EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

EUI Working Paper **ECO** No. 2002/26

A Herding Approach to Merger Waves

Inês Cabral

BADIA FIESOLANA, SAN DOMENICO (FI)

Published in Italy in December 2002
European University Institute
Badia Fiesolana
I-50016 San Domenico (FI)
Italy

A herding approach to merger waves

Inês Cabral*
European University Institute
Via dei Roccettini 9, San Domenico di Fiesole
Italia
ines.cabral@iue.it

This draft: October 2002

Abstract

This paper 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. The herding result holds when the analysis is extended to allow strategic interactions among firms belonging to the same industry.

*I would like to thank Massimo Motta, Karl Schlag and Pedro Barros for very helpful comments and suggestions. All errors are mine.

1 Introduction

Despite the fact that mergers are a ubiquitous phenomenon of economic life, the reasons why firms do merge are yet only poorly understood.[1] Mergers pose a wide variety of research questions due to some stylized facts. World wide merger activity was outstanding during the last decade (peaking in 2000 with a volume of more than $3 trillion US dollars[2]), 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. Indeed, indicators tend to point to a poorer performance relative to the industry. Finally, and perhaps most interestingly, mergers appear to occur in 'waves': 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 merger wave which started in the beginning of the nineties has been referred to as the fifth merger wave.[3] 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 Europe.[4] Given such merger activity, one might expect that the merging strategy would significantly increase the value of the firms and improve their market performance. 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.

This paper claims that an informational cascade can be more responsible for triggering merger waves than relevant economic fundamentals. It 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 analysis is divided in two scenarios: one in which firms belong to different industries and hence the herding effect can be singled out, and another

[1] In 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.

[2] Thomson Financial Securities Data.

[3] The Economist, January 27th 2001, pp. 67.

[4] See Scherer (1980) for a complete description of historical merger waves.

one in which firms belong to the same industry allowing strategic interaction to play a role.

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, though, there are neither theoretical nor empirical studies applying the herding concept to merger activity.

Our model 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.

The first part of the paper deals with herding behaviour in the decision to merge when firms belong to different industries, i.e. ignoring strategic effects between them. 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. 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.

The second part of the paper then checks whether the herding argument holds when allowing for strategic interaction between firms. Instead of learning about the profitability of merging from other industries, it is assumed that firms learn from other firms' behaviour in the same industry. Competition is modelled using the Cournot framework with a small number of initially identical firms. Despite the relevant gains from being an outsider of the merging process, it is shown that herding behavior holds for a wide range of parameters, i.e. for intermediate values of risk, defined as a moderate spread between potential gains and losses.

The relevance of herd behaviour that emerges from these studies may legitimize 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".[5] Many merger deals in recent merger waves might simply have been triggered by other firms' decision

[5] The Economist, January 9th 1999, pp. 21.

to merge, and therefore are more likely to turn out to be a failure.

This paper is organized as follows. Section 1 is composed by the introduction and revision of the relevant empirical and theoretical literature. Section 2 presents the herding model assuming that there is no strategic interaction between firms that is then solved in Section 3. Section 4 explores the probabilities associated with merger cascades. Section 5 extends the model to a setting where firms belong to the same industry and compete in quantities. Finally, Section 6 concludes and discusses ideas for future research.

1.1 Mergers' performance

Evidence from historical merger waves is consistent with the following results. Within a window of less than four months around the date of announcement, 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 Industrial Organization literature on mergers' performance involves a larger time span (three to ten years between pre and post merger period). Results are more extreme and can be summarized as follows: "evidence shows...that the profits from merging companies generally decline after the mergers", Mueller (1993). Many case studies and anecdotal examples illustrate this point based on recent merger deals (e.g. about the DaimlerChrysler merger one can 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"[6]).

Concerning 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 as well as the value of the target firm, and 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.

[6] The Economist, October 14th-20th 2000, pp. 86.

1.2 Herding

The 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.

The reputational herding concept and its basic model where 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.

The 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).

This paper is placed in the first category of the theoretical works even though the idea comes from observed clustering behavior: the 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.

2 The model

This section presents the basic model of the paper which assumes that firms belong to different industries and therefore does not take into account strategic interaction between firms. It attempts to isolate the effect of factors, that concern the merger deal in itself and not competitive issues or features of the industries involved. The cultural cost of merging can be seen as one of those factors. 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 typically 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. 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.

Consider N risk neutral firms that maximize profits and do not belong to the same industry. 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 (i.e. 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 irrealistic if one takes into account that mergers' performance can only be evaluated in the medium-long run).

Note that the *Good* and *Bad* states of the world refer to merger specific properties that do not have to be correlated with states of the economy (expansions and recessions). The correlation addressed here is merger specific and not industry specific.

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 $j = 1, ..., N$, 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 modelisation.[7] 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 $P \cap F = \emptyset$.

Each firm may receive a signal, with probability α,[8] or not receive a signal. Denote s_j, with $j = 1, ..., N$, 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. 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 business magazines, and also through word-of-mouth. 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. 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 firms.[9] 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 receive 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

[7]Note 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.

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

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

when they receive 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 consultancy 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 partner.[10] 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 expand (without acquiring other firms), on the quality of the firm's market research department, on the cohesion among the different departments within the firm. These are some 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 observed actions to make their decision. Two cases will be studied with respect to priors: 'pessimistic' and 'optimistic' ones. 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.

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., $H_k = \{a_1, a_2, ..., a_{k-1}\}$. The information set of player k is composed by H_k, the signal s_k (if it exists) and the initial priors.

3 Solving the model

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.

[10] Choi 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'.

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 it.[11] 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 $E(M \mid m)$ is given by:

$$\begin{aligned} E(M \quad \mid \quad m) &= \frac{\lambda(\alpha\beta)}{P[m]} - \frac{(1-\lambda)\alpha(1-\beta)}{P[m]} > 0 \text{ for } \beta > 1-\lambda \\ P[m] &= \alpha(\lambda\beta + (1-\lambda)(1-\beta)). \end{aligned}$$

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

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

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

$$\begin{aligned} E(M \quad \mid \quad x) &= \frac{\lambda\alpha(1-\beta)}{P[x]} - \frac{(1-\lambda)\alpha\beta}{P[x]} < 0 \text{ for } \beta > \lambda \\ P[x] &= \alpha(\lambda(1-\beta) + (1-\lambda)\beta). \end{aligned}$$

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:

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

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

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 with 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:

$$\begin{aligned} P\left[1 \text{ right} \mid H_2, s_2'\right] &= \frac{\lambda\alpha\beta\alpha\left(1-\beta\right)}{\alpha\beta\alpha\left(1-\beta\right)} = \lambda \\ P\left[2 \text{ right} \mid H_2, s_2'\right] &= \frac{\left(1-\lambda\right)\alpha\beta\alpha\left(1-\beta\right)}{\alpha\beta\alpha\left(1-\beta\right)} = 1-\lambda. \end{aligned}$$

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,

$$\begin{aligned} P\left[1 \text{ right} \mid H_2', s_2\right] &= \frac{\left(1-\lambda\right)\left[\alpha\beta+\left(1-\alpha\right)\right]\alpha\left(1-\beta\right)}{P\left[H_2', s_2\right]}, \\ P\left[2 \text{ right} \mid H_2', s_2\right] &= \frac{\lambda\left[\alpha\left(1-\beta\right)+\left(1-\alpha\right)\right]\alpha\beta}{P\left[H_2', s_2\right]}. \\ P\left[2 \text{ right} \mid H_2', s_2\right] &> P\left[1 \text{ right} \mid H_2', s_2\right] \text{ for } \frac{1}{2} > \lambda > f\left(\alpha,\beta\right). \end{aligned}$$

[12] Other 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 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.

$$\begin{aligned}
P[1 \text{ right} \mid H_3, s_3'] &= \frac{\lambda\alpha\beta\left[\alpha\beta + (1-\alpha)\right]\alpha(1-\beta)}{P[H_3, s_3']}, \\
P[3 \text{ right} \mid H_3, s_3'] &= \frac{(1-\lambda)\,\alpha(1-\beta)\left[\alpha(1-\beta) + (1-\alpha)\right]\alpha\beta}{P[H_3, s_3']}, \\
P[3 \text{ right} \mid H_3, s_3'] &< P[1 \text{ right} \mid H_3, s_3'] \text{ for } \lambda > \frac{1-\alpha\beta}{2-\alpha}.
\end{aligned}$$

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 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 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,

$$\begin{aligned}
P[1 \text{ right} \mid H_4, s_4'] &= \frac{\lambda\alpha\beta\left[\alpha\beta + (1-\alpha)\right]\left[\alpha\beta + \alpha(1-\beta) + (1-\alpha)\right]\alpha(1-\beta)}{P[H_4, s_4']}, \\
P[4 \text{ right} \mid H_4, s_4'] &= \frac{(1-\lambda)\alpha(1-\beta)\left[\alpha(1-\beta) + (1-\alpha)\right]\left[\alpha\beta + \alpha(1-\beta) + (1-\alpha)\right]\alpha\beta}{P[H_4, s_4']} \\
P[3 \text{ right} \mid H_3, s_3'] &= P[4 \text{ right} \mid H_4, s_4'] \ .
\end{aligned}$$

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 kth 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

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

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

[13] In 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.

also play $\{X\}$. Little intuition is provided by this expression. However, considering the case of equal probability of the two states $\left(\lambda = \frac{1}{2}\right)$it simplifies to the following one,

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

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.

3.2 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 $s_3' = \{x\}$.

$$\begin{aligned}P[1 \text{ right} \mid H_3, s_3'] &= \frac{\lambda[\alpha\beta + (1-\alpha)][\alpha\beta + (1-\alpha)]\alpha(1-\beta)}{P[H_3, s_3']}, \\ P[3 \text{ right} \mid H_3, s_3'] &= \frac{(1-\lambda)[\alpha(1-\beta) + (1-\alpha)][\alpha(1-\beta) + (1-\alpha)]\alpha\beta}{P[H_3, s_3']},\end{aligned}$$

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

$$\begin{aligned} \alpha &\leq \overline{\alpha} \Rightarrow \text{No herding occurs, } a_3 = \{X\} \\ \alpha &> \overline{\alpha} \Rightarrow \text{Herding occurs, } a_3 = \{M\} \\ \text{with, } \overline{\alpha} &= \frac{1-\sqrt{\beta(1-\beta)}}{1-\beta(1-\beta)}. \end{aligned}$$

The function $\overline{\alpha}(\beta)$ 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 $\overline{\alpha}(\beta)$, 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 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, $\alpha \leq \overline{\alpha}$) 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 $\alpha's$ for which the result

[14] In 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.

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.

Recall that for α high, i.e. $\alpha > \overline{\alpha}$, 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 *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 $H_l = \{M, M, ..., M\}$. If $l-1$ firms have followed the first firm and decided to merge, and the lth firm has a different signal $s_l' = \{x\}$, the ratios of probabilities are given by

$$\begin{aligned} P\left[1 \text{ right} \mid H_l, s_l'\right] &= \frac{\lambda\left[\alpha\beta + (1-\alpha)\right]\left[\alpha\beta + (1-\alpha)\right]^{l-2}\alpha(1-\beta)}{P\left[H_l, s_l'\right]}, \\ P\left[l \text{ right} \mid H_l, s_l'\right] &= \frac{(1-\lambda)\left[\alpha(1-\beta) + (1-\alpha)\right]\left[\alpha(1-\beta) + (1-\alpha)\right]^{l-2}\alpha\beta}{P\left[H_l, s_l'\right]}. \end{aligned}$$

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 $H_l = \{M, M, ..., M\}$. 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 lth 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'.

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

4 Merger cascades

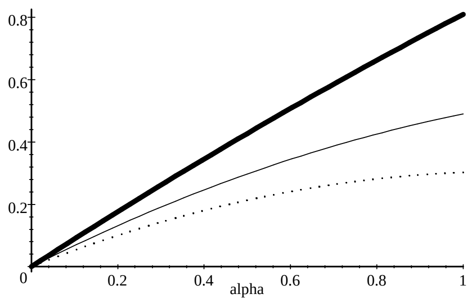
This section shows that the probabilities associated to cascades are quite substantial. Section 3 proved that history $H_3 = \{M, M\}$ under the pessimistic environment is enough to generate a merger cascade starting in firm 3. The probability of this event when the state of the world is *Good* is given by:

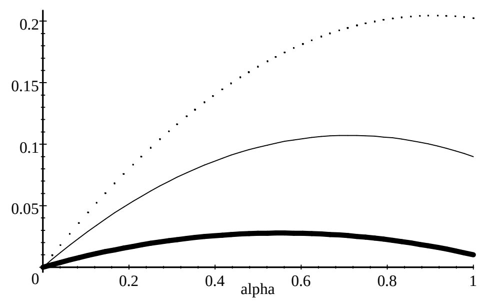
$$P[\{M, M\} \mid Good] = \alpha\beta(\alpha\beta + 1 - \alpha).$$

Recall that α is the probability of receiving information and β corresponds to the precision of that information. It is easy to check that $P[\{M, M\} \mid Good]$ depends positively on both parameters α and β, and can take values in the interval $[0, 1]$. Figure 1 below plots this function for given values of β. Now let the state be *Bad.* As expected, the probability of a merger cascade is smaller, even though not irrelevant for a wide range of the parameters:

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

Figure 1 (left) - $P[\{M, M\} \mid Good]$ and Figure 2 (right) - $P[\{M, M\} \mid Bad]$





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

In this case, the probability depends negatively on β and non monotonically on α, i.e., the derivative with respect to α is positive for $0 < \alpha \leq \frac{1}{2(1-\beta)}$, and negative for $\frac{1}{2(1-\beta)} < \alpha < 1$. Hence, the probability of a merger cascade in the wrong state decreases in the interval $\frac{1}{2(1-\beta)} < \alpha < 1$. For high levels of α and β it is not likely that the two first firms both receive a signal which is wrong.

However, the probability of a merger cascade in the bad state increases in α for $0 < \alpha \leq \frac{1}{2(1-\beta)}$. Given priors, firm 3 knows that firm 1 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 1. This enhances credibility of the first two players decision and makes it possible to generate herding behavior leading to the merger cascade.

To sum up, 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.

Proposition 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 negligeable 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'. ■

4.1 Complete 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
$0 \leq \lambda \leq \frac{1-\alpha\beta}{2-\alpha}$	$\emptyset$
$\frac{1-\alpha\beta}{2-\alpha} < \lambda < \frac{1}{2}$	$\{MM\}, \left\{ \begin{matrix} (MX)^k MM \\ (XM)^k MM \end{matrix} \right\}$
$\lambda = \frac{1}{2}$	$\{XMM\}, \{XXMM\}^*$
$\frac{1}{2} < \lambda \leq \frac{\alpha\beta+(1-\alpha)}{2-\alpha}$	$\{MM...M\}$
$\frac{\alpha\beta+(1-\alpha)}{2-\alpha} < \lambda \leq \frac{\alpha\beta(1-\alpha\beta)}{2\alpha\beta\alpha(1-\beta)+\alpha(1-\alpha)}$	$\{MM\}$
$\frac{\alpha\beta(1-\alpha\beta)}{2\alpha\beta\alpha(1-\beta)+\alpha(1-\alpha)} < \lambda \leq \beta$	$\{M\}, \{XM\}$
$\beta < \lambda \leq 1$	$\forall$

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

The probability of a cascade is a step function and not a continuous function in the parameter λ. This is due to the discreteness of herding decisions 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 in sustainable. Clearly, there may be a history where all firms decide to merge (no one gets a signal not to merge) but the chain of mergers is not due to herd behavior but to firms following its own signal. The threshold corresponds to the 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 $\left(\lambda < \frac{1}{2}\right)$ 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 specific histories are indicated because one can refine this range indefinitely to find herding mechanisms on the action to merge. 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. Histories that do not contain two consecutive 'no mergers' are able to generate a merger cascade.

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 $\{m\}$ 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 (belonging to different 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 queue.[16]

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

Proof: See appendix.

[16] A 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 $\alpha < \overline{\alpha}$.

4.1.1 Waves

The following paragraphs explore the link between prior beliefs and merger waves. 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, or changes in expected behavior of stock markets that affect investors' decisions, for example. 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, the same movement in stock prices can both intensify and refrain merger activity, showing that it can not be considered as an economic fundamental. The example below shows how stock prices can have totally opposite effects in merger activity in different points in time. About the most recent wave, peeking 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 to be possible triggers of 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".[18]

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 [...] acquisition soon, it will cost us more next year because of rising stock prices". To sum up, priors have to do with the expectations of decision makers, and not so much with fundamentals of the economy.

It is important to note that our model 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.

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

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

5 Strategic interaction

In the basic model 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. This section 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 competition effects.

Recall that in quantity setting games, like in the Cournot game, a firm has incentives to free ride on other firms decision to merge since benefits from being an outsider are higher than from being a merging party (see Salant, Switzer and Reynolds (1983), Farrell and Shapiro (1990)).

In this second part of the paper, the herding model is reinterpreted in the following way. A decision to merge is risky given that it may bring efficiency gains or efficiency losses depending on the realization 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 the option not to merge delivers the same payoff to all firms, in the intra-industry setting, 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 is likely to be more difficult to arise in this setting.

We consider an industry of six firms competing on quantities from which three are called at random and sequentially to choose between the actions to 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. For simplicity, demand is linear with intercept normalized to one and firms face constant marginal cost c. 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. When a firm decides to merge, it chooses one partner that will accept any offer as long as it covers at least its opportunity cost.

In stage one firms make their choices after observing all previous actions. 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 $\left(\frac{1}{\epsilon}c\right)$ is more likely to be the marginal cost if the firm decides to merge. The pair $\left(\epsilon, \frac{1}{\epsilon}\right)$ 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.

In the second stage of the game, Nature reveals the state of the world and 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.

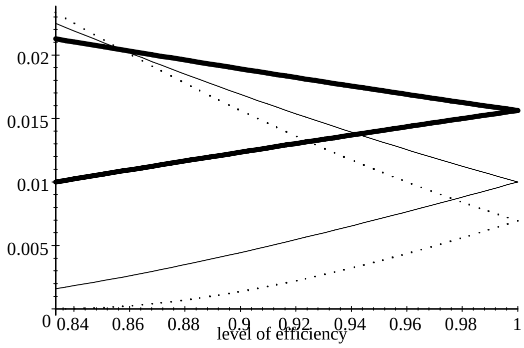
Payoffs are given in the table below. Recall that profits are distributed at the end of the game. The notation $\pi(M \mid MM)$ represents profits from merging given that the other firms also merged and hence there are three firms with equal costs in the post merger market. Payoff $\pi(X \mid MM)$ represents 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 4 - Payoffs of the Strategic Game for $\frac{5}{6} \leq \epsilon \leq 1$.

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

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 if 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 3 - Payoffs from Merging : $\frac{5}{6} < \epsilon < 1$

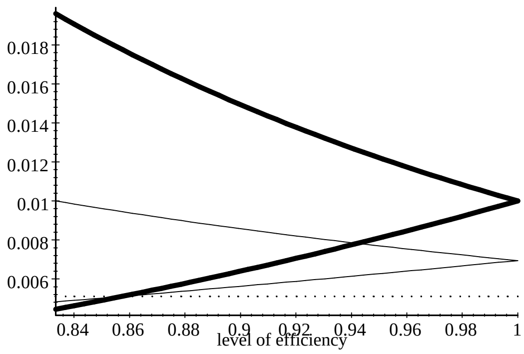


3 Mergers - thick line, 2 Mergers - thin line, 1 Merger - dotted line

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 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 4 - 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 5 (left) - Payoffs for $\frac{3}{4} \leq \epsilon < \frac{5}{6}$ and Table 6 (right) - Payoffs for $\frac{1}{2} \leq \epsilon < \frac{3}{4}$

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

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

In Table 5, the exit of the single merger in case of a Bad state of the world, leaves a market of four equal firms with payoff: $\left(\frac{1-c}{5}\right)^2$. In Table 6, the exit of two loss making mergers leaves a duopoly of identical firms with payoff: $\left(\frac{1-c}{3}\right)^2$.

5.1 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 each 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 notions 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 6 *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 firm 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 7 *The following behavioral strategies are the Perfect Bayesian Equilibrium in pure strategies of the game for parameters:* $\alpha = 1, \beta = 0.75, c = 0.5$.

Recall the herding mechanism of section two. After observing two previous mergers, the third firm would also decide to merge even with the signal supporting the non-merging strategy. The following exercise identifies cases under which herding can occur as an equilibrium strategy for each level of efficiency gains.

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.

In this case, there is no 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.

- *Intermediate Risk*

The "*Intermediate risk*" category includes three ranges of ϵ values. The first table represents the equilibrium for still high efficiency gains or losses where firms may want to exit the market if the Bad state occurs. The other tables illustrates herding when gains are moderate enough not to drive any firm out of the market.

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

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 however not anymore so high so as 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.

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

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 one 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.

In all cases, firm 3 decides to merge after history $\{M, M\}$ even if it gets a signal not to merge because it believes that both previous decision makers had signals 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).

- *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 "*Low risk*" case 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.

- *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

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 like to free ride on others decisions to merge since 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 firm 1 worse off if it is the only merger and therefore it 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.

Proof: See appendix.■

Corollary 8 *Merger waves may also occur when allowing for strategic interaction among firms of the same industry, for intermediate levels of efficiency gains/losses.*

Tedious calculations to find the equilibrium for $\alpha = 0.6$ 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 *per se* 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 maybe the most plausible one given that gains are neither outstanding nor negligeable. 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.

5.2 Sequentiality

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. When called upon to chose 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. By contrasting results with the ones obtained in the non-myopic example, we can disentangle the effects of sequentiality from informational ones.

In the theoretical literature on mergers a strategic explanation for merger waves that relies on sequentiality was put forward (Caves, 1991). Firms find it profitable to merge only if competitors also merge. In a recent paper by Fauli-0ller (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 made on other firms' behaviour.

The behavior of firm 3 is analyzed next 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 $\beta = 0.75$.

Since no inference is made about other firms' actions, the order of moves plays no role in this setting. Like in Fauli-Oller (2000), firms buying in the first place pay a lower price for their targets. The optimal decision rule can be summarized as follows:

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

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 $\{m\}$).

Contrasting with the non-myopic environment, 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 follow their own signals (see appendix). It is interesting to note though, that when ϵ approaches its lower bound $\left(\frac{1}{2}\right)$, it is optimal to merge when firms have a signal $\{x\}$ as long as there are no past mergers. This happens because 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 mergers.[19] 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 9 *Herding behavior never occurs if firms are myopic.*

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

6 Concluding remarks

This paper attempts to make sense out of two facts concerning merger activity. First, mergers tend to occur in waves and second, profits of merging companies often decline after the merger is completed. 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.

In the first part of the paper, Banerjee's herding model is adopted to explain decisions to merge in a setting of firms belonging to different industries. 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. Firms observe the choices of previous decision makers and may receive a signal about the likelihood of the states.

One of the main results is that a merger cascade under 'optimistic' beliefs occurs with higher probability than when firms hold pessimistic beliefs. However, a merger wave starts earlier under pessimistic beliefs than under optimistic ones. Intuitively, 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 trustworthy, two firms merging sequentially are enough to generate the cascade.

The analysis is then extended to the setting of firms belonging to the same industry, hence allowing for strategic effects. The Cournot framework is chosen for its simplicity and for being adverse to merger activity (due to the so called 'free-rider effect'). As a result, herding is more

[19] However, 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.

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. It is shown that for intermediate risks (defined as a moderate spread between potential gains and losses) herding behavior can arise despite the existence of a high outside option, and hence the herding result holds when allowing for strategic effects.

By solving the game under a myopic environment (in which actions of others firms do not influence beliefs of the decision maker), it is then clear that the herding mechanism is entirely due to the information game and not to the sequentiality of the decisions.

Herding on the action to merge may arise both across industries and when firms belong to the same industry. In both settings, under certain conditions, it may be optimal for firms to ignore private information and to follow the herd. This result may help understanding of merger waves characterized by "widespread failure, considerable mediocrity and occasional successes"[20] that so much puzzle economists.

6.1 Future work

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

The herding model developed in this paper is not explaining what triggers or stops a cascade. An exogenous change in the underlying beliefs is responsible for the reversion of the mechanism and the descending part of the wave. By endogenising the behavior of the underlying common beliefs the alternation of peaks and dips could be backed by players expectations, for example.

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 that may lead to failure. A test of the model would be to check whether mergers carried outside the historical waves are more profitable (on average) than merger deals settled during a merger wave, given a time span after the merger took place. Mergers settled

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

outside waves, could suggest more planning before the completion of the deal, would not be the result of herding and should therefore show better performance.

Another prediction of the model is that firms with more 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 suitable for some industries) one could test empirically such implication of the model.

The fact that the herding result holds when allowing strategic interaction among firms is an interesting result. Nevertheless, the model is oversimplified in order to solve for the perfect bayesian equilibrium. A natural extension would be to allow n firms in the strategic game and to solve the model without imposing restrictions in parameters α and β.

7 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

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

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

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

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

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

$$P[1 \text{ right} \mid H_2, s_2] > P[1 \text{ wrong} \mid H_2, s_2] \Rightarrow \{M\}$$

$$H_2' = \{X\}, s_2 = \{m\}$$

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 $f(\alpha,\beta) < \dfrac{1-\alpha\beta}{2-\alpha}$ for $(1-\alpha)^2 - \alpha\beta\alpha(1-\beta) \geq 0$. The ratio $\left(\dfrac{1-\alpha\beta}{2-\alpha}\right)$ 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 $\overline{\alpha}$ of the Optimistic firms section.

Repeat the exercise for $H_2' = \{X\}, s_2' = \{x\}$ and $H_2' = \{X\}, s_2 = \{\emptyset\}$ in which firm 2 will play $\{X\}$.

- Firm 3

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

In the text it is shown that for $\lambda > \dfrac{1-\alpha\beta}{2-\alpha}$ 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 $\lambda > \dfrac{1-\alpha\beta}{2-\alpha}$ is trivially satisfied for $\lambda = \frac{1}{2}$ since $\beta > \frac{1}{2}$.

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

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

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

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

$P[3 \text{ right} \mid H_3, s_3] > P[3 \text{ wrong} \mid H_3, s_3] \Rightarrow \{M\}$

For $H_3 = \{M, X\}, s_3 = \{x\}$ and $H_3 = \{M, X\}, s_3 = \{\emptyset\}$ firm 3 will play $\{X\}$.

Similarly, for $H_3 = \{X, M\}, s_3 = \{m\}$and $H_3 = \{X, M\}, s_3 = \{\emptyset\}$ firm 3 will play $\{M\}$ and for $H_3 = \{X, M\}, s_3 = \{x\}$, firm 3 will play$\{X\}$.

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

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

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

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

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

For $\lambda = \frac{1}{2}$, it simplifies to: $(1-\alpha)^2 - \alpha\beta\alpha(1-\beta) \geq 0$.

A.2 Probability of merger cascades

Proof of proposition 3.5: 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, $P[MMM...M] = (\alpha\beta + 1 - \alpha)^m, m \geq 3$ 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

$l > \frac{\ln\left(\frac{(1-\lambda)\beta}{\lambda(1-\beta)}\right)}{\ln\left(\frac{(\alpha\beta+1-\alpha)}{(\alpha(1-\beta)+1-\alpha)}\right)}$. Both ratios are greater than one. In this setting, $\beta > \frac{\alpha\beta+(1-\alpha)}{2-\alpha} \geq \lambda > \frac{1}{2}$ 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 $\lambda \to \frac{1}{2}$ 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.3 - Strategic effects : checking equilibria

Proof of Proposition 7:

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

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.

The following expressions verify that the pure-strategy equilibria satisfies such conditions for one particular range of parameters given that it would be tedious to describe all cases.

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 $s_1 = \{x\}$. The expected value from playing M is given by:

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

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:

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

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 $s_1 = \{m\}$:

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

$D = b^2 + (1-b)^2 + 2(1-b)^2 b + 2b^2(1-b)$

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

$Q = (1-b)^3 + b^3 + b^2(1-b) + b(1-b)^2 + 2b(1-b)$

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

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

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

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

$q = 2b(1-b)^2 + 2(1-b)b^2$

$E(X) > E(M) \Rightarrow$ Play X.

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

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

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

$d = b^2(1-b) + (1-b)^2 b + b^3 + (1-b)^3$

$E(M) > E(X) \Rightarrow$ Play M.

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 $\{x\}$.

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

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

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

$E(M) > E(X) \Rightarrow$ Play M.

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

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

$E(M) = \frac{(1-b)b^2}{b(1-b)^2+(1-b)b^2}\left(\frac{1-3\epsilon c+2c}{5}\right)^2 + \frac{b(1-b)^2}{b(1-b)^2+(1-b)b^2}\left(\frac{1-3\frac{1}{\epsilon}c+2c}{5}\right)^2$
$E(X) = \frac{(1-b)b^2}{(1-b)b^2+b(1-b)^2}2\left(\frac{1-2c+\epsilon c}{6}\right)^2 + \frac{b(1-b)^2}{(1-b)b^2+b(1-b)^2}2\left(\frac{1-2c+\frac{1}{\epsilon}c}{6}\right)^2$
$E(M) > E(X) \Rightarrow$ Play M
For this range of parameters, this case is identical to the one of $H_3 = \{X, M\}, s_3 = \{m\}$.

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

$E(M) = \frac{b(1-b)^2}{b(1-b)^2+(1-b)b^2}\left(\frac{1-3\epsilon c+2c}{5}\right)^2 + \frac{(1-b)b^2}{b(1-b)^2+(1-b)b^2}\left(\frac{1-3\frac{1}{\epsilon}c+2c}{5}\right)^2$
$E(X) = \frac{b(1-b)^2}{b(1-b)^2+(1-b)b^2}2\left(\frac{1-2c+\epsilon c}{6}\right)^2 + \frac{(1-b)b^2}{b(1-b)^2+(1-b)b^2}2\left(\frac{1-2c+\frac{1}{\epsilon}c}{6}\right)^2$
$E(X) > E(M) \Rightarrow$ Play X
For this range of parameters, this case is identical to the one of $H_3 = \{X, M\}, s_3 = \{x\}$.

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

$E(M) = \frac{b(1-b)^2}{b(1-b)^2+(1-b)b^2}\left(\frac{1-5\epsilon c+4c}{6}\right)^2 + \frac{(1-b)b^2}{b(1-b)^2+(1-b)b^2}\left(\frac{1-5\frac{1}{\epsilon}c+4c}{6}\right)^2$
$E(X) = 2\left(\frac{1-c}{7}\right)^2$
$E(X) > E(M) \Rightarrow$ Play X.
This result implies that firm 3 will also play X when $H_3 = \{X, X\}, s_3 = \{x\}$.

A.4 - Myopic Solution

$$0.92937 < \epsilon \leq 1$$

To play $\{X\}$ is the dominant strategy for all firms independent of past history or received signal. Due to myopia, firms will stop following signals earlier than in the Perfect Bayesian Equilibrium (where this decision rule is optimal for $0.95638 < \epsilon \leq 1$).

$$0.9196 < \epsilon \leq 0.92937$$

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

FIRM 2	$H_2 = \{M\}$	$H_2 = \{X\}$
$s_2 = \{m\}$	X	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	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 impedes the occurrence of three mergers in the Good state of the world.

$$0.715\,13 < \epsilon \leq 0.90552$$

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.

$$0.5 < \epsilon \leq 0.715\,13$$

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.

8 References

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.

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.

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.

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., "Takeover Waves", Journal of Economics and Management Strategy, forthcoming.

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.

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

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.

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

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.

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

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

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.

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.