.49
0.9	0.0891	0.1764	0.2619	0.3456	0.4275	0.5076	0.5859	0.6624	0.7371	0.81

Table 2 - Probability of N merger decisions in the *Bad* state of the world.

β, α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.55	0.0425	0.0801	0.1127	0.1404	0.1631	0.1809	0.1937	0.2016	0.2045	0.2025
0.7	0.0279	0.0516	0.0711	0.0864	0.0975	0.1044	0.1071	0.1056	0.0999	0.09
0.9	0.0091	0.0164	0.0219	0.0256	0.0275	0.0276	0.0259	0.0224	0.0171	0.01

A.3 Probability of merger cascades

Proposition 3.5 *The probability of a merger cascade is higher under optimistic beliefs than under pessimistic ones.*

Proof: For $1 > \alpha > \frac{1-\sqrt{\beta(1-\beta)}}{1-\beta(1-\beta)}$ optimistic firms herd on the action to merge after history $\{MM\}$. Under optimistic beliefs, the probability of two consecutive mergers in the Good state is given by: $P[MM] = (\alpha\beta + 1 - \alpha)^2$. Under pessimistic beliefs, the probability of two consecutive mergers in the Good state is given by: $P[MM] = \alpha\beta(\alpha\beta + 1 - \alpha)$. Clearly, $(\alpha\beta + 1 - \alpha) > \alpha\beta$. The same is true for the Bad state of the world.

For $\alpha \leq \frac{1-\sqrt{\beta(1-\beta)}}{1-\beta(1-\beta)}$ optimistic firms need at least three consecutive mergers to trigger the cascade. Hence, $P[MMM\dots M] = (\alpha\beta + 1 - \alpha)^m$, $m \geq 3$ where m is the number of consecutive mergers. Under pessimistic beliefs, the probability of two consecutive mergers in the Good state is given by: $P[MM] = \alpha\beta(\alpha\beta + 1 - \alpha)$. The first expression is greater than the latter one if $m < \frac{\ln(\alpha\beta(\alpha\beta+1-\alpha))}{\ln(\alpha\beta+1-\alpha)}$. Note that both nominator and denominator are negative numbers.

Now, after an history of l mergers, herding (among optimistic firms) will start in round $l + 1$ if

$$l > \frac{\ln\left(\frac{(1-\lambda)\beta}{\lambda(1-\beta)}\right)}{\ln\left(\frac{(\alpha\beta+1-\alpha)}{(\alpha(1-\beta)+1-\alpha)}\right)}. \text{ Both ratios are greater than one. In this setting, } \beta > \frac{\alpha\beta+(1-\alpha)}{2-\alpha} \geq \lambda > \frac{1}{2} \text{ and } \alpha \leq \frac{1-\sqrt{\beta(1-\beta)}}{1-\beta(1-\beta)}.$$

We want to show that for these values of the parameters, $l < m$.

Take $\lambda \rightarrow \frac{1}{2}$ that maximizes l (the numerator is decreasing in λ) and denote the expression by l^{\max} .

$$l^{\max} = \ln\left(\frac{\beta}{(1-\beta)}\right) / \ln\left(\frac{(\alpha\beta+1-\alpha)}{(\alpha(1-\beta)+1-\alpha)}\right)$$

$m > l^{\max}$ if

$$|\ln(\alpha\beta(\alpha\beta+1-\alpha))| \ln\left(\frac{(\alpha\beta+1-\alpha)}{(\alpha(1-\beta)+1-\alpha)}\right) > |\ln(\alpha\beta+1-\alpha)| \ln\left(\frac{\beta}{(1-\beta)}\right)$$

The inequality is satisfied for $1 > \beta > \frac{1}{2}$ and $\frac{1-\sqrt{\beta(1-\beta)}}{1-\beta(1-\beta)} \geq \alpha > 0$.

A.4. Expected utility of the third player

The two critical situations are the following ones: $H_2 = [X, X]$, $s_3 = \{m\}$ and $H_2 = [X, M]$, $s_3 = \{x\}$. In the first case firm 3 plays $\{M\}$ (no herding) for

$$\left[(1-\alpha)^2 - 2\alpha^2\beta(1-\beta)\right][\beta - (1-\lambda)] + \alpha^2\beta(1-\beta)(1-2\lambda) \geq 0$$

and $\{X\}$ otherwise. In the second case firm 3 plays $\{X\}$ (no herding) for

$$\alpha \geq \frac{2\lambda - 1}{2[\beta - (1-\lambda)]}$$

and $\{M\}$ otherwise. When states are equally likely to occur, $\lambda = \frac{1}{2}$ and when the Bad state occurs with higher probability, $\lambda = \frac{1}{4}$, the latter constraint is not binding (condition on history $[X, M]$, x is always satisfied and the third player chooses $\{X\}$).

Consider the case $\lambda = \frac{1}{2}$. Expected utility of firm 3 is given by:

$$2EU_3 = \alpha(2\beta - 1)[3(1 - \alpha) + 2\alpha\beta\alpha(1 - \beta)] + \alpha^3 [\beta^3 - (1 - \beta)^3] \text{ if } \alpha \leq \bar{\alpha}$$

$$2EU_3 = \alpha(2\beta - 1)(1 - \alpha)(2 + \alpha) + 3\alpha\beta\alpha(1 - \beta)(2\beta - 1) + \alpha^3 [\beta^3 - (1 - \beta)^3] \text{ if } \alpha > \bar{\alpha}$$

$$\text{where } \bar{\alpha} = \frac{1 - \sqrt{\beta(1 - \beta)}}{1 - \beta(1 - \beta)}$$

For the relevant values of β those thresholds are as follows:

β	herding
0.55	$\alpha \geq .58701$
0.6	$\alpha \geq .59073$
0.7	$\alpha \geq .60677$
0.8	$\alpha \geq .6387$
0.9	$\alpha \geq .70212$

When the bad state occurs more frequently, $\lambda = \frac{1}{4}$, only high values of β are meaningful otherwise firms would not follow its own signals. Given the restriction on β , the no herding condition on history $H = [X, X]$, $s_3 = \{m\}$ is satisfied for any given value of α . Investment recommendations would be as follows:

Table 4 - Investment Recommendations when $\lambda = \frac{1}{4}$

β , Investment decision:	$\alpha = \text{information/ideas}$	$\beta = \text{precision}$
0.8	$0 < \alpha \leq .40095$	$.40095 < \alpha < 1$
0.9	$0 < \alpha \leq .51159$	$.51159 < \alpha < 1$

Note that the range of values under which firms should invest in α is substantially larger. When there is a higher probability that the state is Bad, firms should give priority to get signals because by merging without a signal firms are probably choosing the wrong action. Hence, if states occur with equal probability firms would tend to invest in private information precision, whereas if the state of the world is most likely adverse to merger activity firms would tend to invest in general business information (i.e., ideas).

Chapter 4

Strategic Effects in Merger Waves

Abstract: Last chapter has shown that merger waves can be generated by herding behavior when firms belong to different industries. This chapter shows that the herding mechanism holds when firms belong to the same industry. An example, using the Cournot framework, is able to combine the strategic effects and the herding mechanism. In the model, a decision to merge is risky given that it may bring efficiency gains or efficiency losses according to the realization of the state of the world. Firms receive signals about the likelihood of such states. It is found that herding can arise as an equilibrium of the Perfect Bayesian game for intermediate values of risk. The result is independent from sequentiality of the decisions and is entirely due to the information interplay.

4.1 Introduction

The five Historical Merger Waves are of very different natures. Whereas in the first and second movements (1887-1904 and 1916-1929, respectively) horizontal mergers were the most frequent type of deal, the merger wave of the sixties, (also called the Conglomerate Merger Wave), was characterized by pure conglomerates mergers. The merger wave of the eighties is also of diversifying type, but vertical mergers played also a relevant role. The current wave, with the peak in 2000, seems to come closer to the early predominance of horizontal merger deals, even though it is not possible to conclude about it yet.

The previous chapter shows that a herding mechanism can explain the phenomenon of merger waves when firms belong to different industries. Such finding may help to understand why there are unprofitable merger waves (the one of the sixties is a clear example), when firms generally have the choice of remaining independent. A legitimate question is then to ask whether the herding argument holds when firms belong to the same industry. This chapter attempts to shed some light on this question.

It is important to remind that the herding mechanism does not imply that a merger wave must be unprofitable. In the herding model used, if a merger cascade is triggered and the state of the world turns out to be good, then the result is a profitable merger wave. The idea of herding is somehow connected with failure because it would not easily arise as a possible explanation for the waves if there was no evidence of merger failures. The reason is that if merger waves would be generally profitable, some economic fundamentals or firms' strategic behavior would probably be considered the rational incentives for a decision to merge. There would be no room for a 'less noble' behavior like herding. In the model, if firms ignore private information to follow other firms' decision to merge and the state turns out to be favorable to merger activity, the herding firms benefit from a positive externality. If the opposite occurs and the state happens to be adverse to merger activity, then the decision makers who set up the informational cascade, impose a negative externality in all herding firms.

By intra-industry waves it is meant horizontal merger waves. If firms compete in the same market, merger decisions will affect not only the payoffs of merger participants but the payoffs of outsider firms. This can be due to the simple concentration effect, given that the number of operative firms is reduced, or through changes in production costs as a consequence of the merger that indirectly affects competitors. The analysis of the herding mechanism under the intra-industry setting requires that this payoff interaction is taken into account. Moreover, decisions can now be strategically taken in order to influence future players beliefs and actions. Not surprisingly, the equilibrium computations are demanding and overall tedious, even after all simplifying assumptions.

The main finding is that indeed, the herding mechanism holds in the case of strategic behavior between firms belonging to the same industry. Different decision rules characterize the equilibrium of the game depending on the level of risk associated to the action to merge. Herding emerges as a possible equilibrium strategy for intermediate risks that happens to be the largest interval of possible values of risk. By solving the game for myopic firms, it is shown that the merger wave is entirely due to the herding mechanism and does not depend on sequentiality of the decision making process.

4.2 Literature

This paper combines the herding literature with the industrial economics literature on horizontal mergers. A more complete description of the herding literature is given in the previous chapter. In the paragraph below, the major contribution of this paper to the classical works on informational cascades is clarified. The rest of this section deals with the theoretical models of horizontal mergers and explains the choice of our setup.

In the standard informational cascades models, a decision of one individual is heavily influenced by other individuals' decisions. Nevertheless, payoffs are, in general, totally independent from the other

players' choices. In other words, the payoff of player i is uniquely determined by the choice of player i and not affected by the choices of other players. In the Banerjee (1992) setting, randomly chosen decision makers have sequentially to choose to invest in a single asset. There are infinitely many assets but only one delivers a strictly positive return. For all the others, the return is zero. The payoff of the unique non-zero-return asset is independent from the number of actual investors. In the Bikhchandani *et al.* (1992) model, a sequence of individuals decides whether to adopt or reject some behavior. There is the same cost of adopting and the same gain of adopting for all players independently of the number of actual adopters. This paper follows again the Banerjee herding model, extending it in order to allow for strategic interaction among players. The discussion is now turned to the topic of horizontal mergers.

The effects and desirability of horizontal mergers were addressed by Salant *et al.* (1980) in the symmetric Cournot oligopoly with linear demand and cost functions. The authors concluded that any merger agreement consisting of fewer than 80% of the industry is not profitable. Perry and Porter (1985) and Deneckere and Davidson (1985) studied merges in a general setting with asymmetric firms and under Bertrand competition, respectively. One of the findings is that under quite general conditions there are no incentives to merge in quantity setting games (in the absence of efficiency gains) unless firms merge to monopoly since quantities are strategic substitutes, whereas in price setting games firms have incentives to merge since prices are strategic complements. Kamien and Zang (1990) deal with the mechanism by which mergers can be achieved. These authors introduce the concept of endogenous mergers and consider both the centralized and decentralized Cournot game. While the centralized setting corresponds to the common approach to mergers, in the decentralized setting, an owner who acquired several firms may wish to operate any number of them in competition with each other. The authors show that monopolization of an industry through acquisition by one owner of his rivals is limited. In other words, complete monopolization of the industry is possible only if it is initially small. Under this setting, in a latter paper (Kamien and Zang, 1991) the authors include a production cost advantage achieved only through a merger agreement. It is found that despite the double incentive to merge - reduction of cost and reduction of competition - complete or substantial partial monopolization of an industry with sufficiently many firms cannot be obtained in equilibrium. Finally, in their 1993 paper, the authors introduce the issue of sequentiality in the merger game. The question of whether the number of firms in the industry can be whittled down through successive rounds of mergers to a number at which complete monopolization is possible. The authors conclude that while sequential acquisition makes it easier to monopolize an industry, there may still be some limits to monopolization (only when there are three or fewer firms in the initial market structure). Barros (1998) analyses the behavior of asymmetric firms in an endogenous merger model. The model is able to predict which mergers are more likely to occur and supports the economic intuition that

asymmetries in merger participants should be higher when the initial market concentration is low. A paper by Fauli-Oller (2000) addresses directly the issue of sequentiality. In Cournot setting with cost asymmetries, the author shows that mergers raise the profitability of future mergers. The author also points out that the fact that early takeovers are cheaper than later ones, due to an increasingly concentrated market may trigger the merger wave. The analysis brings out that the combination of these two effects with the level of asymmetries plays a major role in generating waves.

The simple setup presented in this paper benefited greatly from the findings of the papers mentioned above. The Cournot model is chosen for its simplicity and for imposing strong obstacles to merger profitability as the literature has shown. The aim is not only to find mergers in equilibrium but to find mergers caused by herding behavior. The analysis developed in the 'myopic' section intends to isolate the effects of sequentiality and hence to confirm that results depend on the role of players' beliefs. Asymmetries through cost differentials are the stimuli to merger decisions in a market of initially identical firms. To keep the setting of the previous chapter (in order to make a fair comparison of results), firms are chosen at random to take an action, and therefore there is no endogenous modelling.

4.3 The model

In the basic model of chapter two, firms are assumed to have no influence on each other payoffs, (as they belong to different industries), so that herding can be isolated from other possible justification for the merger waves. The present chapter studies the herding argument when firms belong to the same industry. When allowing for strategic effects due to mergers between firms in the same industry, it is possible to evaluate the impact of both herding and concentration effects. Despite the fact that there are now relevant gains from being an outsider of the merging process, it is shown that herding behavior still holds for a wide range of parameters.

Consider an industry of six firms competing on quantities. For simplicity, demand is linear with intercept normalized to one and firms face constant marginal cost $c = \frac{1}{2}$. The fact that mergers are not profitable in a Cournot setting legitimizes firms' pessimistic beliefs. Hence, without a signal firms would never take a decision to merge. From the six firms in the industry, three are called randomly one after the other to choose between the actions merge $\{M\}$, or not to merge $\{X\}$. The other three firms have a passive role, they can be picked as a partner or otherwise remain independent. When a firm decides to merge, it chooses one partner that will accept any offer assuring it at least its opportunity cost. Before making their choice, firms observe all previous actions. In the second stage of the game, Nature reveals the state of the world. Finally, firms choose quantities. Firms can always set quantity to zero and exit the market if its profits would be negative under the current state of the world. Hence, in the last stage of the game, firms compete in quantities given the number of operative firms in the

market.

As in the basic model, firms may receive a signal, before choosing an action, that conveys information about merger profitability. A signal $\{m\}$ tells a firm that a merger enhances efficiency by reducing the cost of production. On the contrary, a signal $\{x\}$ tells a firm that a merger brings inefficiencies by increasing the cost of production. Let $\epsilon \in [0, 1]$ be the potential cost reduction from a merger in the good state of the world. With no merger, $\epsilon = 1$, and when the state of the world is adverse costs can increase by $\frac{1}{\epsilon}$. The signal $\{m\}$ indicates that (ϵc) is likely to be the marginal cost if the firm decides to merge whereas signal $\{x\}$ indicates that $(\frac{1}{\epsilon}c)$ is more likely to be the marginal cost if the firm decides to merge. The pair $(\epsilon, \frac{1}{\epsilon})$ tries to approximate the symmetry in payoffs we had in chapter three to allow for a reasonable comparison between inter and intra-industry merger waves. Before the game starts, players believe that states are equally likely to occur. Again, α is the probability of a firm receiving a signal, and β is the probability that the signal is right.

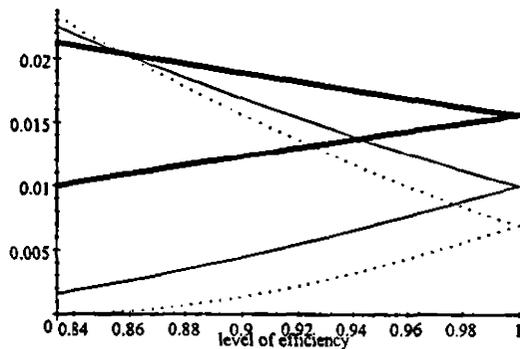
Payoffs are given in the table below. Recall that profits are distributed at the end of the game. The notation $\pi(M | MM)$ represents profits from merging given that the other firms also merged and hence there are three firms with equal costs in the post merger market. Payoff $\pi(X | MM)$ represents profits from not merging given that the other firms have merged. In this case there are 4 firms in the post merger market: the two merged ones with cost ϵc or $\frac{1}{\epsilon}c$, and the two non-merged ones with cost c . If no firm decides to merge, payoffs are not affected by the state of the world.

Table 1 - Payoffs of the Strategic Game for $\frac{5}{6} \leq \epsilon \leq 1$.

Payoffs	Good State (G)	Bad State (B)
$\pi(M MM)$	$(\frac{1-\epsilon c}{4})^2$	$(\frac{1-\frac{1}{\epsilon}c}{4})^2$
$\pi(M MX)$	$(\frac{1-3\epsilon c+2c}{5})^2$	$(\frac{1-3\frac{1}{\epsilon}c+2c}{5})^2$
$\pi(M XX)$	$(\frac{1-5\epsilon c+4c}{6})^2$	$(\frac{1-5\frac{1}{\epsilon}c+4c}{6})^2$
$\pi(X MM)$	$(\frac{1-3c+2\epsilon c}{5})^2$	$(\frac{1-3c+2\frac{1}{\epsilon}c}{5})^2$
$\pi(X MX)$	$(\frac{1-2c+\epsilon c}{6})^2$	$(\frac{1-2c+\frac{1}{\epsilon}c}{6})^2$
$\pi(X XX)$	$(\frac{1-c}{7})^2$	$(\frac{1-c}{7})^2$

Note that quantities are chosen after Nature has revealed the state of the world. Therefore, losses are bounded from below since firms can always set quantities to zero and become inoperative. The higher potential gains/losses, the most likely this is to occur. For $\epsilon < \frac{5}{6}$ a single merger in the market would make losses in the bad state of the world and if $\epsilon < \frac{3}{4}$ a merger would make losses even there is another merger in the market. In these cases, it is assumed that loss-making firms set quantities to zero and become inoperative. The case of $\epsilon < \frac{1}{2}$ is not studied given that it implies radical cost changes and would raise other questions by turning all firms inoperative in the bad state of the world.

Figure 1 - Payoffs from Merging : $\frac{5}{6} < \epsilon < 1$

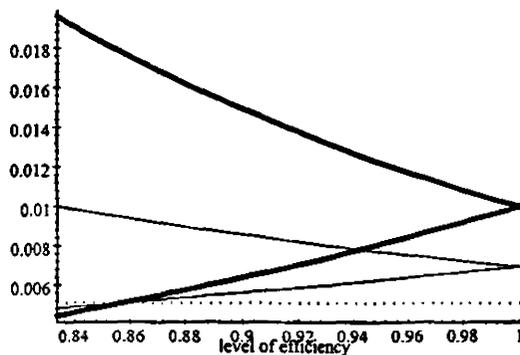


3 Mergers - thick line, 2 Mergers - thin line, 1 Merger - dotted line

Negative slope if state is Good, positive slope if state is Bad.

When ϵ is close to one payoffs in the Good and Bad state of the world coincide. The highest payoff corresponds to the the profit of a merged firm in a market with two other mergers, followed by the profit of a merged firm in a market with another merger and finally, the profit of a merged firm in a market with no other mergers. Note that such order is inverted for $\frac{6}{7} > \epsilon$.

Figure 2 - Payoffs from Not Merging : $\frac{5}{6} < \epsilon < 1$



2 Mergers - thick line, 1 Merger - thin line, No Mergers - dotted line

Negative slope if state is Bad, positive slope if state is Good.

Payoffs for smaller values of ϵ are given by the following tables:

Table 2 - Payoffs for $\frac{3}{4} \leq \epsilon < \frac{5}{6}$

Payoffs	Good State (G)	Bad State (B)
$\pi(M MM)$	$(\frac{1-\epsilon c}{4})^2$	$(\frac{1-\frac{1}{2}c}{4})^2$
$\pi(M MX)$	$(\frac{1-3\epsilon c+2c}{5})^2$	$(\frac{1-3\frac{1}{2}c+2c}{5})^2$
$\pi(M XX)$	$(\frac{1-5\epsilon c+4c}{6})^2$	0
$\pi(X MM)$	$(\frac{1-3c+2\epsilon c}{5})^2$	$(\frac{1-3c+2\frac{1}{2}c}{5})^2$
$\pi(X MX)$	$(\frac{1-2c+\epsilon c}{6})^2$	$(\frac{1-c}{5})^2$
$\pi(X XX)$	$(\frac{1-c}{7})^2$	$(\frac{1-c}{7})^2$

Table 3 - Payoffs for $\frac{1}{2} \leq \epsilon < \frac{3}{4}$

Payoffs	Good State (G)	Bad State (B)
$\pi(M MM)$	$(\frac{1-\epsilon c}{4})^2$	$(\frac{1-\frac{1}{2}c}{4})^2$
$\pi(M MX)$	$(\frac{1-3\epsilon c+2c}{5})^2$	0
$\pi(M XX)$	$(\frac{1-5\epsilon c+4c}{6})^2$	0
$\pi(X MM)$	$(\frac{1-3c+2\epsilon c}{5})^2$	$(\frac{1-c}{3})^2$
$\pi(X MX)$	$(\frac{1-2c+\epsilon c}{6})^2$	$(\frac{1-c}{5})^2$
$\pi(X XX)$	$(\frac{1-c}{7})^2$	$(\frac{1-c}{7})^2$

In Table 2, the exit of the single merger in case of a Bad state of the world, leaves a market of four equal firms with payoff: $(\frac{1-c}{5})^2$. In Table 3, the exit of two loss making mergers leaves a duopoly of identical firms with payoff: $(\frac{1-c}{3})^2$.

4.4 Equilibrium

Before proceeding, a brief description of the equilibrium refinement needed to solve this game is presented. The concept of subgame perfection is not applicable in games of incomplete information, even if players observe one another's actions at the end of each period. Given that players do not know the other's types, the beginning of a period does not form a well defined subgame until the players' posterior beliefs are specified, and so it is not possible to test whether the continuation strategies are a Nash-equilibrium. The Perfect Bayesian equilibrium concept extends subgame perfection to games of incomplete information. It results from combining the notion of subgame perfection, Bayesian equilibrium and Bayesian inference.

We now look for the Perfect Bayesian Equilibrium (PBE) of the game. In this setting of incomplete information with observable actions, a PBE consists of a pair (σ_i, μ_i) where σ_i is the behavioral strategy and μ_i is the system of posterior beliefs for each player i . Beliefs must be consistent with the strategies which have to be optimal given the beliefs. As it was said before, because of this circularity, the PBE can not be determined by backward induction, even when players move one at a time.

Definition 4.1 *A Perfect Bayesian Equilibrium is a set of strategies and beliefs such that at any stage of the game, strategies are optimal given the beliefs and the beliefs are obtained from equilibrium strategies and observed actions using Bayes' rule.*

Denote by θ_i the type of player i . More formally, each pair of strategies and beliefs (σ_i, μ_i) of each player i has to satisfy the four following conditions:

- (1) Sequential rationality: strategy $\sigma_i(\theta_i)$ is optimal for each type θ_i after every sequence of events.
- (2) Initial beliefs are correct for each player i .
- (3) Action-determined beliefs: only a players' actions influence the other players' beliefs about his type.
- (4) Bayesian updating: player i 's action at history h is consistent with the other players' beliefs about player i at h , given σ_i . Players' beliefs are derived using Bayes' rule from their observation of player i 's action.

In our game each player can receive signals $\{m\}, \{x\}$ or no signal $\{\emptyset\}$ and these constitute its possible set of types. The states of the world can be Good or Bad. Even with only three players (only three firms are called upon to choose an action) that have to decide between actions $\{M\}$ or $\{X\}$, the game tree ends up with 432 terminal nodes (6^3 times the two states of the world). We will then make the simplifying assumption that all three firms get a signal ($\alpha = 1$) which reduces the number of terminal nodes to 128 ($4^3 \times 2$). The game is solved for signal precision β equal to 0.75 (the parameterization used in the previous chapter for comparative statics) and marginal cost c equal to 0.5.

Proposition 4.1 *The following behavioral strategies are the Perfect Bayesian Equilibrium in pure strategies of the game for parameters: $\alpha = 1, \beta = 0.75, c = 0.5$.*

The optimal decision rules for each player are organized in four cases according to the risk involved in the action to merge. The first case, called 'High Risk', corresponds to the lowest interval of values of ϵ . The smaller is ϵ , the higher the potential gains can be, and consequently, the higher the potential losses. The fourth case, called 'Cournot' corresponds to almost no risk given that ϵ is very close to one.

The solution of the game is presented with the help of tables indicating the action of each player given possible past histories H_i and signals s_i . The whole derivation with the corresponding beliefs is can be found in the appendix.

- High Risk

$$0.5 < \epsilon \leq 0.60164$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	M

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	M	M
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	M	M	M	M
$s_3 = \{x\}$	X	X	X	X

For high potential efficiency gains, firm 1 decides to merge irrespective of its signal. Gains from merging are very high if the state of the world is Good. If the state of the world turns out to be Bad, and the other two firms did not choose the action to merge, firm 1 will exit the market. Given that action of firm 1 is uninformative, firms 2 and 3 will act according to their signals and no herding occurs.

- Intermediate Risk

$$0.60164 < \epsilon \leq 0.8\bar{3}$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	X

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	M	M
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	M	M	M	M
$s_3 = \{x\}$	M	X	X	X

Potential gains are not anymore so high to induce firm 1 to make a bet and then shut down in case of a Bad state. After observing history $\{M, M\}$, firm 3 believes that both firms 1 and 2 had a

signal to merge and therefore it will herd. By doing this firm 3 is forgoing the benefits from being an outsider in a highly concentrated market. Moreover, if the Bad state would occur, firms 1 and 2 could exit leaving firm 3 and its partner as duopolists. Nevertheless, firm 3 believes that the previous firms are more likely to be right and follows the herd.

$$0.8\bar{3} < \epsilon \leq 0.86847$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	X

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	M	M
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	M	M	M	X
$s_3 = \{x\}$	M	X	X	X

In this and the following tables, no firm will decide to exit the market in any contingency. Gains are not so high and therefore losses can not be that dramatic to lead firms to shut down. The herding argument from above holds. Firm 3 will again act irrespective of its signal and herd after history $\{M, M\}$ but will now also herd on the action $\{X\}$ after a history with no mergers.

$$0.86847 < \epsilon \leq 0.90552$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	X

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	M	X
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	M	M	M	X
$s_3 = \{x\}$	M	X	X	X

The herding behavior of firm 3 still holds in this table, but now firm 2 will also herd on the action not to merge. Given that potential gains from merging are now more modest, benefits from being an outsider start to be more relevant.

- Low Risk

$$0.90552 < \epsilon \leq 0.91959$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	X

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	M	X
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	M	M	M	X
$s_3 = \{x\}$	X	X	X	X

For low efficiency gains, the herding result can no longer be sustained. To be an outsider on a market with two mergers is now more attractive to firm 3, even though firm 3 believes the other firms had a signal to merge. Nevertheless, firm 3 will still follow its signals to merge unless it observes a history with no mergers.

$$0.91959 < \epsilon \leq 0.95638$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	X

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	M	X
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	M	X	X	X
$s_3 = \{x\}$	X	X	X	X

Payoffs from merging both in the Good or Bad states get significantly close. Profits from being an outsider get more relevant and therefore firm 3 will play always $\{X\}$ unless all three firms had a signal to merge.

• The Cournot Case

$$0.95638 < \epsilon \leq 1$$

FIRM 1	
$s_1 = \{m\}$	X
$s_1 = \{x\}$	X

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	X	X
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	X	X	X	X
$s_3 = \{x\}$	X	X	X	X

For insignificant efficiency gains the classical Cournot case is recovered. All firms would like to free ride on others decisions to merge and see the number of firms in the market reduced at no cost. Due to the free riding effect, no firm will merge in equilibrium.

Proof: One has to verify that each pair of behavioral strategies and beliefs for each player i satisfies the following conditions:

- 1) Sequential rationality: check that for each level of ϵ players have no incentives to deviate (can be found in the appendix)
- 2) Correct initial beliefs: before the game starts, all players believe that the states of the world (Good and Bad) are equally likely to occur.
- 3) Action-determined beliefs: only a player's action influences the other players beliefs about his "type" (the signal he received).
- 4) Bayesian updating: players update beliefs about other players using Bayes' rule. When the behavior of a certain player contradicts his strategy, a new conjecture about his type is formed, which is the basis for future Bayesian updating.

In the appendix it is verified that the pure-strategy equilibria satisfies such conditions for each level of efficiency gains/losses. ■

Corollary 4.1 *Merger waves also occur when allowing for strategic interaction among firms of the same industry for intermediate levels of efficiency gains/losses.*

Recall the herding mechanism of chapter three. After observing two previous mergers, the third firm would also decide to merge even with the signal supporting the non-merging strategy. The previous exercise has as main purpose to analyze the behavior of the third firm which receives signal $\{x\}$ and faces history of two past mergers $\{M, M\}$. Hence, the interest is to identify cases under

which herding can occur as an equilibrium strategy for each level of efficiency gains. A discussion of each equilibrium case is given below.

The first case, referred to as "High risk" is not able to generate herding on the action to merge because firm 1's action is uninformative. Potential gains are so high that firm 1 takes the risk to merge no matter its signal. It can be shown that a different type of herding emerges for values of ϵ smaller than 33%. For such high potential efficiency gains, all firms would decide to merge irrespectively of their own signal. The result relies on the fact that losses are bounded from below since a firm can always exit the market in case it took the wrong action.

The "Intermediate risk" case includes three ranges of ϵ values. The first table represents the equilibrium for higher efficiency gains where firms may exit the market if the Bad state occurs. The other tables illustrates herding when gains are moderately enough not to drive any firm out of the market. Firm 3 decides to merge after history $\{M, M\}$ even if it gets the opposite signal because it believes that both firms had a signal to merge. Given that signals have equal quality, firm 3 believes it must have received a wrong signal and decides to ignore it. This is particularly interesting in the strategic case because the outside option of firm 3 is now also very attractive. Recall that in the basic model the decision not to merge would involve no risk and no variation of profit. In the strategic case though, firm 3 now faces a highly concentrated market and is able to make high profits as an outsider, especially if the merging firms are wrong. Due to this reason, firm 3 will not follow the herd for higher values of ϵ (as shown in the first case).

The case of "Low risk" illustrates the situation in which firm 3's outside option facing history $\{M, M\}$ is so high compared to the benefits from merging, that firm 3 will decide not to merge unless it gets signal $\{m\}$. If it gets signal $\{x\}$, firm 3 will decide not to merge and will get significant profits from being an outsider even if it took the wrong decision.

Finally, the "Cournot case" illustrates the classical result that there are no incentives to merge in quantity setting games (with small or no efficiency gains) unless more than 80% of the market is involved on the merger (see Salant Switzer and Reynolds). All firms would benefit from a more concentrated market but being an outsider is always more advantageous than becoming a merging party. Hence, firms would free ride on others decisions to merge and not incur the costs of 'buying' a competitor. When efficiency gains are very small, firm 1 knows that if it decides to merge, no other firm will follow, even if they get signal $\{m\}$. The fact that payoffs from the Bad and Good state almost coincide makes it worse off if it is the only merger and therefore firm 1 will play always $\{X\}$. Due to the free rider problem, no firm will decide to merge both in the static Cournot game and in the sequential move game that is represented in this last case.

Probabilities

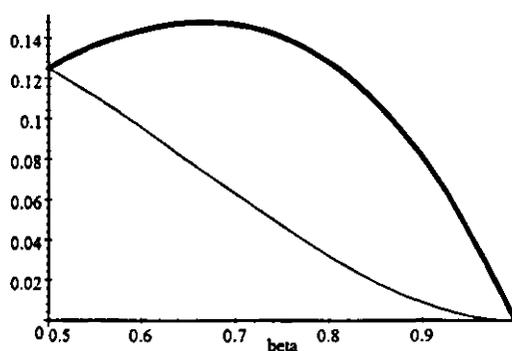
For the parameterization above, in which all firms receive a signal, the probability of a merger

cascade in the Bad state of the world must be smaller than in the inter-industry case studied in chapter three (corresponds to 5% for $\beta = 0.75$). It requires both the first and second firms to have the wrong signal and the third one to have the correct one. The figure below illustrates the probabilities of a merger cascade in both states.

$$P[\text{herding in M} \mid \text{Bad}] = (1 - \beta)^2 \beta$$

$$P[\text{herding in M} \mid \text{Good}] = \beta^2 (1 - \beta)$$

Figure 3 - Probability of herding in the action to Merge



Thick line: Good state, Thin line: Bad state.

As expected, the probability of a merger cascade when the state of the world is Good is non monotonic in β . (It is monotonic when the state of the world is Bad since its maximum occurs when $\beta = \frac{1}{3}$). For low signal precision, $\frac{1}{2} < \beta < \frac{2}{3}$, the probability that firms receive opposite signals is increasing in β whereas above this threshold such probability drops rapidly to zero as β approaches 1.

Tedious calculations to find the equilibrium for $\alpha = 0.6$ (the same parametrization of chapter three) show that a similar pattern of equilibrium exists, based on the levels of risk, and that herding occurs with positive probability. However, for some values of ϵ , in the region where all firms remain operative in the market, there is no equilibrium in pure strategies. We believe that the analysis of the game for $\alpha = 1$ simplifies significantly the calculations and provides a good intuition for what happens if $\alpha < 1$.

This exercise shows that the herding argument holds when allowing for strategic interaction among firms as long as potential efficiency gains or losses are moderate. Despite being artificial, the model is able to identify herding behavior in the Cournot framework which is adverse to merger activity. Note that the range of efficiency gains under which herding (in the action to merge) occurs with positive probability is the largest ($0.60164 < \epsilon < 0.90552$), and the most 'plausible' one given that gains are neither outstanding nor negligible. This simple model reveals the basic intuition, which we conjecture extends to a larger number of firms. With a higher number of firms, the range of potential gains and

losses compatible with herding would shift downwards (towards smaller values of ϵ) but the herding mechanism should arise with positive probability.

4.5 Myopic firms

In the Perfect Bayesian equilibria of the game, each firm takes into account the effect its action has on the follower decision makers. This corresponds to the condition 'action-determined beliefs' in which players' actions influence the other players' beliefs about their type. In this context, players who received a signal to merge are one type and players who received a signal not to merge are the other type. Hence, a firm may decide upon an action, not only based on its direct impact in payoffs, but in order to strategically induce a particular behavior of others.

In this section we drop the Perfect Bayesian Equilibrium conditions and we solve the game for myopic firms. Myopic firms are defined as follows. When called upon to choose an action, a myopic firm takes into account only the impact of its choice on the payoffs. It disregards the effect its decision may have on the choices of future decision makers. This corresponds to ignore the inference firms can make based on others actions and therefore to a setting of "less-intelligent" firms. By contrasting results with the ones obtained in the non-myopic example, we can disentangle the effects of sequentiality from informational ones.

4.5.1 Sequentiality

In the literature on mergers one can find a strategic explanation for merger waves that relies on sequentiality (Caves, 1991). Firms find it profitable to merge only if competitors also merge. In a recent paper by Fauli-Oller (2000) it is shown, in a simple Cournot setting, that previous mergers stimulate future mergers. Analogously, in a setting of endogenous mergers, Kamien and Zang (1993) analyze the impact of sequential acquisition on industry concentration. Therefore, it is relevant to ask whether the results from last section are due to herding behavior or sequentiality of the merger decisions. When facing a history of two past mergers, firm 3 can merge because it is on average profitable to do so, disregarding inference on other firms' signals. In order to check this the behavior of firm 3 is analyzed given past history but disregarding the information game. Firms observe actions of previous firms but are myopic. As before, firms get a signal $\{m\}$ or $\{x\}$ that is wrong with probability $(1 - \beta)$, states are equally likely to occur and $\beta = 0.75$.

Due to the fact that no inference is made about other firms' actions, the order of moves plays no role in this setting. For the range of values of ϵ under which all firms remain operative in the market ($\frac{\sigma_0}{\sigma_1} < \epsilon < 1$). Like in Fauli-Oller (2000), firms buying in the first place pay a lower price for their targets. Results can be summarized as follows:

- After signal $\{x\}$ play $\{X\}$
 - play $\{M\}$ if there are no past mergers and $\epsilon < 0.92937$
- After signal $\{m\}$:
 - play $\{M\}$ if there is one past merger and $\epsilon < 0.9196$
 - play $\{M\}$ if there are two past mergers and $\epsilon < 0.90552$
 - otherwise play $\{X\}$.

Note that for all levels of ϵ there will never be herding after history $H_3 = \{M, M\}$. Furthermore, the free riding effect seems to dominate for high values of ϵ such that firm 3 will not even follow a signal to merge. For the interval $\frac{5}{6} < \epsilon < 1$, the profitability of a merger is increasing in the number of mergers. Hence, a market with three mergers would yield the highest payoff per firm. Nevertheless, note that we obtain the opposite results about the incentives to merge. The higher the number of already existing mergers, the less likely the third firm is to merge (with signal $\{m\}$). The increase in the threshold of ϵ under which firm 3 plays M after a signal to merge is due to a sharp decrease in the outside option as we move from history $\{M, M\}$ to $\{X, X\}$. The benefits from being an outsider in a market with two mergers are so high that potential cost reductions to convince firm 3 to merge have to be higher than in any other case.

For intermediate values of ϵ , firms basically follow their own signals (see appendix for a complete description). It is interesting to note though, that when ϵ approaches its lower bound ($\frac{1}{2}$), it is optimal to merge when firms have a signal $\{x\}$ as long as there are no past mergers. This is due to the fact that for $\frac{1}{2} < \epsilon < \frac{5}{6}$ the pattern of merger profitability is reversed. Under this range of ϵ , a single merger in the industry is more profitable than two or three mergers¹. Clearly, a firm is willing to take this risk given the possibility to shut down in the case of a Bad state of the world.

Proposition 4.2 *Herding behavior never occurs if firms are myopic.*

Proof: From the discussion above and the tables in appendix.

4.5.2 Initial beliefs

One could also think that results depend heavily on the chosen parametrization $\beta = 0.75$ which assumes a high probability that signals are right.

It is easy to check that even if firm 3 bases its decision only on initial beliefs about the state of the world ($\frac{1}{2}, \frac{1}{2}$), the herding result is not recovered. The optimal strategy of firm 3 is summarized in the table below. Again, no inference is made about other players' decisions.

¹However, the ranking of industry profits does not change: the highest industry profit is achieved when there are three mergers, followed by two mergers and finally one merger.

Table 4 - Prior beliefs: $\mu(G) = \mu(B) = \frac{1}{2}$

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$ or $H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$\frac{5}{6} < \epsilon < 1$	X	X	X
$\frac{3}{4} < \epsilon < \frac{5}{6}$	X	M for $\frac{3}{4} < \epsilon < .77391$ X for $.77391 < \epsilon < \frac{5}{6}$	M
$\frac{1}{2} < \epsilon < \frac{3}{4}$	X	M	M

In the first range of ϵ , the ranking of payoffs is the standard one, where the higher the number of mergers in the industry, the higher the individual payoff of a merged firm. Firm 3 has no incentives to merge because the free rider effect dominates: being an outsider is more profitable than being a merging party. In the other two ranges of ϵ , the highest payoff is achieved when there is only one merger in the industry and the state of the world turns out to be the Good one. Clearly, a firm is betting on this high profit running the risk of leaving the market in case Nature reveals a Bad state. As expected, the higher the potential efficiency gains, the higher the incentives to merge. It is interesting to note though, that even for the highest potential efficiency gains, the free rider effect dominates after an history with two past mergers and the outcome $\{M, M, M\}$ can never be an equilibrium.

4.6 Conclusion

In the previous chapter it was found that herding behavior may generate merger cascades when firms belong to different industries. The present chapter analyzes whether the herding result holds when firms belong to the same industry.

The Cournot framework is chosen for its simplicity and for being adverse to merger activity. This adversity comes from the fact that outsider firms benefit more from a merger deal than the merger participants - the so called 'free-rider effect'. While when firms belong to different industries, a decision not to merge delivers a payoff of zero with certainty, in the intra-industry setting, it generally delivers a positive return. Firms have incentives to free ride on other firms' decisions to merge and achieve a more concentrated market at no cost. As a result, herding is more difficult to arise in this setting.

The Bad and Good states of the world are reinterpreted in terms of potential efficiency gains or losses from merging. In the Good state of the world, a merger brings cost reductions due to synergies or cost savings, and in the Bad state of the world, a merger increases the costs of the merging parties. This can be due to managerial problems, cultural costs of mixing different teams of workers or, incompatibilities in the production process. As before, firms get signals about these states of the world that are more likely to be right than wrong. For simplicity, the analysis is developed from an industry of six initially identical firms.

The Perfect Bayesian Equilibrium shows that for intermediate risks (defined as a moderate spread between potential gains and losses) herding behavior can arise. Moreover, the range of values under which herding behavior (in the action to merge) occurs with positive probability is the largest of the plausible set of risks. For intermediate risks, after observing two past mergers, it is optimal for the third firm to merge irrespectively of its signal, and despite the existence of a high outside option.

In the literature on horizontal mergers, the sequentiality of merger decisions is known to play an important role in providing incentives to merge. By solving the game under a myopic environment (in which actions of others firms do not influence beliefs of the decision maker), it is shown that merger waves can only arise when all firms get a positive signal. Facing an already concentrated market with two mergers, the third firm will always decide not to merge after a negative signal. Furthermore, for small potential gains, the third firm will even ignore a 'good' signal and choose not to merge given the two existing mergers. It is then clear that herding is entirely due to the information game and not to the sequentiality of decisions.

The analysis of the previous chapter is extended to horizontal mergers allowing for strategic effects. It is shown that the herding mechanism of the basic model holds for moderate efficiency gains. Therefore, herding on the action to merge may occur in both inter- and intra-industry setups. The fact that it may be optimal for firms to ignore private information and to follow the herd may lead to merger waves. This result helps to understand waves characterized by "widespread failure, considerable mediocrity and occasional successes"² that so much puzzle economists.

²Scherer and Ross (1990) on results of empirical studies about post merger performance of companies during the third and fourth merger waves.

4.7 Bibliography

Banerjee, A. V., "A Simple Model of Herd behavior", *Quarterly Journal of Economics*, 1992, vol. 107, pp. 797-817.

Barros, P. "Endogenous Mergers and Size Asymmetry of Merger Participants", *Economic Letters*, 1998, vol. 60, pp.113-119.

Bikhchandani, S., Hirshleifer, D., and Welch, I., "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *Journal of Political Economy*, 1992, vol. 100, pp. 992-1026.

Caves, R. E., "Mergers, Takeovers and Economic Efficiency", *International Journal of Industrial Organization*, 1989, vol. 7, pp. 151-174.

Deneckere and Davidson, "Incentives to form Coalitions with Bertrand Competition", *Rand Journal of Economics*, 1985, vol. 16, pp. 473-486.

Farrell J., and Shapiro, C., "Horizontal Mergers: An equilibrium analysis", *American Economic Review*, 1990, vol. 80, pp 107-126.

Fauli-Oller, R., "Takeover Waves", 2000, *Journal of Economics and Management Strategy*.

Fudenberg, D. and Tirole, J., *Game Theory*, 1995, The MIT Press, Cambridge, Massachusetts; London, England.

Kamien, M. and Zang, I., "The Limits of Monopolization Through Acquisition", *Quarterly Journal of Economics*, 1990, vol. 105, pp. 465-499.

Kamien, M. and Zang, I., "Competitively Cost advantageous Mergers and Monopolization", *Games and Economic Behavior*, 1991, vol. 3, pp. 323-338.

Kamien, M. and Zang, I., "Monopolization by Sequential Acquisition", *Journal of Law, Economics & Organization*, 1993, vol. 9, pp. 205-229.

Mueller, D. C., *The Determinants and Effects of Mergers: An International Comparison*, Ed. 1980, Oelgeschlager, Gunn & Hain, Cambridge, MA.

Mueller, D. C., "U.S. Merger Policy and the 1992 Merger Guidelines", *Review of Industrial Organization*, 1993, vol. 8, pp. 151-162.

Osborne, M. and Rubinstein, A., *A Course in Game Theory*, 1994, The MIT Press, Cambridge, Massachusetts; London, England.

Perry, M. and Porter, R., "Oligopoly and the Incentive for Horizontal Merger", 1985, *American Economic Review*, vol. 75, pp. 219-227.

Salant, S., Switzer, S. and Reynolds, R., "Losses from Horizontal Mergers: The effects of an exogenous change in industry structure on Cournot-Nash equilibrium", 1983, *Quarterly Journal of Economics*, vol. 98, pp 185-199.

Scherer, F.M., *Industrial Market Structure and Economic Performance*, 1980, 2nd Edition, Rand

McNally College Publishing Company, U.S.A.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

4.8 Appendix

A.1 - Checking equilibria

The purpose of this sections is to illustrate the mechanism to check for the equilibria claimed in the strategic model (intra-industry). Only one particular range of parameters is analyzed thoroughly given that it would be tedious to describe all cases. Finally, the payoff-tables used to deal with the complete range of ϵ are provided.

Consider the following interval under which there is herding and all firms are operative in equilibrium.

$$\frac{5}{6} < \epsilon \leq 0.86847$$

• Firm 1:

Suppose that firm 1 receives signal $s_1 = \{x\}$.

$$s_1 = \{x\}$$

The expected value from playing M is given by:

$$E(M) = \frac{(1-b)b}{k} \left(\frac{1-\epsilon c}{4}\right)^2 + \frac{b(1-b)}{k} \left(\frac{1-\frac{1}{2}c}{4}\right)^2 + \frac{(1-b)^2 b}{k} \left(\frac{1-3\epsilon c+2c}{5}\right)^2 + \frac{b^2(1-b)}{k} \left(\frac{1-3\frac{1}{2}c+2c}{5}\right)^2 + \frac{(1-b)^3}{k} \left(\frac{1-5\epsilon c+4c}{6}\right)^2 + \frac{b^3}{k} \left(\frac{1-5\frac{1}{2}c+4c}{6}\right)^2$$

$$\text{where } k = 2(1-b)b + (1-b)^2 b + b^2(1-b) + (1-b)^3 + b^3.$$

The odd terms correspond to probabilities if the state of the world is Good whereas even terms correspond to probabilities in the Bad state. A signal is wrong with probability $(1-b)$. By playing M , the possible histories are : MMM, MXM, MXX . Firm 1 believes that firm 3 will herd after observing $\{MM\}$ and therefore, the two first fractions are a product of two and not three terms. All terms are multiplied by the initial beliefs (probability $\frac{1}{2}$ for each state of nature) that cancel out with the denominator.

The expected value from playing X is given by:

$$E(X) = \frac{(1-b)b^2}{x} 2 \left(\frac{1-3c+2\epsilon c}{5}\right)^2 + \frac{b(1-b)^2}{x} 2 \left(\frac{1-3c+2\frac{1}{2}c}{5}\right)^2 + \frac{(1-b)^2 b}{x} 2 \left(\frac{1-2c+\epsilon c}{6}\right)^2 + \frac{b^2(1-b)}{x} 2 \left(\frac{1-2c+\frac{1}{2}c}{6}\right)^2 + \frac{(1-b)^2}{x} 2 \left(\frac{1-c}{7}\right)^2 + \frac{b^2}{x} 2 \left(\frac{1-c}{7}\right)^2$$

$$\text{where } x = 2(1-b)b^2 + 2b(1-b)^2 + (1-b)^2 + b^2.$$

Payoffs are multiplied by two because there are two potential merging firms remain independent. The possible histories are now: XMM, XMX, XXX .

For the chosen values of the parameters it is easy to check that $E(X) > E(M)$ and hence firm 1 will play X .

Analogously, for $s_1 = \{m\}$

$$s_1 = \{m\}$$

$$\begin{aligned} E(M) &= \frac{(b)^2}{D} \left(\frac{1-\epsilon c}{4}\right)^2 + \frac{(1-b)^2}{D} \left(\frac{1-\frac{1}{\epsilon}c}{4}\right)^2 + \frac{(1-b)(b)^2}{D} \left(\frac{1-3\epsilon c+2c}{5}\right)^2 + \frac{b(1-b)^2}{D} \left(\frac{1-3\frac{1}{\epsilon}c+2c}{5}\right)^2 + \\ &+ \frac{(1-b)^2 b}{D} \left(\frac{1-5\epsilon c+4c}{6}\right)^2 + \frac{b^2(1-b)}{D} \left(\frac{1-5\frac{1}{\epsilon}c+4c}{6}\right)^2 \\ D &= b^2 + (1-b)^2 + 2(1-b)^2 b + 2b^2(1-b) \\ E(X) &= \frac{b^3}{Q} 2 \left(\frac{1-3c+2\epsilon c}{5}\right)^2 + \frac{(1-b)^3}{Q} 2 \left(\frac{1-3c+2\frac{1}{\epsilon}c}{5}\right)^2 + \frac{(1-b)b^2}{Q} 2 \left(\frac{1-2c+\epsilon c}{6}\right)^2 + \frac{b(1-b)^2}{Q} 2 \left(\frac{1-2c+\frac{1}{\epsilon}c}{6}\right)^2 + \\ &+ \frac{b(1-b)}{Q} 2 \left(\frac{1-\epsilon}{7}\right)^2 + \frac{(1-b)b}{Q} 2 \left(\frac{1-\epsilon}{7}\right)^2 \\ Q &= (1-b)^3 + b^3 + b^2(1-b) + b(1-b)^2 + 2b(1-b) \end{aligned}$$

For the values of the parameters, $E(M) > E(X) \Rightarrow \text{Play } M$.

• Firm 2:

The second firm can observe two possible histories: $H_2 = \{M\}$, $H_2 = \{X\}$ and for each case might get a signal $\{m\}$ or $\{x\}$.

$$H_2 = \{M\}, s_2 = \{x\}$$

$$\begin{aligned} E(M) &= \frac{b(1-b)}{2b(1-b)} \left(\frac{1-\epsilon c}{4}\right)^2 + \frac{b(1-b)}{2b(1-b)} \left(\frac{1-\frac{1}{\epsilon}c}{4}\right)^2 \\ E(X) &= \frac{b(1-b)^2}{q} 2 \left(\frac{1-2c+\epsilon c}{6}\right)^2 + \frac{(1-b)b^2}{q} 2 \left(\frac{1-2c+\frac{1}{\epsilon}c}{6}\right)^2 + \frac{b^2(1-b)}{q} 2 \left(\frac{1-3c+2\epsilon c}{5}\right)^2 + \frac{b(1-b)^2}{q} 2 \left(\frac{1-3c+2\frac{1}{\epsilon}c}{5}\right)^2 \\ q &= 2b(1-b)^2 + 2(1-b)b^2 \\ E(X) &> E(M) \Rightarrow \text{Play } X. \end{aligned}$$

$$H_2 = \{M\}, s_2 = \{m\}$$

$$\begin{aligned} E(M) &= \frac{b^2}{b^2+(1-b)^2} \left(\frac{1-\epsilon c}{4}\right)^2 + \frac{(1-b)^2}{b^2+(1-b)^2} \left(\frac{1-\frac{1}{\epsilon}c}{4}\right)^2 \\ E(X) &= \frac{b^2(1-b)}{d} 2 \left(\frac{1-2c+\epsilon c}{6}\right)^2 + \frac{(1-b)^2 b}{d} 2 \left(\frac{1-2c+\frac{1}{\epsilon}c}{6}\right)^2 + \frac{b^3}{d} 2 \left(\frac{1-3c+2\epsilon c}{5}\right)^2 + \frac{(1-b)^3}{d} 2 \left(\frac{1-3c+2\frac{1}{\epsilon}c}{5}\right)^2 \\ d &= b^2(1-b) + (1-b)^2 b + b^3 + (1-b)^3 \\ E(M) &> E(X) \Rightarrow \text{Play } M. \end{aligned}$$

Analogously for $H_2 = \{M\}$ and $s_2 = \{x\}$, $s_2 = \{m\}$.

One finds that firm 2 will follow its own signal independently of the history.

• Firm 3:

The third firm can observe four possible histories: $H_3 = \{M, M\}$, $H_3 = \{M, X\}$, $H_3 = \{X, M\}$, $H_3 = \{X, X\}$ and for each case might get a signal $\{m\}$ or $\{x\}$.

$$H_3 = \{M, M\}, s_3 = \{x\}$$

$$E(M) = \frac{(1-b)b^2}{(1-b)b^2 + b(1-b)^2} \left(\frac{1-\epsilon\epsilon}{4}\right)^2 + \frac{b(1-b)^2}{(1-b)b^2 + b(1-b)^2} \left(\frac{1-\frac{1}{2}c}{4}\right)^2$$

$$E(X) = \frac{(1-b)b^2}{(1-b)b^2 + b(1-b)^2} 2 \left(\frac{1-3\epsilon c + 2\epsilon\epsilon}{5}\right)^2 + \frac{b(1-b)^2}{(1-b)b^2 + b(1-b)^2} 2 \left(\frac{1-3c + 2\frac{1}{2}c}{5}\right)^2$$

$$E(M) > E(X) \Rightarrow \text{Play } M.$$

This result implies that firm 3 will also play M when $H_3 = \{M, M\}$, $s_3 = \{m\}$.

$$H_3 = \{M, X\}, s_3 = \{m\}$$

$$E(M) = \frac{(1-b)b^2}{b(1-b)^2 + (1-b)b^2} \left(\frac{1-3\epsilon c + 2c}{5}\right)^2 + \frac{b(1-b)^2}{b(1-b)^2 + (1-b)b^2} \left(\frac{1-3\frac{1}{2}c + 2c}{5}\right)^2$$

$$E(X) = \frac{(1-b)b^2}{(1-b)b^2 + b(1-b)^2} 2 \left(\frac{1-2c + \epsilon\epsilon}{6}\right)^2 + \frac{b(1-b)^2}{(1-b)b^2 + b(1-b)^2} 2 \left(\frac{1-2c + \frac{1}{2}c}{6}\right)^2$$

$$E(M) > E(X) \Rightarrow \text{Play } M$$

For this range of parameters, this case is identical to the one of $H_3 = \{X, M\}$, $s_3 = \{m\}$.

$$H_3 = \{M, X\}, s_3 = \{x\}$$

$$E(M) = \frac{b(1-b)^2}{b(1-b)^2 + (1-b)b^2} \left(\frac{1-3\epsilon c + 2c}{5}\right)^2 + \frac{(1-b)b^2}{b(1-b)^2 + (1-b)b^2} \left(\frac{1-3\frac{1}{2}c + 2c}{5}\right)^2$$

$$E(X) = \frac{b(1-b)^2}{b(1-b)^2 + (1-b)b^2} 2 \left(\frac{1-2c + \epsilon\epsilon}{6}\right)^2 + \frac{(1-b)b^2}{b(1-b)^2 + (1-b)b^2} 2 \left(\frac{1-2c + \frac{1}{2}c}{6}\right)^2$$

$$E(X) > E(M) \Rightarrow \text{Play } X$$

For this range of parameters, this case is identical to the one of $H_3 = \{X, M\}$, $s_3 = \{x\}$.

$$H_3 = \{X, X\}, s_3 = \{m\}$$

$$E(M) = \frac{b(1-b)^2}{b(1-b)^2 + (1-b)b^2} \left(\frac{1-5\epsilon c + 4c}{6}\right)^2 + \frac{(1-b)b^2}{b(1-b)^2 + (1-b)b^2} \left(\frac{1-5\frac{1}{2}c + 4c}{6}\right)^2$$

$$E(X) = 2 \left(\frac{1-\epsilon}{7}\right)^2$$

$$E(X) > E(M) \Rightarrow \text{Play } X.$$

This result implies that firm 3 will also play X when $H_3 = \{X, X\}$, $s_3 = \{x\}$.

Relevant Thresholds

These thresholds mark a change in the behavior of at least one player. The solution in the paper presents a smaller number of intervals given that some changes in behavior are out of the equilibrium path. Given that a systematic check for all equilibria would be too long and tedious to include in the appendix, these thresholds allow the reader to check his computations with mine.

- $0.5 < \epsilon \leq 0.60164$
- $0.60164 < \epsilon \leq 0.71164$
- $0.71164 < \epsilon \leq 0.75$
- $0.75 < \epsilon \leq 0.8\bar{3}$
- $0.8\bar{3} < \epsilon \leq 0.86847$
- $0.86847 < \epsilon \leq 0.90552$
- $0.90552 < \epsilon \leq 0.91959$
- $0.91959 < \epsilon \leq 0.95638$
- $0.95638 < \epsilon \leq 1$

A.2 - Myopic Solution

$$0.92937 < \epsilon \leq 1$$

FIRM 1		FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_1 = \{m\}$	X	$s_2 = \{m\}$	X	X
$s_1 = \{x\}$	X	$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	X	X	X	X
$s_3 = \{x\}$	X	X	X	X

Due to myopia, firms will stop following signals earlier than in the Perfect Bayesian Equilibrium. Recall that in the PBE this decision rule is optimal for $0.95638 < \epsilon \leq 1$.

$$0.9196 < \epsilon \leq .92937$$

FIRM 1		FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_1 = \{m\}$	M	$s_2 = \{m\}$	X	M
$s_1 = \{x\}$	X	$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	X	X	X	M
$s_3 = \{x\}$	X	X	X	X

$$0.90552 < \epsilon \leq 0.9196$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	X

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	M	M
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	X	M	M	M
$s_3 = \{x\}$	X	X	X	X

For these three intervals, profits from being an outsider are so high that it is optimal for the third firm not to merger after a history of two past mergers irrespectively of its signal. The free rider effect dominates and unables the occurence of three mergers in the Good state of the world.

$$0.71513 < \epsilon \leq 0.90552$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	X

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	M	M
$s_2 = \{x\}$	X	X

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$H_3 = \{X, M\}$	$H_3 = \{X, X\}$
$s_3 = \{m\}$	M	M	M	M
$s_3 = \{x\}$	X	X	X	X

For intermediate levels of ϵ firms simply follow their own signals.

$$0.5 < \epsilon \leq 0.71513$$

FIRM 1	
$s_1 = \{m\}$	M
$s_1 = \{x\}$	M

FIRM 2	$H_2 = \{M\}$	$(H_2 = \{X\})$
$s_2 = \{m\}$	M	(M)
$s_2 = \{x\}$	X	(M)

FIRM 3	$H_3 = \{M, M\}$	$H_3 = \{M, X\}$	$(H_3 = \{X, M\})$	$(H_3 = \{X, X\})$
$s_3 = \{m\}$	M	M	(M)	(M)
$s_3 = \{x\}$	X	X	(X)	(M)

For the high risk interval of ϵ , a single merger in the Good state of the world yields the highest payoff. Firm 1 will decide to merge irrespectively of its signal and will shut down in case the Bad state occurs. Histories in brackets are off the equilibrium path.

