Topics in Industrial Organization

PhD Thesis in Economics

Nuno Alvim

Thesis Supervisor: Prof. David Pérez-Castrillo

International Doctorate in Economic Analysis (IDEA)

Universitat Autònoma de Barcelona

Barcelona, Spain
Department of Economics and Economic History

June, 2012.

DEDICATION

I dedicate this tome to my family and to my wife Teresa. None of this would be possible without them.

Above all my gratitude to God from whom everything emanates.

ACKNOWLEDGMENTS

I would like to thank David Pérez-Castrillo for his valuable advice and his infinite patience. Without him all this work would not have been possible. I would also like to thank the rest of the faculty who have been present in this journey. All from whom I learned and all with whom I discussed, specially Inés Macho-Stadler, Pedro Rey Biel, Pau Olivella, Stefano Trento and Xavier Martínez-Giralt. My research and my skills greatly benefited from their presence. A final word to remember a great teacher that helped in the first baby steps: Toni Calvó-Armengol.

In my path I was also fortunate to have great professors in my undergraduate studies. Without the interest and passionate teaching of economics and economic intuition of Pedro Pita Barros and Diogo Lucena I would have probably not developed the interest and pleasure of thinking about economic mechanisms.

I would like to thank the people with whom I have worked side by side the last years. My colleagues from the IDEA program were essential to acquire tools, learn how to communicate ideas and clarify problems. Specially important were António Freitas, Bing Guo, Joaquin Naval, Xu Jing and Pau Pujolàs.

A special word for my friend and coauthor Tiago Pires who has the ability to construct challenges.

Finally I want to highlight the presence of many friends who made this journey a more pleasant one. My friends back in Estoril and in Lisbon who make me feel at home whenever I go there, my new friends in Barcelona that helped me construct a home in this city and specially my family, where home actually is.

Last but not least, I thank Fundação Ciência e Tecnologia, from the Portuguese Science Ministry, for trusting my skills and providing financial support throughout 4 years with the Grant SFRH/BD/36752/2007. Without it I could not have done this work.

TABLE OF CONTENTS

CHAPTER 1.

INTRODUCTION

Industrial Organization concerns the strategic interaction between firms. How do firms take decisions when the reactions of others are taken into account? In this dissertation I analyze three specific problems faced by firms. The goal is to better understand firms' decisions and their consequences in market competition. More specifically I analyze firms' cooperation decisions relating them with their market consequences. I also analyze the firms' entry choices in new markets.

The second chapter concerns the impact of R&D cooperations between firms on their decisions to merge. The third chapter analyzes the specific market of venture capital and the venture capitalists decisions to cooperate through co-investments. The fourth chapter takes a behavioral economics approach to a classical endogenous market entry timing problem. It analyzes the consequences of different beliefs about the state of demand on the firm's market leadership.

In chapter 2, I investigate the strategic interaction between research and development (R&D) cooperations and subsequent mergers. I present a theoretical model linking stable R&D cooperation networks with the decisions to merge, followed by the product market competition.

I find that mergers and R&D cooperations are complements. This means that there are mergers that only exist because firms construct certain R&D cooperation networks. If cooperations were not allowed, then firms would choose not to merge. I also find that there are cooperations among firms that would not exist if the firms did not anticipate in subsequent mergers.

I also find that cooperations with non-merging parties are the key links, contradicting the intuition that firms create links to substitute for a merger at a later stage. The intuition is based on the fact that the non-merging firm's cooperations create a market asymmetry in terms of technological efficiency. This asymmetry worsens the competition outside option of at least one of the merging firms. Hence these firms will become more willing to be apart of a merger. The consequence is that the non-merging firm gains veto power over the merger and will only enable it if the resulting merged company is not too competitive.

In chapter 3 I present a model that sheds light on how differently experienced Venture Capital firms compete and why they cooperate in the early stages of venture financing. I consider that the decision to cooperate is made by the financiers, but it must be approved by the owner of the project.

I find that there are welfare gains from cooperating because syndications of Venture Capitalists evaluate projects more effectively and provide more value-added services. When the entrepreneur is taken into account, Venture Capitalists cannot appropriate all the gains from a venture. The entrepreneur retains part of the project by keeping the competition outside option. Therefore I present a model of cooperation without collusion. I also find that projects with greater potential are more likely to be syndicated and that larger investments do not necessarily lead to more cooperation. I test these two implications in an empirical analysis and find that projects with both a larger potential and a larger investment are more likely to be syndicated.

Finally chapter 4 analyzes firms' decisions about the entry timing in markets.

Specifically, I analyze how these decisions are affected by firms' beliefs about the state of the world. A second contribution of the paper is to study whether firms have incentives to distort beliefs and to become optimistic about the state of demand.

The paper considers an endogenous timing model with incomplete information about demand. I show that with Bayesian firms there exists a unique perfect Bayesian equilibrium where firms with optimistic beliefs produce in the first period while firms with pessimistic beliefs only produce in the second period. I also find that when firms could choose to be overconfident they choose not to be. Nevertheless, they are weakly better by having the option to do so.

CHAPTER 2.

STRATEGIC INTERACTION BETWEEN R&D COOPERATIONS AND MERGING ACTIVITY IN A NETWORK FRAMEWORK

2.1 Introduction

Cooperation among firms has always been a controversial issue in economics. On the one hand, it may lead to an increase in technological efficiency, but on the other hand, it may reduce market competition, creating market inefficiencies. The European Union law protects firm cooperation, as long as it is conducted under the scope of Research and Development (R&D) activities. This, however, raises the question, "Is this always desirable?"

If firms change their technological efficiency due to R&D cooperation, one should reasonably expect changes in the final product market structure. In this paper, we address the question of the influence of R&D collaboration on firms' merger decisions. We present a theoretical model where firms create R&D cooperation, thereby forming an R&D network. Later, firms may decide to merge given the existent network. We question whether there are mergers that would not exist if firms did not cooperate. As noted by Hagedoorn and Lundan (2001), mergers and inter-firm cooperations have experienced parallel growth over time. We propose the possibility that there is some causality in this co-movement and investigate whether the increase in alliances may cause an increase in the merging activity.

In industries where R&D capabilities are key assets of firms, alliances may work as way to gain access to other companies' assets, and they are, therefore, expected to have an impact in the market structure. Cooperations may make a firm more attractive to others and increase the desirability to be a target of an acquisition. Moreover, cooperations may improve the bargaining position on a merger, hence making it more likely.

The automotive industry is an example of an industry where R&D is an important activity. R&D advances are key to the creation of new products as well as to the improvement of the production process. In this industry, one can observe many inter-firm alliances by which cooperating firms improve their production processes. Is it possible that these inter-firm alliances induce mergers in this industry? Take the case of the Renault/Nissan collaboration. These two companies signed a cooperation agreement in March 1999. In September 1999, Renault acquired Dacia Motors and in 2000 Renault acquired a large share in Samsung Motors. Our model supports the view that these two subsequent mergers may have been facilitated by the existence of the Renault/Nissan cooperation agreement. The rationale for the Renault/Dacia merger is that by cooperating, Renault and Nissan create asymmetries towards other competitors. This asymmetry makes the acquired firm more willing to be acquired because the competition, if no merger occurs, becomes more severe. In this case, R&D cooperations can be viewed as complements to mergers, as one favors the other.

The Renault and Samsung merger presents one more important feature. By the merger date, Nissan was cooperating with both Renault and Samsung. Hence, by the end of 1999, the situation was as follows. Nissan was the center of two bilateral cooperation agreements, one with Renault and another with Samsung Motors. Our model predicts either a merger between Renault and Samsung, which eventually took place, or the creation of another cooperation between Renault and Samsung. The determining issue is the cost of a merger relative to a cooperation. If the former is not too large relative to the latter, then a merger takes place. As Samsung Motors was looking for a buyer since 1998, one can speculate that the merging costs were not too great. A relatively low merging cost, when compared to the cooperation cost, and the cooperation structure between the companies may have been a determinant for the merger. The intuition is that if Renault and Samsung did not act, they would suffer a significant competitive disadvantage towards Nissan.

To study the influence of R&D collaborations on mergers, we present a model where three homogeneous firms form a network of R&D cooperations. Cooperation leads to a decrease in the marginal cost for cooperating parties, but it also entails a fixed cost. Cooperating on R&D improves the production process at the cost of, for instance, setting up a joint R&D team. After having decided on the cooperation alliances, two firms may decide to merge. A merger also entails a fixed cost and leads to an improvement in technological efficiency. A merger enables the new firm to make joint decisions on competition. In the last stage, the remaining firms compete in quantities.

The main finding of our paper is that the existence of R&D cooperations makes mergers more likely. We also find that the incentives to merge are only present if there exists enough asymmetry in the firms' non-merging outside options. Firms will only merge if the competitive disadvantage of at least one of the merging parties is great enough without merging.

In terms of cooperations, we also find that introducing the possibility to merge induces firms to create R&D networks that would otherwise not be stable. Therefore, anticipating a merger changes the incentives to cooperate. Finally, we note that to make a merger more likely, there is the need for a cooperation link with a non-merging party. This implies that the non-merging party must always anticipate and profit from the merger. If this was not the case, then the non-merging company would not create the necessary market asymmetry. In our illustration, it is plausible to think that Nissan anticipated the Renault/Dacia merger. In fact, we know that Nissan uses Dacia technology to produce its Qashqai model. In this sense, the intuition that firms cooperate to merge at a later stage is not confirmed by our model. Accordingly, we posit that mergers are not substitutes for cooperations.

Our work is related to three strands of literature: horizontal mergers, R&D cooperation networks and R&D cooperation and market competition. In their seminal paper, Salant et al. (1983) note that there is no incentive for a twofold merger when firms are homogeneous and there are no efficiency gains. Kerry and Porter (1985) introduce capacity to create some asymmetry. We provide a similar exercise but consider that the asymmetry arises from a previous stage where firms endogenously form R&D cooperations. We are, therefore, also somewhat consistent with the literature on R&D cooperation networks. We actually use the same structure as Goyal and Joshi (2001) for the R&D network formation game, but we add an interim merger stage. The authors[1] find that the only stable structures are the empty, the complete and the dominant group (where if a firm cooperates, then it cooperates with all cooperating firms) networks. In our paper, we contradict that result, as firms acknowledge that the

[1]Other authors who obtained a similar result are Okumura (2009), Goyal and Joshi (2003) and Billand and Bravard (2004).

market structure may change in the subsequent period. Furthermore, we observe that firms may exhibit a greater variety of R&D network structures.

A closely related problem regards the role of R&D cooperations on sustaining collusion. Martin (1995) and Cabral (2000) find that R&D may facilitate the maintenance of collusion, as it may be used as a punishment device in the event that a firm deviates. While their result is similar to ours in the sense that R&D cooperations change the final market competition, we extend it to the more extreme merger possibility. More closely related to our paper, Kabiraj and Mukherjee (2000) model the choice between independent or cooperative R&D. The authors find, consistent with our results, that the probability of a merger between cooperating firms is much lower than that of a merger between an innovative and a non-innovative firm. Thus, the intuition is that mergers and R&D cooperations are not substitutes for one another.

In section 2, we develop the model, and in section 3, we analyze the scope of the interaction. Section 4 presents a welfare analysis, and section 5 concludes the paper. All proofs are included in the appendices.

2.2 The model

The setup We consider a market with three homogeneous firms that decide with whom to cooperate on R&D. After deciding their R&D cooperation network, firms may decide to merge. In the final stage, the remaining companies compete a la Cournot in the final product market, and profits are realized.

An R&D cooperation consists of a link between two firms. The existence of a cooperation induces a reduction in the final product's marginal cost of γ_1. Cooperating also implies that each cooperating firm incurs a fixed cost f_{RD}. Similarly, a merger is also interpreted as a link between firms. Merging implies a fixed cost f_m and a marginal cost reduction of γ_2. If two cooperating firms decide to merge, then their R&D cooperation link is substituted by a merging link. Merging also implies that final product market decisions are made jointly.

In our model, firms form networks of cooperations. Let us clarify the network notation.

Let $\mathcal{N} = \{1, 2, 3\}$ be the set of ex-ante identical firms. The pairwise relation between the firms is captured by the binary variable $g_{i,j}$. Let $g_{i,j} = 1$ denote the existence of a relation in a network between firms i and j, and $g_{i,j} = 0$ the nonexistence of it. Note that $g_{i,j} = g_{j,i}$. As we have two distinct networks, let g^{RD} and g^m be the matrixes that describe all the existent pairwise relations of all firms in $\mathcal{N}$, the former being the R&D cooperations and the latter the merging relations. Let also $g + g_{i,j}$ denote the replacement, in network g, of link $g_{i,j} = 0$ by $g_{i,j} = 1$, and $g - g_{i,j}$ denote the converse, i.e., the replacement of link $g_{i,j} = 1$ by $g_{i,j} = 0$. As we have three firms, there are only four types of possible network structures: the complete network, where $g_{i,j} = 1$ for all i and j; the empty network, where $g_{i,j} = 0$ for all i and j; the star network, where there is a unique firm i with $g_{i,j} = 1$ for all j and the other firms only have a link with the central one; and finally the dominant network where $g_{i,j} = 1$ only for companies i and j.

We adopt the notation that upper case letters refer to companies that make decisions in the final product market. A company may be an individual firm or two merged firms. Firms are denoted with lower case letters. In the case that no merger occurs, the companies coincide with firms.

We consider that the competition authority only allows for a twofold merger to prevent a monopoly situation. Therefore, we exclude the grand coalition of this game to simplify the analysis and to highlight the interactions between collaborations and mergers. In the last stage, each of the companies, either two or three, seeks to maximize profits by facing a linear demand $p(Q) = \alpha - \sum_I q_I$, where q_I is the quantity produced by each company.

The original marginal cost is γ_0. If, in the last stage, there are still three firms, i.e., there was no merger, then each of them has the following cost function:

$$C_I\left(g^{RD}, q_I\right) = \left(\gamma_0 - \gamma_1 n_I\left(g^{RD}\right)\right) q_I + n_I\left(g^{RD}\right) f_{RD}, I = 1, 2, 3$$

where $n_I\left(g^{RD}\right)$ is the number of neighbors in the R&D network (either zero, one or two).

If some merger occurred, we assume that production is allocated to the firm with the lowest marginal cost. Hence, the cost functions, taking into account the possibility of a merger, are:

$$\begin{aligned} C_I\left(g^{RD}, g^m, q_I\right) &= \left(\gamma_0 - \gamma_1 n_I\left(g^{RD}, g^m\right) - \gamma_2 m_I\left(g^m\right)\right) q_I + \\ &\quad n_I\left(g^{RD}, g^m\right) f_{RD} + m_I\left(g^{RD}, g^m\right) f_m, \end{aligned}$$

where $n_I\left(g^{RD}, g^m\right) = \sum_j g_{ij}^{RD}\left(1 - g_{ij}^m\right)$ are the remaining R&D cooperations, after being substituted by mergers. The number of merging partners is $m_I\left(g^m\right)$. In our setting, this number is either 0 for no merger or 2 for two-firm merger.

Throughout the paper, we will consider that $f_{RD} \leq f_m$. This assumption means that an R&D project is not more costly than a merger. If the former consists of establishing a team, then the latter is joining two whole companies. We also consider that $0 < \gamma_1 < \gamma_2$ reflecting that if there is some possible cost reduction when firms cooperate, then, when joined, firms should be able to do at least as well in terms of production cost efficiency. In joint R&D projects, one should expect moral hazard and free riding problems, which should not exist, or at least should have less impact, within a single firm. Finally, to ensure that marginal costs are positive and that firms always produce a positive quantity, we place the following restrictions on parameters $\gamma_0 - \gamma_1 - 2\gamma_2 > 0$ and $\gamma_2 < \frac{\alpha - \gamma_0}{2}$.

<u>The market competition stage</u> We will use Subgame Perfect Nash equilibrium as the relevant equilibrium concept of the whole game and, hence, solve it by backward induction. We start by analyzing the product market competition stage. With the marginal cost and number of competing companies determined in the first two stages, companies compete in quantities in the last stage. By solving the Cournot game, we obtain the following equilibrium quantities for each company I:

$$q_I = \frac{\alpha - N\left(g^{RD}, g^m\right) c_I\left(g^{RD}, g^m\right) + \sum_{J \neq I} c_J\left(g^{RD}, g^m\right)}{N\left(g^{RD}, g^m\right) + 1},$$

where $N\left(g^{RD}, g^m\right)$ is final number of companies competing in the market (either 2 or 3) and $c_I\left(g^{RD}, g^m\right)$ is the resulting marginal cost of company I. The corresponding profits of the remaining companies will, hence, be

$$\begin{aligned}\Pi_I &= \left(\frac{\alpha - N\left(g^{RD}, g^m\right) c_I\left(g^{RD}, g^m\right) + \sum_{J \neq I} c_J\left(g^{RD}, g^m\right)}{N\left(g^{RD}, g^m\right) + 1}\right)^2 \\ &\quad - n_I\left(g^{RD}, g^m\right) f_{RD} - m_I\left(g^m\right) f_m.\end{aligned}$$

These profits are net of the link costs, f_{RD} and f_m, but must still be shared among the participants of the merger, if one occurs.

The merger stage As previously noted, in this stage, firms only decide whether to merge, and they will do so anticipating the competition in the following stage and taking the R&D network as given.

We assume that firms follow an exogenous sharing rule, namely, the equal sharing of surplus. This rule attributes to each of the participants of the merger half of the net surplus realized by the merged company, discounting payoffs that the participating firms would realize by themselves. This choice is made for tractability. We make one more assumption for this stage such that when multiple equilibria arise, we assume that there is an exogenous rank between the possibilities of merger, thus allowing us to break the multiplicity of equilibria. In any case, if there are incentives for some specific pair of firms to merge, these incentives prevail.

This stage is formalized as a second network game, but it could also be viewed as a coalition formation game, as the structure of the network does not influence the result. Let π_i^{ij} be the final profits obtained by firm i when merging with j and, consequently, sharing market profits Π_I, where $I = ij$. If no super-index is presented, then no merger occurred and, hence, $\pi_i = \Pi_I$. The stability concept that we use throughout the paper is based on the concept of pairwise stability, which will formally be defined in the R&D stage. In this case, where only one merger is allowed and where the no merger case is the starting point, stability is tantamount to the conditions highlighted by Horn and

Persson (2001). For i to merge with j, the following must be satisfied:

1. $\pi_i^{ij} + \pi_j^{ij} \geq \pi_i + \pi_j$ and

2. $\pi_i^{ij} \geq \pi_i^{ik}$ when $\Pi_I \geq 0$ for $I = ik$ and

3. $\pi_j^{ij} \geq \pi_j^{jk}$ when $\Pi_I \geq 0$ for $I = jk$ for all i, j, k

and for no merger to be stable, we require that

$$\pi_i^{ij} + \pi_j^{ij} \leq \pi_i + \pi_j \text{ for all } i, j.$$

For each of the four possible R&D network structures (empty, one cooperation only, one center firm collaborating with the other two and complete network), the decision to merge is different for a given pair of fixed costs.

There are two types of mergers: mergers between cooperating firms and mergers between non-cooperating firms. We analyze these two types of mergers for each R&D network.

For a given marginal cost reduction due to a merger, a merger between non-cooperating firms will be profitable if the merger fixed cost is small enough. This threshold is independent of the cooperating fixed cost. This is true for all possible networks (except the complete network, where all firms are cooperators). Firms face the trade-off between the merging cost advantage and competition reduction and the merging fixed cost and the sharing of the joint profits.

The merging of cooperating firms is profitable if the merger cost relative to the cooperating cost is small enough. In this case, firms analyze how great the extra fixed cost of moving from a cooperation to a merger is. Therefore, the merging cost threshold is dependent on the cooperation cost. The trade-off that firms face is similar to that of the non-cooperators in the merger. The difference is that now firms evaluate the merger relative to a cooperation and take both the relative costs and the relative marginal cost reductions into account.

Thus, the question now is how the R&D structures influence the merging decision. Is it more likely that cooperators merge in a market with only one R&D cooperation or in a market with full cooperation? Additionally, what happens between non-cooperators? Do R&D cooperations always facilitate mergers or are they also able to prevent them?

For the remainder of the paper, we adopt the following notation: if threshold f_m^x depends on f_{RD}, it will be denoted $f_m^x(f_{RD})$, and otherwise just f_m^x. The superscript denotes which is the underlying R&D network such that $x \in \{e, d, s, c\}$, where the options are empty, dominant, star or complete.

There is one special case where all firms are always willing to merge.

Proposition 2..1 *For all cooperation fixed costs and for all merging fixed costs such that $f_m \leq f_m^e$, no R&D network can prevent a merger.*

If merging is sufficiently inexpensive, such that when no cooperations exist, firms want to merge, firms will always merge regardless of the R&D cost and R&D network. This means that R&D cooperations cannot be considered as substitutes for mergers. If firms want to merge, then cooperations will not deter them. The underlying rationale is that if it is profitable for firms to merge, the path taken to get to the merger is irrelevant, as ex-ante, it still is profitable to merge. This does not mean that the R&D

cooperations are neutral, as they give firms different bargaining positions when merging and, therefore, have some consequences even though the final market outcome does not change. As such, the R&D cooperations can only influence the distribution of profits.

Lemma 2..1 *The merging fixed cost above which non-cooperating firms find merging unprofitable, is greater in the star than in the dominant network, which is, in turn, greater than in the empty network, i.e.,* $f_m^e < f_m^d < f_m^s$.

The scope to merge is enhanced due to the existence of R&D networks. The intuition stems from the difference in profits that firms can achieve when merging. There are two contradictory effects produced by the R&D network structures. Favoring the possibility to merge is the effect that cooperations have on final profits, that is, more cooperations between the final merged company with the non-merging firm imply greater joint profits and, hence, greater incentive to merge. Therefore, the star network may still induce a merger even if the dominant network would not do so. After a star network leads to a merger, the merged company has two links with the firm that stayed out of the merger, whereas a dominant R&D network would leave the merged company with only one link. The same holds for the comparison of the dominant network with the empty network.

The second effect, which weakens the previous effect, and one that may be created by the R&D network is the asymmetry that cooperations create in the market. The sum of the profits of independent asymmetric merging firms is less than the sum of the profits of symmetric firms. This effect reduces the incentive for non-cooperators in the dominant network structure to merge. The latter effect is always less than the former, and therefore, we conclude that R&D cooperations favor mergers among non-cooperating firms.

Lemma 2..2 *The merging fixed cost above which cooperating firms find merging unprofitable is greater in the complete network than it is in the star network, which, in turn, is greater than that in the dominant network, i.e.,* $f_m^d(f_{RD}) < f_m^s(f_{RD}) < f_m^c(f_{RD})$.

The result and the intuition are similar to the previous lemma. The existence of R&D cooperations increases the set of possible mergers. More cooperations lead to more profits under a merger, and the possible asymmetries between merging firms decrease the incentive to merge. In this lemma, unlike the previous one, the thresholds depend on the R&D fixed cost. This is because merging firms have already incurred costs in the R&D fixed cost, and they evaluate the merger profitability in terms of the additional cost that a merger would represent. The R&D cooperations would be transformed into a merger, and thus, only the additional costs would have to be bared.

Proposition 2..2 *The existence of R&D cooperations increases the set of merging costs for which firms are willing to merge.*

This proposition follows immediately from the two previous lemmas. The set is enhanced in two ways: through the marginal cost reductions that increase the incentive to merge and through the decrease in their relative cost that a substitutable R&D cooperation may represent.

It is also important to note that an increase of the marginal cost reduction due to a merger always leads to an increase in all thresholds, as it only represents an increase in the profits of merging. The same is not true for an increase in the marginal cost reductions due to an R&D cooperation. In a merger between non-cooperating firms, all thresholds are increasing in γ_1. The increase in the marginal cost reductions amplifies the positive effect of the merging profit more than it increases the negative effect of asymmetry. In the merging of two cooperators, the same is not true, as an increase in γ_1 also increases the value of the substituted R&D cooperation. In the most extreme case, where there is only one cooperation that is substituted by a merger, an increase in the R&D marginal cost reduction decreases the incentive to merge.

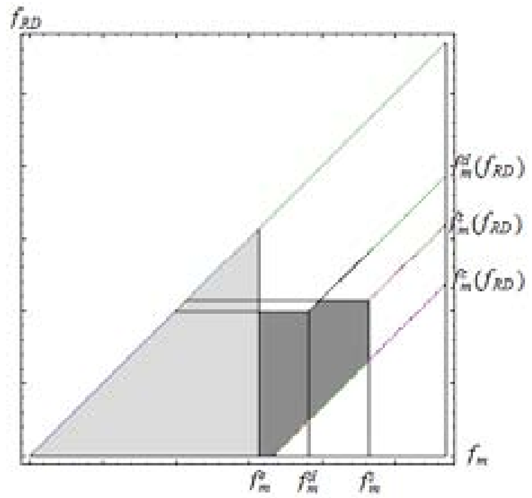
We now know that R&D cooperations may induce mergers that would otherwise not exist. Several questions, however, arise. Are these R&D networks stable? Will firms be willing to construct these cooperations knowing that a merger may follow their decision? For the different R&D and merging costs and given the outcome in the merging stage, what are the stable R&D networks?

The R&D stage In this stage, firms decide with whom to form an R&D cooperation. They will do so knowing that, for a given pair (f_m, f_{RD}), this decision will influence the merger stage and, consequently, the market competition.

As we want to highlight the relationship between R&D and M&A, we will focus on the regions where firms may decide not to merge, depending on the R&D network.

[**Claim 2..1**] If $f_m > f_m^e$, then there are, at most, 9 different regions where a merger may exist, depending on the R&D network structure.

The resulting regions are illustrated in figure 1. The nine possible areas are the ones to the right of f_m^e.



Given this claim, we analyze the R&D collaboration stability for each of these 9 regions. Each region is characterized by different stable mergers, which arise from different underlying stable R&D networks. We can have, for instance, a region where, following an empty R&D network, there will be no merger; following a dominant group network, the two cooperating firms will merge; following a star network, the center will merge with one of the extremes; and following a complete network, a merger will occur. This description corresponds to the area to the left of $f_m^d\left(f_{RD}\right)$. We want to know whether firms have incentives to maintain an R&D structure such that a merger is induced. To do so, let us start by defining what we mean by "maintaining".

The stability concept that we use is based on the pairwise stability from Jackson and Wolinsky (1996), but it has an extra requirement. We allow for the firms to sever all links at once and, hence, demand that they are better off by not doing so.

Definition 2..1 *A network g is stable if:*

1. *For $g_{i,j}=1$, $\pi_i\left(g\right)\geq\pi_i\left(g-g_{i,j}\right)$ and $\pi_j\left(g\right)\geq\pi_j\left(g-g_{i,j}\right)$;*

2. *For $g_{i,j}=0$, $\pi_i\left(g\right)\geq\pi_i\left(g+g_{i,j}\right)$ or $\pi_j\left(g\right)\geq\pi_j\left(g+g_{i,j}\right)$;*

3. For all i in $\mathcal{N}$, $\pi_i(g) \geq \pi_i(g_{-i})$.

The complete formal analysis of each of the 9 regions is quite cumbersome; hence, we present it in appendix 2 where we provide a full description of the thresholds that define each of the regions, as well as the stable R&D networks and the consequent merger decisions.

The following propositions are based on the stability concept and on the results given in appendix 2. We choose to present these propositions, as they provide us the key intuitive results.

We start by analyzing the case where the merger cost is significantly high.

Proposition 2..3 *If the merging fixed cost is great, $f_m \geq f_m^s$, then the stable R&D networks include (depending on f_{RD}) the complete, dominant and empty networks. All of the stable networks induce no merger.*

In this case, no merger will occur because it is too costly. Mergers are anticipated by firms when they construct their R&D network. In this case, the merging cost is very high, which implies that the merger is non-profitable. Therefore, firms will only build R&D networks if no merger follows. As mergers are very costly, the only possible mergers are those where firms would substitute a sufficiently costly R&D cooperation by a merger. The relevant cost would not be the absolute merger fixed cost, but rather the incremental cost of cooperation to merging. Hence, all R&D networks that would induce a merger by substituting a cooperation are not stable.

Corollary 2..1 *If the cost of merging is significantly great, no strategic interactions between merging and cooperating exist.*

R&D cooperations are completed only for the sake of lowering the marginal cost, as firms will never create networks that lead to a merger. If the R&D link cost is too great, then the benefit from cooperation is not sufficient to compensate for it, and hence, the empty network is stable. If a smaller cooperation cost is considered, then the dominant group is uniquely stable. Here, the R&D cooperation costs are sufficiently small and thus induce cooperation but are still too high for the complete network to be the most profitable. For an even lower R&D fixed cost, both the dominant and the complete networks are stable. This happens because players in a dominant network have no interest in forming an intermediate structure, which would be the star, and hence, they never reach the complete status. On the other hand, if the initial situation is a complete network, then firms would rather stay as they are, and hence, they do not create the star network. In this case, severing all links and becoming the outsider of the dominant group network also leaves firms worse off. Below a certain threshold, the complete network is uniquely stable, as cooperating represents a very small cost, and the profits, the gross of the link costs, increase as the number of cooperations increase. This replicates the results of Goyal and Joshi (2003), where merging is not a possibility. In their paper, the stable cooperation structures include the empty, complete and dominant group networks.

The following lemma states the stable R&D networks for an intermediate value of the merging cost.

Lemma 2..3 *If the merging cost is intermediate,* $f_m^e \leq f_m \leq f_m^s$*, then the stable R&D networks include (depending on* f_{RD} *and the other parameters) the complete,* *star, dominant and empty networks.*

Building on this lemma, we can state the stable R&D networks and identify their consequences on the later merger activity.

Proposition 2..4 *If the merging cost is intermediate,* $f_m^e \leq f_m \leq f_m^s$ *and*

1. the R&D cooperations cost f_{RD} *is great, then the empty network is uniquely stable and no merger is induced.*

2. the R&D cooperations cost f_{RD} *is intermediate, then the dominant, star and complete networks are stable, and all induce a merger.*

3. the R&D cooperations cost f_{RD} *is low, then the complete network is uniquely stable, and no merger is induced.*

There are mergers that occur because there is the possibility to cooperate. If cooperations were not allowed for this merging cost, firms would never want to merge. The R&D cooperations create sufficient asymmetry among firms to create the incentives for a merger. For instance, take the case where a complete R&D network induces a merger. By our exogenous rule, firms 1 and 2 will merge. None of these firms has an incentive to sever one R&D cooperation link, as the resultant star network would induce a merger between the extremes. The final equilibrium outcome would still imply that the same two firms would merge.

These firms also have no incentive to sever all of their links, as this would create a network where the other two firms would cooperate. There are two possible consequences in this case. Either the non-cooperating firm still merges, but suffers a bargaining disadvantage and hence receives lower profits, or no merger is induced. In the latter case, the non-cooperating firm would lose, as there would be more competition (three firms in the final product market) and because the firm would be competing against two much more efficient firms.

The non-merging firm also has an incentive to maintain its links. The firm benefits from the competition reduction and from the marginal cost decrease due to the R&D cooperations. This situation yields the firm a greater profit than would be generated by being part of a merger (by severing only one cooperation) or by inducing no merger (by retaining no cooperations).

Corollary 2..2 *If both the merging and cooperation costs are intermediate, mergers and R&D cooperations complement one another.*

From the previous proposition, we conclude that the possibility to cooperate in R&D leads to a larger scope for mergers. In addition to this effect, the fact that firms are able to anticipate a merger also makes the star R&D network stable. When mergers are not a possibility, this network is never stable, which means that mergers, in turn, expand the set of stable R&D networks. R&D cooperations and mergers can be thought of as complements, as the existence of one induces the other.

Note that this complementarity does not hold for all values of the fixed costs. When cooperating is sufficiently inexpensive, firms will cooperate and would prefer to maintain these cooperations rather than replace them by mergers. In this case, mergers are too costly relative to cooperations. The increase in marginal cost reduction from cooperation to merge, and the competition reduction gains are not sufficient to compensate for the increase in the fixed cost from cooperation to merger. Additionally, when R&D cooperations represent a significant cost, no merger is induced. In this case, it is too costly to form any cooperation and, consequently, only the empty R&D network is stable. For these values of the merger's fixed cost, no merger is induced.

There is non-monotonicity of the influence of the R&D decisions on mergers. If the R&D cooperation link is very inexpensive, no merger is induced because the merger is relatively too expensive, and when the R&D cost is very high, no merger is induced because the cost to cooperate and, hence, induce a merger is too great. Firms are only able to induce mergers for intermediate values of the R&D link cost. This has implications for policy makers in that if mergers are, for some reason, to be prevented, then authorities should either make cooperations very expensive or very inexpensive.

Corollary 2..3 *In all stable R&D networks that lead to a merger (except for the complete R&D network), mergers do not substitute for cooperations.*

Furthermore, it is important to stress that firms do not use cooperations as a means to commit to a later merger. The intuition that firms cooperate to achieve less costly future mergers does not apply here because if firms know that they will substitute the R&D cooperation by a later merger, they will have either no incentive to create the R&D link in the first place or no need for the link to realize the merger. This point is supported by Hagedoorn and Sadowski (1999), as the authors show that there is no evidence that firms cooperate to integrate themselves for the purpose of eventually inducing a merger. Mergers are induced through cooperations, but the fundamental cooperations are not performed among the merging firms. This point is also observed in the automobile industry. The cooperation between Renault and Nissan did not lead to a merger between them but, rather, to mergers where one of the firms participated.

2.3 The scope to induce a merger

The results from the previous section show that there is scope to induce a merger through R&D cooperations, but one should consider the extent of this scope and the degree to which the scope is dependent on the parameters. It is of interest to check the consequence of changes of some parameters. The previous propositions only concern existence, and it may be the case that we can guarantee nonexistence in some cases.

First, consider a change in the R&D cooperation marginal cost reduction.

Proposition 2..5 *If the R&D cooperation marginal cost reduction is equal to zero,* $\gamma_1 = 0$*, then no merger is induced through an R&D cooperation.*

There are three reasons that influence a merger decision, including the marginal cost of efficiency gains as considered by the firms, the possibility of substituting an R&D link by a merger and the change in market structure. The latter reason is influenced by the asymmetry of firms. In the case where the marginal cost gains from R&D cooperation are zero, no asymmetry is created in the R&D network and no marginal cost efficiency gains from cooperating exist. When firms are perfectly homogeneous, we find the result of Salant et al. (1983) is that firms do not have individual incentives to merge only for market structure motives.

The reason regarding the cost efficiencies is excluded, as we are only considering values of the merging fixed cost for which a merger with no R&D cooperations does not exist, i.e., $f_m \geq f_m^e$. The cooperation substitution reason should be taken into account in the first stage when R&D cooperations are created. As stated before, for values of the fixed costs for which firms want to substitute a cooperation by a merger, but where firms ex-ante do not wish to merge, firms would not create the substitutable links in the first place. Hence, gains from cooperation, as they create asymmetry in the market, are necessary for there to be R&D-induced mergers.

Let us now consider the consequences of a change in marginal cost reduction caused by a merger, γ_2. When no stable R&D network induces a merger, i.e., due to a very large merger cost, a change in γ_2 has no influence in the final competition stage as mergers are simply not considered.

For the cases where a strategic interaction between mergers and cooperations exists, the value of the merger efficiency gain matters. In fact, all induced mergers need a collaboration link with some firm that is not going to be a part of the merger. The incentives to form these necessary links depend on γ_2, and hence, a change in the marginal cost reduction due to a merger has an impact on the results.

Proposition 2..6 *If the marginal cost reduction due to a merger is large relative to the cooperation cost reduction, $\gamma_2 \geq \frac{1}{8}(\alpha - \gamma_0) + \gamma_1$, then no merger is induced through an R&D cooperation.*

When the marginal cost reduction due to a merger is relatively large, then either the empty or the dominant group network is stable. The important characteristic is that neither of them induce a merger. As previously stated, for a merger to be induced, it is necessary for the non-merging firm to cooperate with one or two merging firms. In some sense, the firm that stays outside of the merger has veto power with respect to the merger. By not cooperating with anyone, this firm can avoid all mergers. If the marginal cost reduction due to a merger, γ_2, is large, then the firm resulting from the merger is highly competitive, which harms the non-merging firm. The non-merging firm faces a trade-off between the benefits of cooperating and having less market competition on the one hand, while, on the other hand, facing the efficiency of the resulting competitor. If γ_2 is large enough, the competitor is very efficient, and hence, the outsider does not want to induce the merger. As a result, only the cooperation structures that induce no merger will be stable.

From the previous proposition, we learn that mergers require the consent of the non-merging firm, but this is not enough. Mergers must also have the consent of the merging firms. When merging becomes too costly, the merging firms are less eager to induce a merger.

Corollary 2..4 *The lower bound of the R&D cooperation costs, f_{RD}, below which firms cooperate but do not merge, increases with the merging costs, f_m.*

From this corollary, we see that the larger the merging cost, the easier it is to sustain cooperation among firms without inducing a merger. When the merging cost is significantly large, we already know, by proposition 8, that there will be no merger. However, when restricting the analysis to the regions where a merger may be induced, a greater merging cost and a low cooperation cost causes firms to be less prone to merge simply because cooperating becomes relatively inexpensive. In this case, the complete network would be created, and no merger would be induced.

2.4 Welfare Implications

The important question regarding R&D activity is whether the activity should be facilitated. In our model, R&D cooperations decrease marginal costs and are, therefore, welfare enhancing. On the other hand, however, these cooperations may change the market structure and decrease competition, which could harm welfare. The final effect that a merger has on welfare is also unclear, as it represents an increase in technological efficiency and results in smaller market competition. Thus gives rise to the question, "Which is the best social outcome?"

In the previous case, if the size of the fixed costs were negligible, then the socially efficient outcome would be one that induces a merger. There are various effects with different directions. The consumer surplus is reduced by the decrease in market competition, but this negative effect is countervailed by the increase in technological efficiency of the companies. From the perspective of the company, a merger increases the profits due to the decrease in market competition and to the lower marginal costs. As the total effect is positive, a cooperation structure that leads to a merger is the socially efficient network.

It is important, however, to take the cooperations and merging costs into account, as they are major drivers of our results. Therefore, we will consider different relative sizes of the cooperations and merging costs and verify whether the stable equilibria correspond to the social optimum. Prior to a discussion of the details, we note that when the merging cost is significantly large and the stable R&D networks induce no merger, then the social planner does not need to consider the merger possibility and will subsidize cooperations only if the distortions created by the necessary taxing are smaller than the gains from the improved technology. For these regions, because the analysis is straightforward, we will concentrate on the areas where a merger may be induced.

Proposition 2..7 *If γ_2 is small such that $f_{RD}^{-1}(f_m^c) < f_{RD}^W(f_m)$ and the cooperation cost is such that $f_{RD}^{-1}(f_m^c) \leq f_{RD} \leq f_{RD}^W(f_m)$, then in the private equilibrium, the star and complete networks are stable and induce a merger but the social optimum would be a complete cooperation network with no merger.*
If γ_2 is large such that $f_{RD}^{-1}(f_m^c) > f_{RD}^W(f_m)$ and the cooperation cost is such that $f_{RD}^W(f_m) \leq f_{RD} \leq f_{RD}^{-1}(f_m^c)$, then in the private equilibrium, the complete network is stable and induces no merger, and the social optimum would be either that the complete or the star cooperation networks lead to a merger.

As stated in the proposition, there is a case where firms induce a socially undesirable merger. The increase in profits does not outweigh the increase in fixed costs and the reduction in consumer surplus, as cooperations are substituted by a relatively expensive merger. The social first-best would be to allow the cooperations because doing so would benefit the complete network, thereby resulting in maximum cooperation among firms and avoiding mergers. It appears that, in some instances, firms decide to substitute their cooperations by mergers too soon because the firms fail to consider the impact of competition reduction on the total welfare. Cooperations create an externality by increasing the merger opportunity.

The second case is when a merger is not induced, although it would be socially efficient to do so. In this case, the technological efficiency gains due to a merger are significant. Thus, the impact of a merger on social welfare is positive, but in this case, firms substitute these cooperations by mergers too late.

Corollary 2..5 *There exists a subsidy (tax) on R&D cooperations such that* $f_{RD}^{-1}(f_m^c) + s = f_{RD}^W(f_m)$. *In this case, the private optimum coincides with the social optimum.*

Given a merging cost, consider a situation where the equilibrium outcome does not coincide with the maximum welfare outcome, and firms merge too soon. How can the cooperation costs be influenced to improve welfare?

If there is a sufficiently large increase in the cooperation costs, firms will decide not to cooperate and, hence, do not merge. The socially undesirable merger, as well as the socially desirable cooperations, is avoided. The other possible choice is for the cooperation cost to be decreased by a sufficient amount through a subsidy. This reduction would cause cooperations to be so inexpensive that firms would rather not substitute them with mergers. In this way, the social optimum would coincide with the private equilibrium. Thus, a movement towards the social first best can be achieved by a reduction in the cooperation costs, and again, the only question that remains is whether the tax distortions are greater or less than the increase in welfare. The second case is when firms merge too late, in which case R&D cooperations should be taxed so firms can substitute them for mergers.

2.5 Concluding Remarks

In our model, we have analyzed the strategic interaction between R&D cooperations among firms and their subsequent M&A decisions, and we find that R&D cooperations may enlarge the scope for mergers, as firms are able to influence their technological efficiency. However, it is important to note that the existence of the R&D cooperations does not reduce the scope of mergers, i.e., if firms find it ex-ante profitable to merge, no cooperation can prevent it.

We also find that when cooperation costs are sufficiently low relative to the merging costs, then firms do not substitute cooperations by mergers, whereas when the cooperation costs are significantly large relative to the merging costs, then no cooperation is created, and hence, no merger occurs. This represents a non-monotonic relationship between R&D cooperations and the induction of the merger.

Regarding the non-merging firm, we find that it maintains veto power over the possible merger, as a collaboration with the non-merging firm is necessary for the merger to be induced. The implication of this result is that induced mergers will only be possible if the resulting firm is not too efficient as this would harm the non-merging firm. Finally, we note that there is scope to induce a merger where otherwise none would exist only if R&D cooperations create asymmetry in the final product market. If no cost reductions would exist when the firms cooperate, then the mergers would only be induced if they would substitute for R&D links. However, in this case, these R&D links would not be stable in the first place.

With respect to welfare, we find that merging firms substitute their cooperation by a merger too soon (late) when the technological efficiency gain is small (large). If R&D cooperations were subsidized, then firms would maintain their cooperations for a larger pre-subsidy R&D cost. With a subsidy of the right size, it would be possible to cause the private equilibrium to coincide with the social optimum.

2.6 Appendix

In Appendix 1 we present the formal proofs of the propositions and the lemmas presented throughout the paper. These are the most important results. In Appendix 2 we present the claims, and their respective proofs, that formally define the 9 regions under analysis.

Appendix 1 **Proof Proposition 1.** The threshold f_m^e is always smaller than f_m^d, f_m^s and $f_m^c(f_{RD})$ that are defined in the second appendix. Hence all R&D networks below f_m^e induce a merger. ■

Proof Lemma 1. The thresholds for the merger for non-cooperating firms are: f_m^e, f_m^d and f_m^s.

$f_m^s - f_m^d = \frac{1}{9}\gamma_1\left(\alpha - \gamma_0 + \frac{21}{8}\gamma_1 + 4\gamma_2\right) > 0$

$f_m^d - f_m^e = \frac{1}{9}\gamma_1\left(\alpha - \gamma_0 - \frac{5}{8}\gamma_1 + 4\gamma_2\right) > 0.$ ■

Proof Lemma 2. The thresholds for the merge of cooperating firms are: $f_m^d(f_{RD})$, $f_m^s(f_{RD})$ and $f_m^c(f_{RD})$.

$f_m^c(f_{RD}) - f_m^s(f_{RD}) = \frac{1}{9}\gamma_1\left(\alpha - \gamma_0 + \frac{15}{4}\gamma_1 + 4\gamma_2\right) > 0$

$f_m^s(f_{RD}) - f_m^d(f_{RD}) = \frac{1}{9}\gamma_1\left(\alpha - \gamma_0 - \frac{7}{4}\gamma_1 + 4\gamma_2\right) > 0$ ■

Proof Proposition 2. Follows from the previous lemmas. ■

Proof Claim 1. To obtain these 9 regions it is enough to intersect all conditions of appendix 2, and only consider the areas for which the merge cost, f_m, is larger than f_m^e. ■

Proof Proposition 3. For a merging cost above f_m^s firms will only merge if they substitute a sufficiently costly R&D cooperation. Hence, if firms find merging too costly they will ex-ante not create the R&D structures that would lead to a merger.

The stable R&D structures are created for the sake of lowering marginal cost through cooperation. Which are the stable ones depends on the R&D cooperation cost. A large R&D cooperation cost implies the empty network, whereas a low R&D cooperation cost implies the complete network. For an intermediate R&D cooperation cost both the dominant and the complete networks are stable. ■

Proof Corollary 1. Follows directly from the previous proposition. ■

Proof Lemma 3. Follows from the claims of appendix 2. ■

Proof Proposition 4. If merging costs intermediate then and there exist stable networks that induce a merger: great R&D empty no merger; intermediate R&D dominant, star and complete induce merger; low R&D complete induces no merger) ■

Proof Corollary 2. Follows directly from the previous proposition. ■

Proof Corollary 3. Follows directly from the previous proposition. ■

Proof Proposition 5. If $\gamma_1 = 0$ then $f_m^e = f_m^d = f_m^s$. Therefore the area where cooperations and mergers are complements does not exist. ■

Proof Proposition 6. The last binding restriction for stability is of the firm that is at the center of the star network. The star network induces a merger between the extremes. The restriction is in order for this firm not to severe all the links and induce the empty network, which induces no merge. The threshold value for the cooperation cost is $f_{RD} \leq \frac{1}{18}(\alpha - \gamma_0 + 2\gamma_1 - 2\gamma_2)^2 - \frac{1}{32}(\alpha - \gamma_0)^2$, and this threshold is non-positive for $\gamma_2 \geq \frac{1}{8}(\alpha - \gamma_0) + \gamma_1$. For these values of γ_2 not even the most robust merger inducing network is stable, and hence no merger occurs. ■

Proof Corollary 4. The lower bound is defined by the threshold $f_m^c(f_{RD})$, that is used in the proof of lemma 3. This threshold is obviously increasing in f_{RD}. ■

Proof Proposition 7. Define the threshold that makes the society indifferent between a merge and a complete network with no merge $f_{RD}^W = \frac{15}{64}(\alpha - \gamma_0 + 2\gamma_1)^2 - \frac{1}{18}\left[(\alpha - \gamma_0 + 2\gamma_1 + 4\gamma_2)^2 + (\alpha - \gamma_0 + 2\gamma_1 - 2\gamma_2)^2 + 2(\alpha - \gamma_0 + 2\gamma_1 + \gamma_2)^2\right] + f_m$.

As seen before, for $f_{RD} \geq f_{RD}^{-1}(f_m^c)$ firms choose to merge. If $f_{RD}^W > f_{RD}^{-1}(f_m^c)$ then there is an interval in between the two thresholds where it is not socially desirable to merge, and where firms do so. It also implies that for $f_{RD} \leq f_{RD}^{-1}(f_m^c)$ firms only cooperate, which coincides with the social first best.

If $f_{RD}^W < f_{RD}^{-1}(f_m^c)$ then there is an interval in between the two thresholds where it is socially desirable to merge, and where firms do not do so. For $f_{RD} \leq f_{RD}^{-1}(f_m^c)$ firms only cooperate, which does not coincide with the social first best.

$f_{RD}^W - f_{RD}^{-1}(f_m^c) = \frac{1}{576}\left(3(\alpha - \gamma_0 + 2\gamma_1)^2 - 64\gamma_2^2\right) > 0$, for γ_2 small enough.

$f_{RD}^W - f_{RD}^{-1}(f_m^c) = \frac{1}{576}\left(3(\alpha - \gamma_0 + 2\gamma_1)^2 - 64\gamma_2^2\right) < 0$, for γ_2 large enough. ■

Proof Corollary 5. Follows from the previous proposition. The subsidy (tax) should be such that the post subsidized (taxed) cooperation cost is equal to f_{RD}^W. ■

Appendix 2 Let us first define the merger thresholds before analyzing the R&D network stability.

The merging cost that makes two non cooperators indifferent between merging or not is (i) under the empty network $f_m^e = \frac{1}{18}(\alpha - \gamma_0 + 4\gamma_2)^2 - \frac{1}{16}(\alpha - \gamma_0)^2$, (ii) under the dominant group network $f_m^d = \frac{1}{18}(\alpha - \gamma_0 + \gamma_1 + 4\gamma_2)^2 - \frac{1}{32}\left((\alpha - \gamma_0 - 2\gamma_1)^2 + (\alpha - \gamma_0 + 2\gamma_1)^2\right)$ and (iii) under the star network $f_m^s = \frac{1}{18}(\alpha - \gamma_0 + 2\gamma_1 + 4\gamma_2)^2 - \frac{1}{16}(\alpha - \gamma_0)^2$.

The merging cost that makes two cooperators indifferent between merging or not is

(i) under the dominant group network $f_m^d(f_{RD}) = \frac{1}{18}(\alpha - \gamma_0 + 4\gamma_2)^2$

$-\frac{1}{16}(\alpha - \gamma_0 + 2\gamma_1)^2 + f_{RD}$

(ii) under the star network $f_m^s(f_{RD}) = \frac{1}{18}(\alpha - \gamma_0 + \gamma_1 + 4\gamma_2)^2$

$-\frac{1}{32}\left((\alpha - \gamma_0 + 4\gamma_1)^2 + (\alpha - \gamma_0)^2\right) + f_{RD}$

(iii) under the complete network $f_m^c(f_{RD}) = \frac{1}{18}(\alpha - \gamma_0 + 2\gamma_1 + 4\gamma_2)^2$

$-\frac{1}{16}(\alpha - \gamma_0 + 2\gamma_1)^2 + f_{RD}$.

The R&D fixed cost that makes any two firms indifferent between merging with a cooperator or with a non cooperator is (i) under the dominant group network $f_{RD}^d = \frac{13}{36}\gamma_1\left(\alpha - \gamma_0 + \frac{2}{13}\gamma_1 + \frac{16}{13}\gamma_2\right)$ and (ii) under the star network $f_{RD}^s = \frac{1}{2}\gamma_1(\alpha - \gamma_0 + 2\gamma_1)$.

The following claims analyze the stable R&D networks for each of the possible areas defined by the intersection of the previously defined thresholds.

If (f_m, f_{RD}) are such that $f_m^e \leq f_m < f_m^d$ and $f_{RD}^{-1}(f_m^c) \leq f_{RD} < f_{RD}^d$, then:

1. the empty network is stable if $f_{RD} \geq f_{RD}^1(f_m)$ or $f_{RD} \geq f_{RD}^2$;

2. the dominant network is stable if $f_{RD} \leq f_{RD}^1(f_m)$ and $f_{RD} \leq f_{RD}^2$ and either $f_{RD} \geq f_{RD}^3$ or $f_{RD} \geq f_{RD}^4$ and either $f_{RD} \geq f_{RD}^5(f_m)$ or $f_{RD} \geq f_{RD}^4$;

3. the star with center 1 is stable if $f_{RD} \leq f_{RD}^5(f_m)$ and $f_{RD} \leq f_{RD}^6$ and $f_{RD} \leq f_{RD}^4$ and $f_{RD} \geq f_{RD}^7$;

4. the star with center 2 is stable if $f_{RD} \leq f_{RD}^5(f_m)$ and $f_{RD} \leq f_{RD}^6$ and $f_{RD} \leq f_{RD}^4$ and $f_{RD} \geq f_{RD}^7$ and also $f_{RD} \leq f_{RD}^8$;

5. the star with center 3 is stable if and $f_{RD} \leq f_{RD}^6$ and $f_{RD} \leq f_{RD}^4$ and also $f_{RD} \leq f_{RD}^8$;

6. the complete network is stable if $f_{RD} \leq f_{RD}^7$ and $f_{RD} \leq f_{RD}^9$ and also $f_{RD} \leq$

$f^{10}_{RD}(f_m)$.

Proof. 1. The empty network is stable if either it does not payoff to form a link and merge, above $f^1_{RD}(f_m)$, or to form a link and stay out of the merge, above f^2_{RD}. $f^1_{RD}(f_m) = \frac{1}{18}(\alpha - \gamma_0 + \gamma_1 + 4\gamma_2)^2 - \frac{1}{32}\left[\begin{array}{c} (\alpha - \gamma_0 - 2\gamma_1)^2 + \\ 2(\alpha - \gamma_0)^2 - (\alpha - \gamma_0 + 2\gamma_1)^2 \end{array}\right] - f_m$ and $f^2_{RD} = \frac{1}{9}(\alpha - \gamma_0 + \gamma_1 - 2\gamma_2)^2 - \frac{1}{16}(\alpha - \gamma_0)^2$.

2. The dominant network is stable if the link is not severed, these are the two conditions derived trivially from the previous point, and if no other link is created. The non creation of another link is guaranteed if either it is better to be non merging in a dominant than non merging in a star, $f^3_{RD} = \frac{1}{9}\left[(\alpha - \gamma_0 + 2\gamma_1 - 2\gamma_2)^2 - (\alpha - \gamma_0 + \gamma_1 - 2\gamma_2)^2\right]$ or it is better to be merging as an outsider in a dominant than in a star $f^4_{RD} = \frac{1}{18}(\alpha - \gamma_0 + \gamma_1 + 4\gamma_2)^2 - \frac{1}{32}\left[(\alpha - \gamma_0 - 2\gamma_1)^2 + 2(\alpha - \gamma_0)^2 - (\alpha - \gamma_0 + 2\gamma_1)^2\right]$ and also if either it is better to be in a merging in a dominant than non merging in a star $f^5_{RD}(f_m) = \frac{1}{18}\left[\begin{array}{c} 2(\alpha - \gamma_0 + 2\gamma_1 - 2\gamma_2)^2 \\ -(\alpha - \gamma_0 + \gamma_1 + 4\gamma_2)^2 \end{array}\right] - \frac{1}{32}\left[\begin{array}{c} (\alpha - \gamma_0 + 2\gamma_1)^2 \\ -(\alpha - \gamma_0 - 2\gamma_1)^2 \end{array}\right] + f_m$ or it is better to be merging as an outsider in the dominant than be merging in a star and we have again f^4_{RD}.

3. The star with center 1 is stable if the links are not severed. The first condition ensures that the center of the star prefers to remain there, and not be transform it into a dominant one where (as we are talking of company 1, and due to our exogenous rule) it would merge with the outsider. The second condition ensures that the center of the star does not want to severe both links and form an empty network $f^6_{RD} = \frac{1}{18}(\alpha - \gamma_0 + 2\gamma_1 - 2\gamma_2)^2 - \frac{1}{32}(\alpha - \gamma_0)^2$. The third condition ensures that the extremes do not want to severe the link and become an outsider of the dominant R&D. The fourth

condition ensures the extremes don't want to create the complete network, because the it is better to be in a merge, than stay out of it in a complete network $f_{RD}^{7}(f_m) = \frac{1}{18}\left[2(\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2-(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2\right]+f_m$.

4. Obviously all the restrictions are the same, one just has to take into account that the firm may now be in a dominant group and not merge, when the others do so. This possibility did not exist for firm 1 due to the exogenous rule. That the firm does not want to do this is guaranteed by the last restriction where $f_{RD}^{8}=\frac{1}{9}\left[(\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2-(\alpha-\gamma_0+\gamma_1-2\gamma_2)^2\right]$.

5. For star with center 3 the condition for firm 3 not to severe a link is that it is better off being the center of star than being in a dominant R&D and not merging, i.e. $f_{RD}\leq f_{RD}^{8}$. Not severing all the links is also ensured by $f_{RD}\leq f_{RD}^{6}$ and the last requirement, $f_{RD}\leq f_{RD}^{4}$ is that the extremes do not want to severe a link themselves, as all firms are indifferent between this network or the complete one.

6. For the complete network we have to ensure the first condition, where firms are firms are better off by being the outsider of the merge in the complete than by becoming an extreme of the star, i.e. $f_{RD}\leq f_{RD}^{7}$. Merging firms should also be better off in this situation than by severing both links and becoming an outsider of a dominant group. This defines threshold $f_{RD}^{9}=\frac{1}{18}\begin{bmatrix}(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2\\-(\alpha-\gamma_0+\gamma_1+4\gamma_2)^2\end{bmatrix}-\frac{1}{32}\begin{bmatrix}(\alpha-\gamma_0-2\gamma_1)^2\\-(\alpha-\gamma_0+2\gamma_1)^2\end{bmatrix}$. At last we must have that the same is true for the non merging firm, and hence $f_{RD}^{10}(f_m)=\frac{1}{18}\left(2(\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2-(\alpha-\gamma_0+\gamma_1+4\gamma_2)^2\right)$ $-\frac{1}{32}\left((\alpha-\gamma_0-2\gamma_1)^2-(\alpha-\gamma_0+2\gamma_1)^2\right)+f_m$. ■

For the next claim we consider the same problem, but consider a larger cooperation cost. We will be in the area where a dominant group now induces a merger between the two cooperating firms, and everything else is as before. This is so because cooperation is now too costly and both firms in the dominant group now prefer to substitute their link by a merger.

If (f_m, f_{RD}) are such that $f_m^e \leq f_m < f_m^d(f_{RD})$ and $f_{RD}^d \leq f_{RD} < f_{RD}^s$, then the empty network is the uniquely stable network.

Proof. The dominant can not be stable because if firms do not want to merge with no R&D, and the R&D is fully substituted by a merge, firms would not want to cooperate, as the final outcome would be a merge.

The star network would be stable for $f_{RD} \leq \frac{1}{18}\begin{bmatrix} (\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2 \\ -\frac{1}{2}(\alpha-\gamma_0+4\gamma_2)^2 \end{bmatrix} - \frac{1}{2}f_m$ and $f_{RD} \leq \frac{1}{18}\begin{bmatrix} (\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2 \\ -2(\alpha-\gamma_0-2\gamma_2)^2 \end{bmatrix} - f_m$, which ensures that no firm wants to sever a link, and also for $f_{RD} \geq \frac{1}{18}\begin{bmatrix} 2(\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2 \\ -(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2 \end{bmatrix} + f_m$, which ensures that the complete network is not created. The intersections of these condition with the conditions defining the considered area is empty.

The complete network would be stable for $f_{RD} \leq \frac{1}{18}\begin{bmatrix} 2(\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2 \\ -(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2 \end{bmatrix} + f_m$ and $f_{RD} \geq \frac{1}{18}\left[(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2 - 2(\alpha-\gamma_0-2\gamma_2)^2\right] - f_m$ and $f_{RD} \leq \frac{1}{18}\left[(\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2 - (\alpha-\gamma_0-2\gamma_2)^2\right]$, where all the three conditions guarantee that no link is severed. Again the intersection of the conditions with the conditions defining the considered area is empty.

As all are non stable, and at least one stable network has to exist, then it has to be the empty one. ■

We can still further consider a larger cooperation cost, and this would lead us to an area where both the dominant and the star network would induce a merger such that some cooperation link would be substituted.

If (f_m, f_{RD}) are such that $f_m^e \leq f_m < f_m^d(f_{RD})$ and $f_{RD}^s \leq f_{RD} < f_m$, then the empty network is the uniquely stable network.

Proof. The proof for the dominant group network is analogous to the previous one.

The star network is also never stable because, for the considered region, the merging extreme of a star would always want to severe a link, and become an outsider of a dominant group. The threshold value for this not to be true is $f_m \leq \frac{1}{18}\left[(\alpha-\gamma_0+\gamma_1+4\gamma_2)^2 - 2(\alpha-\gamma_0-2\gamma_2)^2\right] - \frac{1}{32}\left[(\alpha-\gamma_0+4\gamma_1)^2 - (\alpha-\gamma_0)^2\right]$, which does not intersect with the considered region.

The complete network is never stable because the non merging firm would always prefer to severe a link, and remain as non merging, but with one less link. The region for this not to be true is $f_{RD} \leq \frac{1}{9}\left[(\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2 - (\alpha-\gamma_0+\gamma_1-2\gamma_2)^2\right]$, which does not intersect with the considered region.

As all described networks are non stable, and at least one stable network has to exist, then it has to be the empty one. ■

For the next claim we consider a larger merger cost, such that the dominant group would rather keep the cooperation link.. Everything else is the same as in the previous claim except for this new feature.

If (f_m, f_{RD}) are such that $f_m^d(f_{RD}) \leq f_m < f_m^s(f_{RD})$ and $f_{RD}^s \leq f_{RD}$ then the empty network is the uniquely stable network.

Proof. The dominant group is not stable because in this region it is too costly too cooperate. They would only do so if $f_{RD} \leq \frac{1}{16}\left[(\alpha-\gamma_0+2\gamma_1)^2-(\alpha-\gamma_0)^2\right]$, which does not intersect the considered region.

The star network is not stable because the merging extreme is better off by severing their link, as it is too costly to merge. The considered area is only defined for $f_m \geq \frac{1}{18}\left[(\alpha-\gamma_0+4\gamma_2)^2+(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2-(\alpha-\gamma_0+\gamma_1+4\gamma_2)^2\right]$ $-\frac{1}{32}\left[(\alpha-\gamma_0)^2-(\alpha-\gamma_0+4\gamma_1)^2+2(\alpha-\gamma_0+2\gamma_1)^2\right]$, whereas the condition for the merging extreme not to severe the link (and become an outsider of a non merging dominant group) is $f_m \leq \frac{1}{18}(\alpha-\gamma_0+\gamma_1+4\gamma_2)^2-\frac{1}{32}\begin{bmatrix}(\alpha-\gamma_0+4\gamma_1)^2+\\ 2(\alpha-\gamma_0-2\gamma_1)^2-(\alpha-\gamma_0)^2\end{bmatrix}$. The difference between these two thresholds is $\frac{10}{9}\gamma_1^2$, which implies that they never intersect, and hence that the star network is never stable.

The complete network is also not stable. The reasoning is the same as in the previous proposition.

Again, as all described networks are non stable, and at least one stable network has to exist, then it has to be the empty one. ∎

The next case to be considered is where also the star network induces no merge, and everything else is as before.

If (f_m, f_{RD}) are such that $f_m^s(f_{RD}) \leq f_m < f_m^c(f_{RD})$ and $f_m^s \leq f_m$ then the empty network is the uniquely stable network.

Proof. The considered area is only defined for $f_{RD} \geq \frac{1}{16}\left[\begin{array}{c}(\alpha-\gamma_0+2\gamma_1)^2\\ -(\alpha-\gamma_0)^2\end{array}\right]$, and for $f_m \geq \frac{1}{18}(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2 - \frac{1}{16}(\alpha-\gamma_0)^2$.

The dominant group network is never stable. The condition for the link not to be destroyed is $f_{RD} \leq \frac{1}{16}\left[(\alpha-\gamma_0+2\gamma_1)^2-(\alpha-\gamma_0)^2\right]$ which does not intersect the considered area.

The star network is not stable. The condition that the extreme of the star does not want to severe his link is $f_{RD} \leq \frac{1}{16}\left[(\alpha-\gamma_0-2\gamma_1)^2-(\alpha-\gamma_0)^2\right]$, which again does not intersect the considered area.

The complete is not stable. The condition for a merging firm not to severe a link, and thus become an extreme of a star is $f_m \leq \frac{1}{18}(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2-\frac{1}{16}(\alpha-\gamma_0)^2$, which clearly does not intersect the considered area.

As all described networks are not stable, and at least one stable network has to exist, then it has to be the empty one. ■

In the following area we consider the case where no merge never happens. Independently of the existing R&D network, the merge cost is so large that no merge takes place. The only reason for firms too cooperate, is the reduction in the marginal cost.

If (f_m, f_{RD}) are such that $f_m^c(f_{RD}) \leq f_m$ and $f_m^s \leq f_m$ then:

1. the empty network stable if $f_{RD} \geq f_{RD}^{11}$;
2. the dominant group is stable if $f_{RD}^{12} \leq f_{RD} \leq f_{RD}^{11}$;
3. the star network is never stable;
4. the complete network is stable if $f_{RD} \leq f_{RD}^{13}$.

Proof. 1. The empty network is stable if it is too costly to cooperate: $f_{RD}^{11} = \frac{1}{16}\left[(\alpha-\gamma_0+2\gamma_1)^2-(\alpha-\gamma_0)^2\right]$.

2. The dominant group can only be stable for values above below f_{RD}^{11} and above

$f_{RD}^{12} = \frac{1}{16}\left[(\alpha-\gamma_0)^2-(\alpha-\gamma_0-2\gamma_1)^2\right]$ because otherwise both the outsider of the R&D and one of the insiders would be willing to form the star network.

3. The star network is never stable because in this area one of two things may happen. It is better for the extremes of the star to create one more link, and form the complete network, or it is better for one of the extremes to delete the link and become isolated in a dominant group network.

4. The complete network is stable below $f_{RD}^{13} = \frac{1}{32}\begin{bmatrix}(\alpha-\gamma_0+2\gamma_1)^2- \\ (\alpha-\gamma_0-2\gamma_1)^2\end{bmatrix}$ because otherwise the firms would have incentive to severe all the links at once, and become an outsider of a dominant group network. ■

Note that in the previous claim $f_{RD}^{12} \leq f_{RD}^{13}$, which implies that the complete and the dominant networks are both stable at the same time. More details on the intuition can be found in Goyal and Joshi (2003).

The next claim concerns the area where the only R&D structure that induces a merger is the star network.

If (f_m, f_{RD}) are such that $f_m^d \leq f_m < f_m^s$ and $f_{RD} \leq f_{RD}^{-1}(f_m^c)$ then

1. the empty network is never stable;
2. the dominant group R&D network is stable if $f_{RD} \geq f_{RD}^{14}$ or $f_{RD} \geq f_{RD}^{15}$;
3. the star network is never stable;

4. the complete R&D network is stable if $f_{RD} \leq f_{RD}^{13}$.

Proof. 1. The empty network is never stable because this area is restricted to $f_{RD} \leq \frac{1}{16}\left[(\alpha-\gamma_0+2\gamma_1)^2-(\alpha-\gamma_0)^2\right]$ and for these values of the cooperation costs it always pays off to at least form a dominant group cooperation.

2. The firms in the dominant group network never want to severe the link, as the considered region is always below the threshold $f_{RD} = \frac{1}{16}\left[(\alpha-\gamma_0+2\gamma_1)^2-(\alpha-\gamma_0)^2\right]$, above which a deletion would occur. The conditions for the stability are, hence, only concerning the non creation of the star network. $f_{RD}^{14}(f_m) = \frac{1}{18}(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2 - \frac{1}{16}(\alpha-\gamma_0-2\gamma_1)^2 - f_m$ ensures that the outsider of the dominant group R&D does not want to become an extreme of the star, and $f_{RD}^{15} = \frac{1}{18}(\alpha-\gamma_0+2\gamma_1-2\gamma_2)^2 - \frac{1}{32}(\alpha-\gamma_0+2\gamma_1)^2$ ensures that a firm in the dominant group does not want to become the center of a star.

3. The star network is never stable because for this area the extremes of the star will always create a link, and form the complete network. The area is defined only for $f_{RD} \leq \frac{1}{16}(\alpha-\gamma_0+2\gamma_1)^2 - \frac{1}{18}(\alpha-\gamma_0+2\gamma_1+4\gamma_2)^2 + f_m$, and the condition for the complete network not to be formed is exactly the complementary.

4. For the stability of the complete R&D network the binding restriction is to ensure that no firm wants to delete all the links, and become an outsider of the dominant group R&D network, and hence is the same as in the previous proposition.

■

The next claim concerns the region where both the dominant group and the star network induce a merge, but both the empty network and the complete network do not induce a merge.

If (f_m, f_{RD}) are such that $f_m^c(f_{RD}) \leq f_m < f_m^d$ then

1. the empty network is stable if $f_{RD} \geq f_{RD}^{16}$ or $f_{RD} \geq f_{RD}^{2}$;
2. the dominant network is never stable;
3. the star network is never stable;
4. the complete network is always stable.

Proof. 1. The empty network is stable if it is better not to induce the merge, from the point of view of the one who will merge, or from the point of view of the one who stays out of the merge. The first condition defines $f_{RD}^{16} = \frac{1}{18}(\alpha - \gamma_0 + 4\gamma_2)^2 - \frac{1}{32}\left[2(\alpha - \gamma_0)^2 + (\alpha - \gamma_0 - 2\gamma_1)^2 - (\alpha - \gamma_0 + 2\gamma_1)^2\right]$ and the second condition defines $f_{RD}^2 = \frac{1}{9}(\alpha - \gamma_0 + \gamma_1 - 2\gamma_2)^2 - \frac{1}{16}(\alpha - \gamma_0)^2$.

2. The dominant network is never stable because it is restricted by the condition that ensures that the merging insider of the R&D network does not want to become the center of a star, and this condition does not intersect the considered area. The condition is $f_{RD} \geq f_{RD}^5(f_m)$ and the area is defined by $f_{RD} \leq \frac{1}{16}(\alpha - \gamma_0 + 2\gamma_1)^2 - \frac{1}{18}(\alpha - \gamma_0 + 2\gamma_1 + 4\gamma_2)^2 + f_m$.

3. The star network is not stable for the same reason as the previous proposition.

4. The complete network is always stable because the two conditions that have to be met (not to severe one link and not to severe all links) are always both satisfied in the considered are. The first condition is coincides with the definition of the area, namely $f_{RD} \leq \frac{1}{16}(\alpha - \gamma_0 + 2\gamma_1)^2 - \frac{1}{18}(\alpha - \gamma_0 + 2\gamma_1 + 4\gamma_2)^2 + f_m$ and the second one is $f_{RD} \leq \frac{1}{64}\left[6(\alpha - \gamma_0 + 2\gamma_1)^2 - (\alpha - \gamma_0 - 2\gamma_1)^2\right] - \frac{1}{36}(\alpha - \gamma_0 + \gamma_1 + 4\gamma_2)^2 + \frac{1}{2}f_m$, which includes the considered area. ■

The last area that we have to consider is the are where the star and complete

R&D networks induce a merge, but the empty and dominant group network do not.

If (f_m, f_{RD}) are such that $f_m^d(f_{RD}) \leq f_m$ and also $f_m^d \leq f_m < f_m^s$ and $f_{RD}^{-1}(f_m^c) \leq f_{RD} < f_{RD}^s$ then

1. the empty network is stable if $f_{RD} \geq f_{RD}^{17}$;
2. the dominant group network is stable if $f_{RD}^{17} \geq f_{RD} \geq f_{RD}^{14}$ or $f_{RD}^{17} \geq \; f_{RD} \geq f_{RD}^{15}$;
3. the star network is stable if $f_{RD} \leq f_{RD}^{6}$ and $f_{RD}^{14}(f_m) \leq f_{RD}$;
4. the complete network is stable if $f_{RD}^{14}(f_m) \leq f_{RD} \leq f_{RD}^{7}(f_m)$.

Proof. 1. The empty network is stable if it does not payoff to cooperate, and this condition defines $f_{RD}^{17} = \frac{1}{16}\left[(\alpha - \gamma_0 + 2\gamma_1)^2 - (\alpha - \gamma_0)^2\right]$.

2. The dominant network is stable when no link is severed, and when no link is created. The no severing condition is $f_{RD}^{17} \geq \; f_{RD}$. The conditions for a link not to be created are the same as in the area where only the star network induces a merger between the two extremes.

3. The star network is stable if it does not payoff, for the center of the star, to severe both links at once, and for the extremes of the star to severe their link. The first condition is $f_{RD}^{6} = \frac{1}{18}(\alpha - \gamma_0 + 2\gamma_1 - 2\gamma_2)^2 - \frac{1}{32}(\alpha - \gamma_0)^2$ and the second is $f_{RD}^{14}(f_m) = \frac{1}{18}(\alpha - \gamma_0 + 2\gamma_1 + 4\gamma_2)^2 -$
$\frac{1}{16}(\alpha - \gamma_0 - 2\gamma_1)^2 - f_m$.

4. The complete network is stable if the merging firms do not want to severe both links, which corresponds to$f_{RD}^{14}(f_m) \leq f_{RD}$, and if the non-merging firm does not want to severe one of his links, which is ensured by $f_{RD}^{7}(f_m) = \frac{1}{18}\begin{bmatrix} 2(\alpha - \gamma_0 + 2\gamma_1 - 2\gamma_2)^2 \\ -(\alpha - \gamma_0 + 2\gamma_1 + 4\gamma_2)^2 \end{bmatrix} + f_m$. ■

REFERENCES CITED

1. Billand, P. and Bravard, C.: 2004, Non-cooperative networks in oligopolies, *International Journal of Industrial Organization*, 22(5):593–60.

2. Blonigen, B. and Taylor, C.:2000, R&D Intensity and Acquisitions in High-Technology Industries: Evidence from the US Electronic and Electrical Equipment Industries, *The Journal of Industrial Economics* 48, 47-70.

3. Cabral, L.M.B.: 2000, R&D cooperation and product market competition. *International Journal of Industrial Organization* 18, 1033–1047.

4. Goyal, S. and Joshi, S.: 2003, Networks of collaboration in oligopoly, *Games and Economic Behavior* 43, 57–85.

5. Goyal, S. and Moraga-Gonzáles, J. L.: 2001, R&d networks, *The RAND Journal of Economics* 32(4), 686–707.

6. Haeussler, C.: 2007, Proactive versus reactive M&A activities in the biotechnology industry, *The Journal of High Technology Management Research*, 17(2) 109-123.

7. Hagedoorn, J., and Lundan, S. M.: 2001, Strategic technology alliances: Trends and patterns since the early eighties. In A. Plunket, C. Voisin, & B. Bellon (Eds.), *The dynamics of industrial collaboration* (pp. 88–101). Cheltenham, UK: Edward Elgar.

8. Hagedoorn, J., Sadowski, B.: 1999, The transition from strategic technology alliances to mergers and acquisitions: an exploratory study, *Journal of Management Studies* 36, 87–107.

9. Higgins, M., Rodriguez, D.: 2006, The outsourcing of R&D through acquisition in the pharmaceutical industry. *Journal of Financial Economics* 80, 351-383.

10. Horn, H. and Persson, L.: 2001, Endogenous mergers in concentrated markets, *International Journal of Industrial Organization* 19, 1213-1244.

11. Jackson, M.,Wolinsky, A.: 1996, A strategic model of social and economic networks. *Journal of Economic Theory* 71, 44–74.

12. Kabiraj, T. and Mukherjee, A.: 2000, Cooperation in R&D and production: a three-firm analysis, *Journal of Economics (Zeitschrift fur Nationalokonomie)* 71: 281 – 304.

13. Martin, S.: 1995, R&D joint ventures and tacit product market collusion, *European Journal of Political Economy* 11, 733–741.

14. Okumura, Y.: (2009), Spatial Competition and Collaboration Networks, mimeo, Kanagawa University.

15. Salant, S., Switzer, S. and Reynolds, R.: 1983, Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium, *The Quarterly Journal of Economics* 98: 185-199.

16. Westbrock, B.: 2010, Natural concentration in industrial research collaboration. *The RAND Journal of Economics*, 41: 351–371.

CHAPTER 3.

COMPETE OR COOPERATE? ENTREPRENEURS' INFLUENCE ON VENTURE CAPITALISTS' DECISIONS.

3.1 Introduction

Cooperation among financial intermediaries is widespread. In the Venture Capital market in the US and Canada, approximately 60% of projects have more than one investor, and in Europe, 30% of ventures are conducted through cooperation (Casamatta and Haritchabalet, 2007). In this paper, we seek to understand the Venture Capitalists' decision to cooperate or to compete as well as the welfare consequences of these cooperations, also known as syndicates.

As financial intermediaries, VCs make investment decisions. As noted by Lerner (1994) and Casamatta and Haritchabalet (2007), VCs screen projects to evaluate whether to undertake a venture. Traditional financial intermediaries, such as banks, also evaluate projects, but VCs focus on industries with considerable uncertainty and hence develop specific expertise in these areas (Bygrave, 1987). VCs invest in highly risky projects and may syndicate for a better screening process. Syndication is a way to obtain access to a second opinion.

VCs also differ from traditional investors in their important managerial impact on ventures. This feature is highlighted, for example, by Amit et al. (2002) and by Casamatta (2003). VCs provide financial management, marketing knowledge and even client and supplier contacts. As noted by Kaplan and Strömberg (2003), VCs tend to become very involved in the daily life of ventures. This is their value-adding role.

Syndication may also be motivated by the desire to improve managerial assistance. Sorensen and Stuart (2001) present evidence that VCs tend to become specialists in certain industries and areas. Inviting another VC to co-invest may be a way to gain access to specific expertise. By syndicating, VCs increase the value added of the venture.

The roles that VCs assume also motivate their syndication decisions. However, we note that the syndication decision should also involve the agreement of all players, including the owner of the innovative project. The agreement of the entrepreneur is crucial. In venture financing, all players become shareholders of the project. Therefore, the original shareholder also contributes to deciding who, or under what conditions, someone may be participate in the project. The entrepreneur cannot promote cooperation among VCs because cooperation is decided solely by the VCs. However, the entrepreneur may prevent syndication by imposing competition.

By allowing for the possibility that entrepreneurs will impose competition, we can better understand their influence in the syndication and competition decisions of VCs. When deciding whether or not to syndicate, Venture Capital firms must consider the entrepreneur. In previous literature, the syndication decision has been restricted to consideration of the VCs. This new focus provides a richer description of the optimal syndication decision, and of its welfare consequences.

In our model, we take the two roles of VCs into account. We consider two risk-neutral heterogeneous VCs. When either of these VCs invests in a venture, they also undertake screening and value-adding tasks. VCs' abilities to realize these tasks differ due to different experiences. A more experienced VC may better understand a project's potential and may have a larger client network. We focus on the role of VCs and, for simplicity, we assume that the entrepreneur cannot influence the project.

We model competition as a first-price auction among VCs. After receiving a signal on the project's quality, VCs bid in terms of shares of the startup that are to be kept by the entrepreneur. A winning VC invests a fixed amount of capital, retains the stipulated shares, and subsequently exerts the managerial assistance. We characterize the unique Bayesian equilibrium of this auction. In particular, we find that the least experienced VC does not obtain profits because they are competed away. The most experienced VC obtains positive profits, which arise from his relative advantage in evaluating and providing managerial assistance. Interestingly, the profits of the most experienced VC are increasing in the investment amount. The reduction in competitive intensity more than compensates the larger investment the VC must realize if he becomes the financier. We also find that under competition, there is over-investment. Projects with an ex-ante negative expected value are financed.

Syndication is modeled as a joint proposal. Instead of competing for the right to be the financiers of the project, VCs present a joint proposal. Before the project is disclosed, the VCs propose a division of shares, stipulating the shares the entrepreneur and each VC retains if the investment is realized. When doing so, the VCs must consider the possibility that the entrepreneur will reject the joint proposal and enforce competition.

We find that syndicated deals can be more efficient than those with competition among VCs, for two reasons. First, there is no over-investment under syndication. VCs are able to jointly evaluate projects, and the information aggregation allows them to avoid projects with negative expected value. Second, there is better value-added technology. The way VCs distribute shares within a syndicate is efficient, and the best managerial abilities are achieved through the team effect. However, the fact that the entrepreneur is able to prevent syndication creates inefficiency. The entrepreneur must be compensated for the fact that there is better screening and, hence, fewer financed projects. The only way to implement this compensation is by increasing the entrepreneur's share in the venture. This leaves the syndicate with a smaller share, which reduces its incentives for value-adding services.

We find that projects with larger potential are more likely to be syndicated if the team effect is important enough. Greater potential can induce an increase in VCs' value-adding services, further increasing the gains of the extra contribution of a potential partner. When analyzing the impact of changes in investment, we find no clear consequence on the syndication decision. On the one hand, by increasing the investment, the competitive intensity for projects decreases. The most experienced VC benefits from this and may want to maintain competition. On the other hand, this situation implies that the outside option of syndication improves, leading to an improvement in the VC's position under syndication as well. The final effect is not clear. In any case, an increase in investment worsens the entrepreneur's payoff.

Government intervention in this sector is significant. In addition to taxing capital gains, governments interfere by means of co-investments. According to the European Venture Capital Association, in 2009, approximately 30% of capital funds in Europe had public origins. Our results imply that increasing the potential of the project by decreasing the tax on gains promotes welfare and increases syndications. A tax reduction policy is more effective for promoting syndication than co-investment, which has no clear effect on the likelihood of syndication.

The main empirical implication of the model is that projects with greater potential are more likely to be syndicated. We test this implication by conducting an empirical analysis on Venture Expert data. We perform a logit regression where the dependent variable is a dummy equal to one if the project has been syndicated in the early and seed investment stages. The project's potential is measured by the market-to-book values of the industry in the previous years. We find that the likelihood of syndication increases with greater expectations for a project's potential. In terms of investment, we find that larger investment is associated with more syndication.

The paper is organized as follows: in the second section we describe the industry, in the third the Individual Competition game and in the fourth the Syndication game. Section 5 compares the two games. In section 6 we draw the Policy Implications, and in section 7 we present a brief empirical analysis of the implications. Section 8 concludes. All proofs are in the appendix.

3.2 The Model

We analyze the Venture Capital market and the financing decisions of an innovative project. Innovative projects are risky. For simplicity, we assume that the project can be either successful or not. If successful, it pays a positive amount that will depend on the value added activities. If non successful, the project pays zero. The probability of success depends on the quality of the project. We assume that a good project is successful with probability p, while, for simplicity, we assume that a bad project is never successful. For the project to be realized an initial investment I is needed.

Quality is ex-ante unknown to all agents. These have a common prior about the probability that the project is good, q_0. We assume that this common prior is not too large, which prevents traditional investors to enter the market. This means that the expected profit for traditional (non-specialized) investors, only relying on the common prior is negative. This assumption is not necessary but it simplifies the model and, moreover, it seems to be very natural in this market.

The entrepreneur is the owner of the innovative project. She does not have the funds to realize the investment and she cannot influence its quality and success probability. She seeks to maximize her expected return.

We assume that there are two heterogeneous Venture Capitalists (VCs) that can finance the project. Heterogeneity arises from differences in experience, α. We denote $\overline{\alpha}$ the level of experience of the most experienced VC and $\underline{\alpha}$ that of the less experienced, with $\overline{\alpha} > \underline{\alpha} > 1/2$. We denote the most experienced VC as $VC_{\overline{\alpha}}$ and the less experienced $VC_{\underline{\alpha}}$.

VCs exert managerial effort, which increases the return of a successful project. A more experienced VC has a more productive effort. We assume that the expected return of a good project is $pRV(e)$, where $V(e)$ is the value added by VCs and R measures the projects' potential. For simplicity we assume that $V(e) = \alpha e$. The effort is non contractible and we assume that it entails a convex cost $c(e)$, with

$$c(e) = \frac{e^{1+\gamma}}{1+\gamma},$$

where $\gamma > 1$.

Besides the managerial effort VCs also evaluate the project. We assume that VCs receive a private signal about the projects quality, which allows them to, imperfectly, do the evaluation. It can be either high ($s = H$) or low ($s = L$). The precision of the signal depends on their experience. More experienced VCs are able to extract more information from the signal. Formally, the signals are such that

$$P\left(H|good\right) = \alpha = P\left(L|bad\right).$$

We also assume that the two signals are independent, conditional on the projects' quality. Hence

$$P\left(H, H|good\right) = \overline{\alpha}\underline{\alpha}.$$

After receiving his signal, a VC updates his belief about the projects' quality. Let $b_{\alpha}\left(s\right)$ be the posterior belief about the quality of the project for a VC with experience α, after receiving signal s. Hence, after receiving a high signal, a VC with experience α assigns the following probability to his belief that the project is good:

$$b_{\alpha}\left(H\right) = \frac{\alpha q_0}{\alpha q_0 + \left(1-\alpha\right)\left(1-q_0\right)}.$$

Similarly, let $b_{\alpha}\left(s, s'\right)$ be the posterior about the quality of the project of a VC with experience α, when signals are public. The first signal of the list $\left(s, s'\right)$ corresponds to the signal of the VC to whom it is referred to.

3.3 The Individual Competition Game

In the individual competition game VCs propose, or not, a contract to the entrepreneur. They do so after they have received a signal. We model the competition as an auction, where each VC "bids" a contract that specifies the share of the project that he will keep if his bid is accepted. A VC proposes a share of equity to be kept by him, and in exchange guarantees the whole investment. The entrepreneur is left with the complementary shares. We assume that VCs are not allowed to present joint proposals, or to share their private information. They compete with each other to become the sole financiers of the project.

Before going to the game, let us define the Net Present Value (NPV) of the project. The NPV is the difference between the expected return of the project and the investment and effort costs. Formally we define it as

$$NPV_{\alpha}\left(e_{\alpha};q_{0}\right)=q_{0}\left(V\left(e_{\alpha}\right)pR-\frac{e_{\alpha}^{1+\gamma}}{1+\gamma}\right)-I,$$

where q_0 is the probability that the project is good.

The NPV will depend on which player we are referring it to in two ways. First it depends on the belief about the projects' quality, updated after the signal, and second it also depends on the effort that the VC exerts.

We assume that the least experienced VC is sufficiently experienced, and hence that there is some effort such that he wants to participate in a project, after observing a high signal, i.e.

$$NPV_{\underline{\alpha}}\left(b_{\underline{\alpha}}\left(H\right)\right)=NPV_{\underline{\alpha}}\left(e_{\underline{\alpha}}^{*};b_{\underline{\alpha}}\left(H\right)\right)>0,$$

where $e_{\underline{\alpha}}^{*}$ is

$$e_{\underline{\alpha}}^{*}=\arg\max NPV_{\underline{\alpha}}\left(e_{\underline{\alpha}};b_{\underline{\alpha}}\left(H\right)\right).$$

This means that there exists some effort level that adds sufficient value as to make the project worthwhile, in expected terms. We also assume that there is some effort level such that the most experienced VC always wants to participate in a project after a high signal, even if the competitor has received a low signal, i.e.

$$NPV_{\overline{\alpha}}\left(b_{\overline{\alpha}}\left(H,L\right)\right) \equiv NPV_{\overline{\alpha}}\left(e_{\overline{\alpha}}^{'}; b_{\overline{\alpha}}\left(H,L\right)\right) > 0,$$

where $e_{\overline{\alpha}}^{'}$ is

$$e_{\overline{\alpha}}^{'} = \arg\max NPV_{\overline{\alpha}}\left(e_{\overline{\alpha}}; b_{\overline{\alpha}}\left(H,L\right)\right).$$

The timing of the game is as follows:

$t = 0$ The entrepreneur approaches the VCs, and each receives a private signal about the project.

$t = 1$ VCs simultaneously propose a share to be kept be them, or do not propose anything. If no offer was made, the game ends and all get zero profits.

$t = 2$ If an offer has been presented, the entrepreneur accepts, or not, an offer and, in the first case, a contract is signed.

$t = 3$ If a contract was signed, investment is made, quality is realized and the winning VC exerts the managerial effort.

Using backward induction, we start by analyzing the (possible) last stage of the game, where the winning VC, of experience α, chooses his effort. If the investment is realized, after a contract that stipulates a share ρ to be kept be the VC has been signed, and the project turns out to be of good quality, he solves

$$\max_{e_{\alpha}} \pi_{\alpha}\left(\rho\right) = \rho\left(\alpha e_{\alpha}\right) pR - \frac{e_{\alpha}^{1+\gamma}}{1+\gamma}.$$

The solution is the ex-post effort

$$e_{\alpha}\left(\rho\right) = \left(\rho\alpha pR\right)^{\frac{1}{\gamma}},$$

which implies expected income for the VC, gross from investment, of

$$\pi_{\alpha}(\rho) = (\rho\alpha pR)^{\frac{1+\gamma}{\gamma}} \frac{\gamma}{1+\gamma},$$

in case the project is good. If the project turns out to be bad, the investment is lost and no effort is done.

Once we know the last period choices we can analyze the acceptance decision by the entrepreneur. Her expected profit of accepting an offer $1-\rho$ is

$$\pi_E(\rho) = (1-\rho)\rho^{\frac{1}{\gamma}} (\alpha pR)^{\frac{1+\gamma}{\gamma}},$$

in case the project is good. Otherwise she has no financial return.

If only one offer is presented, then the entrepreneur will always accept it, as she incurs in no cost by doing so. If two offers are presented she will choose the one that delivers the largest expected profit.

Note that she will not necessarily choose the VC that offers her the largest share. It may be the case that she chooses a smaller share to partner with a more experienced VC.

We now characterize the equilibrium, by solving the first period of our model. Remember that VCs privately observe their signal, and afterwards present, or not, a bid-share simultaneously. Therefore, we look for the Nash Equilibrium of this auction.

In this type of bidding game with asymmetric information, where one of the bidders is more informed than the other, there is typically no equilibrium in pure strategies. We build in the work from Sharpe (1990), von Thadden (2004) and Casamatta and Haritchabalet (2008) to solve this auction. In order to simplify the proposition that describes the equilibrium let us first state some lemmas.

Lemma 3..1 *In the Individual Competition game VCs do not participate after a low signal.*

Lemma 3..2 *There is no pure strategy equilibrium in the Individual Competition game.*

Lemmas 1 and 2 are common results in the literature that analyzes these auctions. Firms do not want to participate after bad news, and the possibility to always undercut prevents the existence a pure strategy equilibrium. Using these lemmas we state the proposition that characterizes the equilibrium strategies in the Individual Competition game.

Proposition 3..1 *The equilibrium of the Individual Competition game is unique and has the following properties:*
(i) When $VC_{\overline{\alpha}}$ *receives a high signal, with probability* $\overline{x}$ *he randomizes according to a continuous distribution function* $\overline{F}\left(\rho^{\frac{1}{\gamma}}(1-\rho)\,\overline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)$ *on* $\left[\widehat{\rho}_{\underline{\alpha}}, 1\right]$*, and with probability* $1-\overline{x}$ *he plays* $\rho^{*}_{\overline{\alpha}} = 1$.
(ii) When $VC_{\underline{\alpha}}$ *receives a receives a high signal, with probability* $\underline{x}$ *he randomizes according to a continuous distribution function* $\underline{F}\left(\rho^{\frac{1}{\gamma}}(1-\rho)\,\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)$ *on* $\left[\widehat{\rho}_{\underline{\alpha}}, 1\right]$*, and with probability* $1-\underline{x}$ *he does not participate.*
(iii) When a VC receives a low signal he does not participate.
All probabilities, distribution functions and supports are constructed in the appendix.

The lower bound of the support over which VCs play their mixed strategy is the share that keeps $VC_{\underline{\alpha}}$ indifferent between (i) winning the auction with certainty, retaining $\widehat{\rho}_{\underline{\alpha}}$, and (ii) not presenting a bid. Hence $\widehat{\rho}_{\underline{\alpha}}$ is the share until which he is willing to compete for the project. As the most experienced VC can anticipate this lower bound, he will also not undercut beneath it.

As stated in the proposition, in equilibrium $VC_{\overline{\alpha}}$ will always present a bid after receiving a high signal, because the venture's expected value after observing a high signal is always positive. Moreover he plays 1 with strictly positive probability. He does this because if he would not, then $VC_{\underline{\alpha}}$ would never offer 1, as he would only win when the most experienced VC did not present an offer. This can only be true if $VC_{\overline{\alpha}}$ has observed a low signal, and hence the project has negative expected returns.

Unlike the most experienced VC, in equilibrium, $VC_{\underline{\alpha}}$ will not always present a bid after receiving a high signal. If he would always participate then $VC_{\overline{\alpha}}$ would not play 1 with strictly positive probability, as he would only win in the cases when the less experienced has received a low signal. These would still be positive profits, but the most experienced could increase the expected return by always playing the lower limit of the support.

The next corollary describes the expected profits of each player in the Individual Competition game. In case VCs have received a low signal they do not present any offer, and obtain zero profits, but in case they observe a high signal they compete as described in the proposition.

Corollary 3..1 *The expected profits, of the less experienced VC are zero and of the most experienced VC are* $E\left[\pi^{C}_{\overline{\alpha}}\right] = P\left(H\right)\left(\frac{b_{\overline{\alpha}}(H)}{b_{\underline{\alpha}}(H)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}} - 1\right) I.$

The expected profits of the entrepreneur are

$$E\left[\pi_E^C(\rho)\right] = P(good, H, H)\, E\left((1-\rho)\,\rho^{\frac{1}{\gamma}}\,(\alpha pR)^{\frac{1+\gamma}{\gamma}}\,|\overline{F}, \overline{x}, \underline{F}, \underline{x}\right) +$$
$$P(good, H, L)\, E\left((1-\rho)\,\rho^{\frac{1}{\gamma}}\,(\alpha pR)^{\frac{1+\gamma}{\gamma}}\,|\overline{F}, \overline{x}\right) +$$
$$P(good, L, H)\, E\left((1-\rho)\,\rho^{\frac{1}{\gamma}}\,(\alpha pR)^{\frac{1+\gamma}{\gamma}}\,|\underline{F}, \underline{x}\right).$$

$VC_{\underline{\alpha}}$ earns zero profits because they are competed away. On average, the gains he has from the good projects he obtains, are the same as the negative profits he gets from the bad projects he finances. The gains from the good projects are not larger due to the competition with the other VC. $VC_{\overline{\alpha}}$ has positive expected profits as he only invests in projects with positive NPV. These profits arise from two sources, both related to his experience. First he has an informational advantage, which allows him to form the belief about quality more accurately; second he also has a more productive effort.

The entrepreneur obtains positive expected profits in three cases. When both the VCs have observed a high signal, and when each of them has received a high signal, but the other has not. In the first case she has largest expected payoff, as she can choose the best out of two proposals. In the other cases she only receives one offer, and therefore has lower expected profits.

The equilibrium described in the proposition does not explicitly show which type of projects are financed, but it is possible to state which patterns the industry exhibits. These are stated in the following corollary.

Corollary 3..2 *In the Individual Competition game (i) there is no underinvestment, but (ii) the industry exhibits overinvestment.*

All projects with positive NPV are financed. As $VC_{\overline{\alpha}}$ always presents a proposal after a high signal, there are no ventures with positive expectation left unfinanced.

On the other hand, there are projects with negative expected value that are financed. There are cases where the less experienced presents an offer and wins the project, but the project has an ex-ante negative expectation. These cases are the ones where the most experienced has received a low signal and the less experienced has received a high signal. If there were a way to aggregate information, then no VC would be willing to invest in this state of nature.

There are two more sources of inefficiency in the equilibrium of the Individual Competition game. First, there are cases where the winning VC does not retain the whole project. In these cases he exerts an effort that is below the efficient one, which would only be obtained when $\rho = 1$. Second, there are cases when $VC_{\overline{\alpha}}$ looses the auction. It is the less experienced VC who will exert effort, and there could be an overall improvement if the same share would have been offered to the most experienced VC.

<u>Individual Competition Comparative Statics</u> We now analyze the consequences on the behavior of VCs of changes in some parameters. We are particularly interested in changes in both the investment and the projects' quality. We then show how these changes affect the individual expected profits.

The first lemma states how the contract proposals depend on the parameters of the model, and then we state the proposition that describes changes in profits.

Lemma 3..3 *An increase in the expected potential of the project pR leads to more* *aggressive bidding, and VCs propose, in expected terms, a smaller share ρ.*

A decrease in the investment I leads to more aggressive bidding, and VCs propose, in expected terms, a smaller share ρ.

Both the increase in project potential and the decrease in the investment make projects more desirable. Hence VCs will require a smaller share to take apart in them. Using this lemma we can describe the changes in expected profits.

Proposition 3..2 *If the expected potential of the project pR increases, then: (i) the expected profits of the entrepreneur increase and (ii) the expected profits of the VCs remain unchanged.*

If the investment I decreases, then: (i) the expected profits of the most experienced VC decrease, (ii) the expected profits of the entrepreneur increase and (iii) the expected profits of the less experienced VC remain unchanged.

An increase in the projects' potential does not lead to an increase in the VCs' profits because they are competed away. This can be inferred directly from their expected profit functions. The change leads to an increase in the share kept by the entrepreneur, and in this way she can appropriate the gains from the larger potential. The sum of expected profits of all agents increases, but there are two contradictory effects that determine this change: on the one hand the sum decreases due to the reduction of efforts caused by the reduction in the share kept by the VC, but on the other hand it increase due to the direct effect of pR on both the return and on efforts.

A decrease in investment reduces the expected profit of $VC_{\overline{\alpha}}$. This is due to the reduction of the expected share retained by VCs, induced by the more aggressive bidding. The reduction of ρ also induces a decrease in efforts, and therefore the value added services decrease, in expectation. As the ex-ante value cannot decrease with a decrease in the investment size, and $VC_{\underline{\alpha}}$ has no change in profits, it must be that the entrepreneur increases her profits.

3.4 Syndication Game

Suppose now that VCs are able to present a joint proposal. The joint proposal is a contract that specifies the shares to be kept by each player, in case the investment is realized. VCs decide among themselves how to share the part of the project the syndicate will keep, before presenting it to the entrepreneur. For simplicity we assume that the most experienced VC has all the bargaining power. In order to implement a joint proposal each VC incurs in a fixed cost F. It is the cost of monitoring the syndication partner, and hence is only incurred in case VCs are called to exert effort in the project.

We assume that the sharing rule between the two VCs does not have to be used to share the investment. We allow for different prices of equity. Let β and $1-\beta$ be the sharing rule of shares within the syndication, and θ the sharing rule of the investment. Therefore, the joint proposal is a vector of shares $(\beta\rho, (1-\beta)\,\rho, (1-\rho)\,;\theta)$. This means that, if the contract is signed, the most experienced VC retains $\beta\rho$ shares, and invests θI. The less experienced keeps $(1-\beta)\,\rho$ shares and invests $(1-\theta)\,I$, and the entrepreneur keeps $(1-\rho)$ shares.

The entrepreneur requires that the proposal is presented before she shows the project. If the joint proposal is rejected, she can still impose the Individual Competition game, by showing the project separately to each VC. If the joint proposal is accepted, VCs observe the signal and share information. After doing so, they jointly decide whether to invest or not. As before, true quality is fully disclosed after the investment.

We assume that efforts in the value added function are substitutes, i.e. $V(e) = \overline{\alpha} e_{\overline{\alpha}} + \underline{\alpha} e_{\underline{\alpha}}$. They are still non contractible, and therefore are decided independently and non cooperatively by VCs.

The decision to invest, or not, depends on the signals they have observed. We maintain the same informational requirements as in the Individual Competition game. The common prior about the quality is not sufficient to make VCs willing to invest, and it is worth to invest in a project if and only if the most experienced has received a high signal. The assumption is that there exists some β' such that the new value added function is such that, it is worthwhile investing after (H, L), i.e.

$$NPV_{\overline{\alpha},\underline{\alpha}}(b(H,L)) = \beta' NPV_{\overline{\alpha}}\left(e'_{\overline{\alpha}}; b(H,L)\right) + (1-\beta') NPV_{\underline{\alpha}}\left(e'_{\underline{\alpha}}; b(L,H)\right) > 0.$$

The timing of the game is as follows:

$t = 0$ The entrepreneur approaches the VCs with an innovative project.

$t = 1$ VCs decide how to share the project among themselves, and present a joint proposal.

$t = 2$ The entrepreneur accepts or rejects the joint proposal.

If it is accepted then the entrepreneur discloses the project jointly and each VC observes his signal and shares the information.

If it is rejected then the entrepreneur discloses the project separately, VCs also receive the signal, and it remains private information. The Individual Competition game is played.

$t = 3$ If the syndication has been accepted, VCs decide jointly whether to invest or not, given their signals and the joint proposal.

$t = 4$ If VCs do not invest, the game ends and all get zero profits. If VCs invest, quality is realized and each VCs exerts his managerial effort and also spends F.

We solve this game using backward induction. If VCs happen to invest, and the project is of high quality, then in the last period, given the stipulated shares of the contract, the most experienced VC solves

$$\max_{e_{\overline{\alpha}}} \pi_{\overline{\alpha}} = \rho\beta\left(\overline{\alpha}e_{\overline{\alpha}} + \underline{\alpha}e_{\underline{\alpha}}\right)pR - \frac{e_{\overline{\alpha}}^{1+\gamma}}{1+\gamma},$$

whereas the less experienced solves

$$\max_{e_{\underline{\alpha}}} \pi_{\underline{\alpha}} = \rho\left(1-\beta\right)\left(\overline{\alpha}e_{\overline{\alpha}} + \underline{\alpha}e_{\underline{\alpha}}\right)pR - \frac{e_{\underline{\alpha}}^{1+\gamma}}{1+\gamma}.$$

Equilibrium efforts are

$$e_{\overline{\alpha}} = \left(\rho\beta\overline{\alpha}pR\right)^{\frac{1}{\gamma}} \text{ and } e_{\underline{\alpha}} = \left(\rho\left(1-\beta\right)\underline{\alpha}pR\right)^{\frac{1}{\gamma}}.$$

Once we know the last period choices we can analyze the acceptance decision by the entrepreneur, as a function of ρ and β. Her expected profit of accepting an offer is

$$\pi_E\left(\rho,\beta\right) = \left(1-\rho\right)\rho^{\frac{1}{\gamma}}\left(\beta^{\frac{1}{\gamma}}\overline{\alpha}^{\frac{1+\gamma}{\gamma}} + \left(1-\beta\right)^{\frac{1}{\gamma}}\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)\left(pR\right)^{\frac{1+\gamma}{\gamma}},$$

in case the project is good. Otherwise she has no financial return. Note that for a constant share of the entrepreneur, syndication allows to increase the expected profit. In this sense, syndicating works as an improvement of the value added technology.

We can now construct the expected profits of the entrepreneur in the Syndication game. These are

$$E\left[\pi_E^S(\rho,\beta)\right] = P(good, H)\left[(1-\rho)\rho^{\frac{1}{\gamma}}\left(\beta^{\frac{1}{\gamma}}\overline{\alpha}^{\frac{1+\gamma}{\gamma}} + (1-\beta)^{\frac{1}{\gamma}}\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)(pR)^{\frac{1+\gamma}{\gamma}}\right]$$

where $P(good, H)$ is the probability that the project is good and that the most experienced VC has received a high signal. She will only receive a joint proposal if $VC_{\overline{\alpha}}$ has received a high signal and will only have a positive return if the project is good.

Proposition 3..3 *If the VCs choose to present a joint proposal, they will present one such that the entrepreneur is indifferent between accepting and imposing the Individual Competition game, i.e. the condition that implicitly defines the syndication equilibrium share, ρ_S, is*

$$E\left[\pi_E^S(\rho_S,\beta)\right] = E\left[\pi_E^C(\rho)\right].$$

From this proposition we infer that the entrepreneur has to be, in expected terms, indifferent between both games. This also implies that if the increase in the value added services due to syndication is not large enough, then it may be the case the entrepreneur ends up with a larger share under syndication than the expected in the Individual Competition. She has positive expected income from the overinvestment cases in the Individual Competition game. This overinvestment income is not present in the syndication game, and therefore she has to be compensated, by keeping a larger share.

The question now is whether VCs want to present a joint proposal or not. In case they do so, their profits will be a function of ρ_S, and in case they do not, they will play the Individual Competition game. Before analyzing the profits of VCs in the Syndication game let us state an efficiency result.

Lemma 3..4 *In the Syndication game, the industry exhibits nor over nor underinvestment.*

As the information is shared all existing information is taken into account in the investment decision. Hence, in informational terms, investment decisions are efficient.

The Syndicate will only invest after $VC_{\overline{\alpha}}$ has received a high signal. The expected profits of each VC, as a function of the joint proposal and the investment share θ are

$$E\left[\pi^S_{\overline{\alpha}}(\rho_S,\beta,\theta)\right] = P(good,H)\left[\beta\rho_S^{\frac{1+\gamma}{\gamma}}\begin{pmatrix}\overline{\alpha}^{\frac{1+\gamma}{\gamma}}\beta^{\frac{1}{\gamma}}\frac{\gamma}{1+\gamma}+ \\ +(1-\beta)^{\frac{1}{\gamma}}\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\end{pmatrix}(pR)^{\frac{1+\gamma}{\gamma}}-F\right]-\theta P(H)I,$$

for $VC_{\overline{\alpha}}$. The less experienced VC has the following expected profits

$$\begin{aligned}E\left[\pi^S_{\underline{\alpha}}(\rho_S,\beta,\theta)\right] &= P(good,H)\left[(1-\beta)\rho_S^{\frac{1+\gamma}{\gamma}}\begin{pmatrix}\overline{\alpha}^{\frac{1+\gamma}{\gamma}}\beta^{\frac{1}{\gamma}}+ \\ +\underline{\alpha}^{\frac{1+\gamma}{\gamma}}(1-\beta)^{\frac{1}{\gamma}}\frac{\gamma}{1+\gamma}\end{pmatrix}(pR)^{\frac{1+\gamma}{\gamma}}-F\right] \\ &\quad -(1-\theta)P(H)I.\end{aligned}$$

We now solve for the first stage of the game, where VCs decide how to share the project. As the most experienced has all the bargaining power, he offers a share $(1-\beta)$ of the final income allocated to the syndicate, in exchange for a share $(1-\theta)$ of the investment, taking into account the outside option of going to the Individual Competition game. His expected profits are

$$\begin{aligned}E\left[\pi^S_{\overline{\alpha}}(\rho_S,\beta^*)\right] &= P(good,H)\begin{bmatrix}\overline{\alpha}^{\frac{1+\gamma}{\gamma}}(\beta^*)^{\frac{1}{\gamma}}(1-\beta^*+\gamma)+ \\ +\underline{\alpha}^{\frac{1+\gamma}{\gamma}}(1-\beta^*)^{\frac{1}{\gamma}}(\beta^*+\gamma)\end{bmatrix}\frac{(pR)^{\frac{1+\gamma}{\gamma}}}{1+\gamma}\rho_S^{\frac{1+\gamma}{\gamma}} \\ &\quad -P(good,H)2F-P(H)I,\end{aligned}$$

where β^* is

$$\begin{aligned}\beta^* &= \arg\max_{\beta,\theta} E\left[\pi^S_{\overline{\alpha}}\left(\rho_S, \beta, \theta\right)\right] \\ &\quad \text{subject to} \\ &\quad E\left[\pi^S_{\underline{\alpha}}\left(\rho_S, \beta, \theta\right)\right] \geq 0 \\ &\quad E\left[\pi^S_E\left(\rho_S, \beta\right)\right] \geq E\left[\pi^C_E\left(\rho\right)\right].\end{aligned}$$

As we have not determined the equilibrium share for the syndicate ρ_S, we cannot determine the share within the syndicate. But note that both restrictions of the program are binding, and hence both the least experienced VC and the entrepreneur remain indifferent between the Individual Competition game and the Syndication game.

Lemma 3..5 *The syndicate will choose the most efficient sharing rule among the VCs.*

The most efficient sharing rule within the syndicate is the β for which the sum of profits the three players is maximized. A more efficient sharing rule within the syndicate leads to more value added services. More value added services are able to increase the share retained by the syndicate ρ_S, and a larger share retained by the syndicate increases total surplus.

This lemma is a direct consequence of the possibility of the most experienced VC to retain all the surplus. If the Syndication game is chosen, all the surplus is concentrated in one player, and hence he will propose β that maximizes it.

We have now described the equilibrium profits of the Syndication game. Both the entrepreneur and the less experienced VC are indifferent between the two games because they will always receive offers that make their participation constraint binding. The decision to present a joint proposal or not relies only on the most experienced VC.

Proposition 3..4 *There will be syndication if and only if*

$$\left[\begin{array}{c}\overline{\alpha}^{\frac{1+\gamma}{\gamma}}\left(\beta^{*}\right)^{\frac{1}{\gamma}}\left(1-\beta^{*}+\gamma\right)+\\+\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\left(1-\beta^{*}\right)^{\frac{1}{\gamma}}\left(\beta^{*}+\gamma\right)\end{array}\right]\frac{(pR)^{\frac{1+\gamma}{\gamma}}}{1+\gamma}\rho_S^{\frac{1+\gamma}{\gamma}}-2F\geq\frac{1}{b_{\underline{\alpha}}(H)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}I.$$

When evaluating whether or not to syndicate the most experienced VC will evaluate under which situation he will obtain more profits. Only his decision matters, as he is able to make the other two players indifferent between accepting his proposals or not.

<u>Syndication Comparative Statics</u> We are interested in describing how the expected profits of the different players change when both the potential of the project and investment change. We want to understand these movements, as they will determine whether the necessary and sufficient condition to syndicate is satisfied.

In order to do so, we first describe how the equilibrium syndication share ρ_S changes with these parameters.

Lemma 3..6 *An increase in the expected potential of the project pR decreases the share retained by the syndicate.*

A increase in the investment I increases the share retained by the syndicate.

This lemma is as a direct consequence of the participation constraint of the entrepreneur. Both changes increase the expected value for the entrepreneur in the Individual Competition game, and this implies that she has to hold a larger share in the Syndication game.

Proposition 3..5 *If the expected potential of the project* pR *increases, then the expected profits of the entrepreneur increase. The expected profits of* $VC_{\overline{\alpha}}$ *increase if the value-added services are sufficiently important.*
If the investment I *increases, then: (i) the expected profits of the entrepreneur decrease, (ii) the expected profits of* $VC_{\underline{\alpha}}$ *remain unchanged and (iii) the change in expected profits of* $VC_{\overline{\alpha}}$ *does not have a clear sign.*

As, in equilibrium, the entrepreneur is indifferent between the two games, her expected profits change as in the Individual Competition game. One consequence of this is that when the investment increases the entrepreneur is left with a smaller share. Therefore the syndicate is left with a larger share, and the value added services will increase. As $VC_{\underline{\alpha}}$ has no bargaining power, $VC_{\overline{\alpha}}$ absorbs all the surplus. Hence, gross of investment, the profit of the most experienced increases. The final effect on profits, which accounts the increase in investment, of the most experienced is not clear.

When the projects' potential increases then the share retained by the syndicate decreases. This in turn will give VC a smaller effort incentive. This is the indirect effect through which an increase in the potential may harm the VCs' profits. When this effect is not too strong, i.e. when the team effect is sufficient as not to decrease the syndicate's share too much, then increasing the potential increases profits. Intuitively, this case should be rather general. If it is not the case, then an increase in the potential leads to a reduction in the projects' income.

3.5 Individual Competition vs. Syndication

Now that both the Individual Competition and the Syndication games are described, it is possible to ask when will each of them prevail over the other. As the entrepreneur and the least experienced VC are indifferent between the two games, the important question is targeted to the most experienced VC. When is it more likely that he prefers to share information, shares and investment, rather than remaining under competition? We now analyze how the necessary and sufficient condition of syndication is more likely satisfied.

Proposition 3..6 *If the team effect is sufficiently important, then larger potential leads to more syndication.*

Under the stated condition, increasing the potential of the project increases the effort incentives. Under syndication, the most experienced VC is able to capture the extra surplus that results from this effort increase, whereas under competition the entrepreneur absorbs all the increase in surplus due to the increase in the potential. Therefore the syndication condition is more likely to be satisfied for larger values of the project potential.

The same result cannot be stated for changes of investment. Reducing the investment decreases the expected share kept by the VCs, and this is true both for the Syndication game and the Individual Competition game. In both of them this causes a reduction in the profits of $VC_{\overline{\alpha}}$. It is not clear in which case the reduction is larger.

In terms of efficiency, syndication has two clear positive effects, compared to competition. It avoids overinvestment and it allows a better value added technology.

But there is a third effect that has no clear sign. As a consequence of the positive profits that the entrepreneur obtains from overinvestment in the Individual Competition game, it may be the case that the share retained by the entrepreneur ends up being larger under Syndication than under Competition. This has a negative effect on efficiency, as the efforts, and therefore value added services, decrease.

If the entrepreneur would not be taken into account, she would not need to be compensated. Hence, only the two positive effects on efficiency would exist. In this case, an increase in the potential of the project would also always increase the likelihood of syndication. An increase in the investment would decrease the likelihood of syndication. The reason is the following: as the entrepreneur is not considered, an increase in investment does not increase the share kept by the syndicate. Hence the total surplus, and consequently the profits of the most experienced, decreases with the level of investment. As under competition his profits increase with the level of investment, the likelihood of syndication decreases.

3.6 Policy Implications

Public intervention in the sector is large. It is estimated that in the U.S. governments agencies account for 10% of the funds raised, and in Europe the number raises to 30%. In this way they interfere with the investment size. Besides this instrument, governments also tax capital gains, and can influence the projects' potential.

What are the implications, in terms of the syndication decision, and in terms of Welfare of the two policies? We have seen how decisions change when pR changes and when I changes. Let us apply them here.

A decrease in the investment, through co-investment policies, leads to no clear predictions, in terms of syndication. The most experienced VC may be better in the Individual Competition game or by choosing the Syndication game.

An increase in the potential of the project induces more syndication, if the team effect is strong enough.

In both cases the entrepreneur always profits from a policy intervention. Both a smaller investment and a larger potential always induce more competition between VCs, and the entrepreneur will achieve better deals. Hence if the main objective is to stimulate the entrepreneurs' side of the market, both policies are a solution.

3.7 Empirical Analysis

Introduction The main implication of our model is that projects with larger potential are more likely to be syndicated. In terms of investment size we find no clear relation. There are two concurrent effects, and we cannot say which is strongest. We now test whether implications are true.

Data Description All data was obtained in the Thomson One Banker database. We use private equity investments data from the US market, and restrict ourselves to first round investments at seed and early stages from 1971 until 2009. We also only use data on deals on industries with more than 30 deals. We obtain 9599 observations. These encompass syndicated and non syndicated deals.

Dependant Variable Our testable implications are all on the impact of differences of the potential of the project and investments on the likelihood of syndication. Therefore we compare syndicated and non syndicated deals, and hence our dependant variable is a dummy variable with value equal to one if the deal is syndicated either in the seed or early stages.

Independent Variables As a proxy for the expected potential of the project we use past Market to Book (MB) values, for the industry where the project is classified. If MB values have been large in the past recent years, then we assume that VCs form expectations on large future MB values. This is always done within each industry, and each industry is defined through a two digit SIC code. For the investment level we use the amount that was invested in that project.

Controlling Variables As controlling variables we include industry dummies and the year the investment was realized. We do so in order to capture any industry or year specific effects. We also include the number of VCs that invest for the first time in the same year as the investment is realized, in order to have a control for boom events and big shifts in expectations in the industry, and on the age of the company when it was financed. Summary statistics are presented in table 1.

Analysis and Results We use the model

$$S_i = \gamma_0 + \sum_{n=1}^{2} \gamma_n \log\left(MB_{t-n,k}\right) + \beta \log\left(equity_i\right) + \sum_{l=1}^{n} \alpha_l Z_l + \epsilon_i$$

where S_i is a dummy for syndicated deals, and Z_l are all the control variables. We use a logit specification with robust standard errors. The results are presented in table 2.

We find that the two coefficients associated with the market to book values are jointly significant. Both exhibit positive values, and therefore we conclude that projects in industries with larger returns are more probably syndicated, when compared with projects in industries with lower returns. This confirmation is line with the theoretical prediction of the base model that larger returns imply a larger gain from syndicating. We also estimate a model with only one lag of the market to book variable and it also is positive and significant. The same is true when we exclude the invested equity. We perform these three regressions to check for the robustness of the result.

We also find that increasing the equity amount invested in the project increases the probability of having a syndicated deal. The theory did not deliver a clear cut prediction, as an increase in the investment led to more profits both in the Individual Competition game and in the Syndication game. The empirical result can be encompassed within our model. Another plausible explanation that could explain this fact is that VCs use syndication as means to lower risk.

3.8 Conclusions

We present a model that compares competition and syndication, taking the entrepreneur into account. We assume that VCs perform screening and value-added and find that if the team effect in the value added roles is strong enough, projects with larger potential have a larger likelihood of syndicating. We also find that larger investments do not necessarily lead to more syndication.

In terms of welfare analysis, syndication has two positive effects: better evaluation and better value-added technology. This leads to an increase in the value added by the industry and to a reduction in overinvestment. But it also has a downside. The entrepreneur has to be compensated for forgone profits in the competition game, and therefore the share kept by the syndicate is sub-optimal.

The sub-optimal share of the syndicate and the fact that larger investments do not necessarily lead to more syndication are direct consequences of the possibility that the entrepreneur has to veto syndications. If she were not taken into account, then a syndicate would always keep the efficient share and an increase in the investment size would lead to a reduction in the likelihood of syndication.

Proof Lemma 1. $NPV_{\overline{\alpha}}\left(b_{\overline{\alpha}}\left(L\right)\right) < 0$ by assumption. Hence the most experienced does not want to participate after a low signal. To show that the less experienced VC does not want to participate it is enough to show that $NPV_{\underline{\alpha}}\left(b_{\underline{\alpha}}\left(L\right)\right) < 0$.

$$NPV_{\underline{\alpha}}\left(b_{\underline{\alpha}}\left(L\right)\right) = P\left(H_{\overline{\alpha}}|L_{\underline{\alpha}}\right) NPV_{\underline{\alpha}}\left(b_{\underline{\alpha}}\left(L,H\right)\right) + P\left(L_{\overline{\alpha}}|L_{\underline{\alpha}}\right) NPV_{\underline{\alpha}}\left(b_{\underline{\alpha}}\left(L,L\right)\right)$$

where the first term is positive and the second negative. By successive substitutions and using the fact that $\underline{\alpha} > \frac{1}{2}$ we find that $NPV_{\underline{\alpha}}\left(b_{\underline{\alpha}}\left(L\right)\right) < NPV_{\underline{\alpha}}\left(q_0\right) < 0$. ■

Proof Lemma 2. Suppose that VC_i anticipates that $\rho_j^*\left(H\right) = \rho_j^*$. The expected profit is:

$$\begin{array}{ll} P\left(L_j|H_i\right)\left(b_i\left(H,L\right)\left(\rho_i\alpha_i pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma} - I\right) & \text{if } \pi_E\left(\rho_i\right) > \pi_E\left(\rho_j^*\right) \\ P\left(L_j|H_i\right)\left(b_i\left(H,L\right)\left(\rho_i\alpha_i pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma} - I\right) + & \\ +P\left(H_j|H_i\right)\lambda\left(b_i\left(H,H\right)\left(\rho_i\alpha_i pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma} - I\right) & \text{if } \pi_E\left(\rho_i\right) = \pi_E\left(\rho_j^*\right) \\ P\left(L_j|H_i\right)\left(b_i\left(H,L\right)\left(\rho_i\alpha_i pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma} - I\right) + & \\ +P\left(H_j|H_i\right)\left(b_i\left(H,H\right)\left(\rho_i\alpha_i pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma} - I\right) & \text{if } \pi_E\left(\rho_i\right) < \pi_E\left(\rho_j^*\right) \end{array}$$

where λ denotes the probability that VC_i wins when the bid is such that the entrepreneur is indifferent. Remember that the expected value of $P\left(L_j|H_i\right)\left(b_i\left(H,L\right)\left(\rho_i\alpha_i pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma} - I\right)$ is negative for $VC_{\underline{\alpha}}$. He would rather not participate in this case. Note that when the entrepreneur is indifferent between the two proposals one can always find a profitable deviation. If λ is large enough, then VC_i does not want to deviate, but then VC_j would want to undercut.

Let us define the indifference shares $\widehat{\rho}_{\underline{\alpha}}$ and $\widehat{\rho}_{\overline{\alpha}}$. These are the shares that make the VCs indifferent between winning the auction with certainty, keeping $\widehat{\rho}$, and (i) for the most experienced VC keeping the whole project only after the competitor has received a low signal, and (ii) for the least experience VC not participating after the competitor has received a high signal.

$\widehat{\rho}_{\overline{\alpha}}$ is implicitly defined by

$$b_{\overline{\alpha}}(H,H)\left(\widehat{\rho}_{\underline{\alpha}}\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I=P\left(L_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b_{\overline{\alpha}}(H,L)\left(\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right),$$

and $\widehat{\rho}_{\underline{\alpha}}$ by

$$b_{\underline{\alpha}}(H,H)\left(\widehat{\rho}_{\underline{\alpha}}\underline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I=0.$$

Using the definition of $\widehat{\rho}$, a pure strategy equilibrium would have to involve the strategies $\widehat{\rho}_{\underline{\alpha}}$ and $\widehat{\rho}_{\overline{\alpha}}$, as these are the shares that deliver the same expected value in each state of the world. It is enough to show that one of them does not hold as a best response.

If $\rho^*_{\overline{\alpha}} < \widehat{\rho}_{\underline{\alpha}}$ then it is optimal for $VC_{\underline{\alpha}}$ not to participate, by definition. This would induce the most experienced to play $\rho^*_{\overline{\alpha}} = 1$, and hence cannot be an equilibrium. If $\rho^*_{\overline{\alpha}} > \widehat{\rho}_{\underline{\alpha}}$ then $VC_{\underline{\alpha}}$ wants deviate and to undercut $\rho^*_{\overline{\alpha}}$. Hence this cannot be an equilibrium. If $\rho^*_{\overline{\alpha}} = \widehat{\rho}_{\underline{\alpha}}$ then $VC_{\underline{\alpha}}$ would rather not participate.

This shows that there is no equilibrium in pure strategies. ■

Proof Proposition 1. The proof follows from the previous lemmas, and the following four extra steps:

Step 1: $\widehat{\rho}_{\underline{\alpha}} > \widehat{\rho}_{\overline{\alpha}}$.

Proof of step 1: First note that $\widehat{\rho}_{\underline{\alpha}} > \widehat{\rho}_{\overline{\alpha}}$. To Suppose not. Then $\widehat{\rho}_{\underline{\alpha}} < \widehat{\rho}_{\overline{\alpha}}$ which implies that

$$\frac{P\left(L_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b\left(L_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)+I}{b\left(H_{\overline{\alpha}}\right)\left(\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}}>\frac{I}{b\left(H_{\underline{\alpha}}\right)\left(\underline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}}.$$

Rearranging terms we get lower bound on the size of the expected profits:

$$(pR)^{\frac{1+\gamma}{\gamma}}>\frac{I}{b\left(L_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\overline{\alpha}\right)^{\frac{1+\gamma}{\gamma}}}\left(1+\frac{1}{P\left(L_{\underline{\alpha}}|H_{\overline{\alpha}}\right)}\left(\frac{b\left(H_{\overline{\alpha}}\right)}{b\left(H_{\underline{\alpha}}\right)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}-1\right)\right).$$

Using $NPV\left(q_0\right)<0$ we can write an upper bound for the expected returns

$$(pR)^{\frac{1+\gamma}{\gamma}}>\frac{I}{q_0\overline{\alpha}^{\frac{1+\gamma}{\gamma}}}.$$

Intersecting the two conditions and simplifying we reach the following condition

$$q_0\overline{\alpha}\left(\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}-1\right)<\left(1-q_0\right)\left(1-\underline{\alpha}\right)\left(1-\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}\right).$$

Note that the left hand side of the inequality is positive, and that the right hand is negative, by construction. This completes the contradiction.

Step 2: VCs mix over the support $\left[\widehat{\rho}_{\underline{\alpha}},1\right]$.

Proof of Step 2: Both VCs will have the same lower bound of the support in either case. If it were not so, the one that had a smaller lower bound could always increase profits by increasing the distribution mass in the interval between these lower bounds.

$VC_{\underline{\alpha}}$ will obviously not play below $\widehat{\rho}_{\underline{\alpha}}$ because that would yield negative profits (by definition).

Step 3: In equilibrium $VC_{\overline{\alpha}}$ plays 1 with strictly positive probability, and $VC_{\underline{\alpha}}$ does not participate with strictly positive probability.

Proof of step 3: If $VC_{\underline{\alpha}}$ plays 1 and $VC_{\overline{\alpha}}$ does not do it with strictly positive probability, then $VC_{\underline{\alpha}}$ has negative expected profits when playing 1. In order for

$VC_{\overline{\alpha}}$ to play 1 with strictly positive probability it must be the case that $VC_{\underline{\alpha}}$ does not participate with strictly positive probability, otherwise he would be better off by playing $\widehat{\rho}_{\overline{\alpha}}$.

Step 4: Now we just have to construct the distribution functions, and the x probabilities, as we know that VCs will play mixed strategies on the defined supports, and with the defined mass points.

To define $\underline{x}$ recall that when $VC_{\overline{\alpha}}$ bids 1 his expected profits must be:

$$\begin{aligned} & P\left(L_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(\left(b\left(L_{\underline{\alpha}},H_{\overline{\alpha}}\right)\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)+ \\ & P\left(H_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(1-\underline{x}\right)\left(\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)= \\ = & \left(\frac{b\left(H_{\overline{\alpha}}\right)}{b\left(H_{\underline{\alpha}}\right)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}-1\right)I, \end{aligned}$$

hence

$$\underline{x}=\frac{b\left(H_{\overline{\alpha}}\right)\left(\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-\left(\frac{b\left(H_{\overline{\alpha}}\right)}{b\left(H_{\underline{\alpha}}\right)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}\right)I}{P\left(H_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)}.$$

To define $\underline{F}\left(\rho^{\frac{1}{\gamma}}\left(1-\rho\right)\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)$ note that for all $\rho_{\overline{\alpha}}$ in $\left[\widehat{\rho}_{\underline{\alpha}},1\right]$ the same expected value has to hold:

$$\begin{aligned} & P\left(L_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b\left(L_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\rho\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)+ \\ & \left(\left(1-\underline{F}\left(\rho^{\frac{1}{\gamma}}\left(1-\rho\right)\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)\right)\underline{x}+\left(1-\underline{x}\right)\right)P\left(H_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\rho\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)= \\ = & \left(\frac{b\left(H_{\overline{\alpha}}\right)}{b\left(H_{\underline{\alpha}}\right)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}-1\right)I, \end{aligned}$$

hence

$$\underline{F}\left(\rho^{\frac{1}{\gamma}}(1-\rho)\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right) = \frac{b\left(H_{\overline{\alpha}}\right)\left(\rho\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-\left(\frac{b\left(H_{\overline{\alpha}}\right)}{b\left(H_{\underline{\alpha}}\right)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}\right)I}{P\left(H_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\rho\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)} * \frac{P\left(H_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)}{b\left(H_{\overline{\alpha}}\right)\left(\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-\left(\frac{b\left(H_{\overline{\alpha}}\right)}{b\left(H_{\underline{\alpha}}\right)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}\right)I}.$$

Note that $\underline{F}\left(\widehat{\rho}_{\underline{\alpha}}^{\frac{1}{\gamma}}\left(1-\widehat{\rho}_{\underline{\alpha}}\right)\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)=0$ and $\underline{F}\left(1\left(1-1\right)\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)=1$.

To calculate $\overline{x}$ and $\overline{F}\left(\rho^{\frac{1}{\gamma}}(1-\rho)\overline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)$ we follow the same method, and obtain

$$\overline{x}=\frac{b\left(H_{\underline{\alpha}}\right)\left(\underline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I}{P\left(H_{\overline{\alpha}}|H_{\underline{\alpha}}\right)\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\underline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)}$$

and

$$\overline{F}\left(\rho^{\frac{1}{\gamma}}(1-\rho)\overline{\alpha}^{\frac{1+\gamma}{\gamma}}\right) = \frac{b\left(H_{\underline{\alpha}}\right)\left(\rho\underline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I}{P\left(H_{\overline{\alpha}}|H_{\underline{\alpha}}\right)\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\rho\underline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)} * \frac{P\left(H_{\overline{\alpha}}|H_{\underline{\alpha}}\right)\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\underline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)}{b\left(H_{\underline{\alpha}}\right)\left(\underline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I}.$$

■

Proof of Corollary 1. Follows from the proof of proposition 1. ■

Proof of Corollary 2. Follows from the proof of proposition 1. ■

Proof of Lemma 3. There are two reasons why there is more aggressive bidding. First the lower bound of the support of the mixed strategies, which is

$$\widehat{\rho}_{\underline{\alpha}}=\left(\frac{I}{b_{\underline{\alpha}}\left(H,H\right)\left(\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}}\right)^{\frac{\gamma}{1+\gamma}}$$

decreases both with an increase in pR and a decrease in I. Second, using the following equation that was used to define the distribution functions,

$$P\left(L_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b\left(L_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\rho\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)+$$
$$\left(1-\underline{F}\left(\rho^{\frac{1}{\gamma}}\left(1-\rho\right)\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)\underline{x}\right)P\left(H_{\underline{\alpha}}|H_{\overline{\alpha}}\right)\left(b\left(H_{\underline{\alpha}},H_{\overline{\alpha}}\right)\left(\rho\overline{\alpha}pR\right)^{\frac{1+\gamma}{\gamma}}\frac{\gamma}{1+\gamma}-I\right)=$$
$$=\left(\frac{b\left(H_{\overline{\alpha}}\right)}{b\left(H_{\underline{\alpha}}\right)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}-1\right)I$$

we can see if pR increases, the right hand side does not change and the NPVs in the left hand side increase. This implies that $\underline{F}\left(\rho^{\frac{1}{\gamma}}\left(1-\rho\right)\underline{\alpha}^{\frac{1+\gamma}{\gamma}}\right)\underline{x}$ must increase, which means that $VC_{\underline{\alpha}}$ participates more often and gives a larger weight to smaller shares. The same can be inferred for a reduction in investment. A similar analysis may be done for a decrease in the investment size. ■

Proof Proposition 2. The expected profits of the less experienced are always zero. The expected profits of the most experienced are increasing in the investment and do not depend on the potential.

As an increase in the investment increases the profits of the most experienced and the total surplus does not increase it must be that the entrepreneur receives less expected profits. The profits of the entrepreneur increase with the potential of the project because of the more aggressive bidding. ■

Proof Proposition 3. The profits from VCs are increasing in the syndicate's share. Hence the VCs will present a proposal with the largest possible share. Therefore the participation constraint of the entrepreneur is binding. ■

Proof Lemma 4. Follow from the Information aggregation. ■

Proof Lemma 5. The most experienced can extract all the rents from the less experienced and from the entrepreneur (for some ρ_S). Hence he will choose β to maximize the contribution of the partner and to maximize the share the syndicate keeps ρ_S. Both these features are of his interest, and increase the total sum of profits. ■

Proof of Proposition 4. The condition for the most experienced VC to have more expected profits under syndication than under competition is

$$P(good, H)\left(\left[\begin{array}{c}\overline{\alpha}^{\frac{1+\gamma}{\gamma}}(\beta^*)^{\frac{1}{\gamma}}(1-\beta^*+\gamma)+\\ \underline{\alpha}^{\frac{1+\gamma}{\gamma}}(1-\beta^*)^{\frac{1}{\gamma}}(\beta^*+\gamma)\end{array}\right]\frac{(pR)^{\frac{1+\gamma}{\gamma}}}{1+\gamma}\rho_S^{\frac{1+\gamma}{\gamma}}-2F\right)-P(H)I\geq$$

$$\geq\ P(H)\frac{b_{\overline{\alpha}}(H)}{b_{\underline{\alpha}}(H)}\left(\frac{\overline{\alpha}}{\underline{\alpha}}\right)^{\frac{1+\gamma}{\gamma}}I-P(H)I,$$

which simplifies to the condition in the proposition. ■

Proof of Lemma 6. Increasing the potential increases the expected profits of the entrepreneur under competition. Hence his expected profits under syndication also have to increase. The only way to do so is by increasing his share in the venture. The opposite holds for an increase in the investment. ■

Proof of Proposition 5. The expected profits of the entrepreneur behave as in the Individual Competition case.

By the previous lemma, an increase in potential leads a reduction in the share kept by the syndicate. If the team effect is strong enough, then the share decrease has a small relative impact, and total increases. The condition is

$$\frac{\partial(pR)^{\frac{1+\gamma}{\gamma}}}{\partial R}>\frac{\partial\rho_S^{\frac{1+\gamma}{\gamma}}}{\partial R}.$$

As the most experienced retains all the surplus, his profits increase.

Increasing the investment makes the share of the syndicate increase, increasing the expected income of the project. As the investment also increases, the final result is not clear. ■

Proof of Proposition 6. If the team effect is sufficiently important, the profits of the most experienced VC under syndication increase with the project's potential. In the Individual Competition game his profits do not change for larger potential. Hence if there is syndication, it should be in large potential projects. ■

3.10 Tables

Table 3..1: Summary statistics

Variable	**Mean**	**Std. Dev.**	**N**
syndicationdummy	0.458	0.498	9599
MB_t_1	11.861	18.311	9599
ageatfinancingmonths	23.581	43.718	8423
entrants	31.909	37.427	9599
equityamountdisclosedusdmil	4.99	7.873	8815

Table 3..2: Logit Robust

Variable	1	2	3
log(MB-t-1)	0.110***	0.140***	0.068
	(0.042)	(0.043)	(0.049)
log(MB-t-2)		0.060	0.088*
		(0.044)	(0.050)
log(equity)	0.734***		0.736***
	(0.027)		(0.027)
log(age)	-0.142***	-0.075***	-0.142***
	(0.021)	(0.018)	(0.021)
year	-0.067***	-0.034***	-0.071***
	(0.005)	(0.005)	(0.006)
log(entrants)	-0.216***	-0.040*	-0.205***
	(0.024)	(0.022)	(0.025)
sic 28	0.415***	0.445***	0.402***
	(0.138)	(0.119)	(0.138)
sic 35	0.028	0.228*	0.024
	(0.150)	(0.136)	(0.150)
sic 36	0.407***	0.640***	0.416***
	(0.136)	(0.119)	(0.136)
sic 38	0.438***	0.356***	0.439***
	(0.136)	(0.119)	(0.136)
sic 48	-0.270	0.008	-0.270
	(0.169)	(0.145)	(0.169)
sic 73	0.024	0.012	0.003
	(0.118)	(0.102)	(0.118)
sic 80	0.112	0.286	0.119
	(0.211)	(0.179)	(0.212)
Constant	133.156***	68.038***	142.328***
	(10.072)	(9.932)	(11.390)
N	7620	8223	7615

REFERENCES CITED

1. Brander, J., Amit, R. and Antweiler, W. (2002), "Venture capital syndication: improved venture selection versus value-added hypothesis". *Journal of Economics and Management Strategy,* 11(3): 423–452.

2. Bygrave, W. (1987), "Syndicated investments by venture capital firms: A networking perspective", *Journal of Business Venturing*, 2: 139-154.

3. Casamatta, C. (2003), "Financing and advising: optimal financial contracts with venture capitalists", *Journal of Finance* 58(5): 2059–2086.

4. Casamatta, C. and Haritchabalet, C. (2007), "Experience, screening and syndication in venture capital investments", *Journal of Financial Intermediation*, 16: 368–398.

5. Casamatta, C. and Haritchabalet, C. (2008), "Competition between Informed Venture Capitalists and Entrepreneurs' Fund-Raising Strategy", Unpublished manuscript. University of Toulouse.

6. Hege, U., Palomino, F., and Schwienbacher, A. (2003) "Determinants of venture capital performance: Europe and the United States", Mimeo.

7. Hochberg, Y. V., Ljungqvist, A. and Lu, Y. (2010), "Networking as a Barrier to Entry and the Competitive Supply of Venture Capital". *The Journal of Finance*, 65: 829–859.

8. Kaplan, S. N. and P. Strömberg (2003), “Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts” *The Review of Economic Studies*, 70, 281-315.

9. Lerner, J. (1994), "The Syndication of Venture Capital Investments", *Financial Management,* 23-3, 16-27.

10. Sharpe, S, (1990), "Asymmetric information, bank lending and implicit contracts: A stylized model of customer relationships", *Journal of Finance* 45, 1069-1087.

11. Sorensen, O. and Stuart, T. (2001), "Syndication Networks and the Spatial Distribution of Venture Capital Investments", *American Journal of Sociology*, 106, 1546-1588.

12. Tykvova, T. (2006) "Who chooses whom? Syndication, skills and reputation". *Review of Financial Economics*, Vol. 16, No.1, 5-28.

13. von Thadden, E. L., (2004), "Asymmetric Information, Bank Lending and Implicit Contracts: the Winner’s Curse", *Finance Research Letters* 1, 11-23.

CHAPTER 4.

AN ENDOGENOUS TIMING MODEL WITH HETEROGENEOUS BELIEFS

4.1 Introduction

The failure of firms and congestion in markets is often associated with firms' optimism and lack of information at the time of entry. This idea is strengthened by results of experiments[1] that found that optimism and imperfect information can lead to excessive and too early market entry. These findings suggest that firms' perceptions and beliefs about the state of the world affect their entry decisions and competition behavior in markets. Empirical evidence also shows that executives are particular prone to display optimism[2] that affects their decisions. In particular, Malmendier and Tate (2003, 2005a, 2005b) showed that executives' optimism affects the firms investment decisions and the cash flows sensitivity. Glaser, Schafers and Weber (2008), using data from Germany, found that firms with optimistic managers invest more. In particular, CEO optimism explains larger capital expenditures while optimism of all managers increases the probability of acquisition (but CEO optimism alone does not). Therefore, there is a need to understand the effects of optimism in market decisions.

Our main research question is how heterogeneous beliefs affect the timing of market entry. We also analyze possible sources of heterogeneous beliefs and how the outcomes of the model differ in each of these explanations. The explanations analyzed

[1] See Camerer and Lovallo (1999) and Brocas and Carrillo (1999)

[2] See for example Langer (1975), Larwood and Whittaker (1977), Weinstein (1980), March and Shapira (1987).

rely on the idea that firms can choose to be optimistic (or pessimistic). Therefore, our paper also answers how optimism (or pessimism) affects entry and whether firms have incentives to become optimistic.

In our framework we analyze the firm's problem as the decision of an entrepreneur or a manager that chooses all the firm's actions. Therefore, our paper is related with the literature on strategic delegation since our model enables to evaluate the benefits and losses of hiring managers with different beliefs about the state of the world or managers that are intrinsically more optimistic or pessimistic than others. Given the importance of this literature, without loss of generality, we will interpret the results as obtained from a model where firm's decision are made by a manager that maximizes firm value.[3]

The proposed framework extends the endogenous timing model of Hamilton and Slutsky (1990) where two players compete in quantities and must decide whether to enter the market at date 1 or at date 2. Our model departs from the standard framework by assuming that firms have incomplete information about demand and by modeling the source of heterogeneous beliefs.

We pursue our analysis in a sequential way. In a first step, we consider that firms are completely uniformed about demand. They only know that demand can be high or low. Firms are Bayesian and so they have subjective beliefs about the value of demand. We allow for firms to have different beliefs and so they might "agree to

[3] We could have interpreted our results as obtained from a model with decisions made by internal organization structures where a specific behavior or beliefs emerge as dominants. In any case, since we concerned with firm behavior in markets rather than with the internal organization of firms, the specific decision process within the firm is not relevant as long as the decision maker maximizes the firm value. Therefore, we assume the inexistence of Principal-Agent problems within the firm.

disagree".

In a second step, we analyze why firms have different beliefs. We do so by adding a stage previous to the entry decisions. In that stage firms can take actions that affect their posterior beliefs. We consider two different frameworks to model this extra stage.

The first one builds in the model proposed by Brunnermeier and Parker (2005) and assumes that firms can choose subjective beliefs. These beliefs might differ from the objective beliefs (priors) and they will be used in the entry game. In this model agents choose beliefs in order to maximize happiness and so firms have optimal but possibly incorrect subjective beliefs. This is a depart from the usual rational expectations assumption. We call it the Optimal Expectations model.

The second framework extends the model proposed in Benabou and Tirole (2002). We assume there exists a period 0 where firms receive a signal about the true demand. Firms have access to a mechanism that allows them to forget the received signal. We call it the Overconfidence model. These two frameworks enable to view the model as a model where firms could choose to be optimistic or pessimistic.

We show that with Bayesian firms there exists an unique perfect Bayesian equilibrium. In that equilibrium firms with more positive beliefs about the state of the world produce in the first period while firms with more negative beliefs about the state of the world produce in the second period. Therefore, the proposed model is consistent with the empirical evidence of excessive and earlier entry of firms with more optimistic managers or entrepreneurs. These results could be extended to an environment where firms with more negative beliefs about the state of the world are defined as firms that follow a maximin rule. We show that, under the suitable assumptions, this equilibrium fits with an equilibrium derived from an Optimal Expectations Model.

We also find that when firms receive a signal about the true value of demand, but one of the firms has access to a mechanism that allows to forget that signal, equilibrium outcomes satisfying forward induction are such that in the equilibrium path the firm with the mechanism to forget does not use it. The firm that could be optimistic chooses not to be. Nevertheless, the firm with the mechanism always moves first, and so it is weakly better by having the mechanism, despite not using it. One interpretation of this result is that firms with the possibility of hiring or delegating firm's decisions to an optimistic manager are better than firms without this possibility, even if this delegation never occurs.

Our paper is essentially related with three topics of economic literature: endogenous timing decisions, optimism and strategic delegation. In classical industrial organization models, firms may play simultaneous or sequentially, but the choice of the game played by firms is normally taken as an assumption and not as a choice of firms. Endogenous timing models go beyond this weakness of classical models by endogeneizing the decision of playing a simultaneous or a sequential game. The seminal paper in the endogenous timing literature is Hamilton and Slutsky (1990)[4]. The authors propose a model with two players who must decide a quantity to be produced in one of two periods before the market clears. If a player commits to a quantity in the first period, she acts as a leader but she does not know the other player's decision. If a player waits until the second period to commit, then she observes the action of the other player in

[4]Despite we are considering Hamilton and Slutsky (1990) as the seminal paper in endogenous timing literature, we should also mention Gal-Or (1987) which analyzes first versus second mover advantages.

the first period.

Endogenous timing models have been explored by many papers in recent years. The literature has tried to find and establish conditions which might lead firms to play either a sequential-move Stackelberg game or a simultaneous-move Cournot game. Some examples are Branco (2008), van Damme and Hurkens (1999, 2004) and Normann (2002). van Damme and Hurkens (1999, 2004) analyzed endogenous timing in a duopoly model where firms have different marginal costs and compete in quantities (1999)/prices (2004). They find that, with risk dominance considerations, the efficient firm moves first, while the inefficient firm waits until the second period either for quantity or price competition. The model with different marginal costs was extended by Branco (2008) who considered that firms are privately informed about their costs. He finds that in the informative perfect Bayesian equilibrium, a firm with a low cost produces in the first period, while a firm with a high cost produces in the second period, after learning the other firm's decision in the first period. Another important paper (and closely related to ours) is Normann (2002). Normann analyzed the Hamilton and Slutsky's endogenous timing model with action commitment in a duopoly with incomplete information, in which one firm knows the state of the demand while the other remains uninformed. He finds that the Cournot equilibrium in the first period and the Stackelberg equilibria with either the informed or the uniformed firm as Stackelberg leader emerge as outcome of that game. Our paper endogeneizes the difference in the information structure and refines the results of Normann.

Our paper is also related with another important topic in economics: optimism. Optimism is something natural to human behavior and has been identified as a fundamental human impulse. Hence, we should expect economic decisions and interactions to be affected by it. In fact Heifetz, Shannon and Spiegel (2007) showed that in a large class of strategic interactions the equilibrium payoffs of optimists may be higher. This happens because the optimism of one of the players leads the adversary to change equilibrium behavior, possibly to the benefit of the optimistic player. This paper proposes that optimism may appear as tendency which takes over. Consequences of optimism have been recently formalized, among others[5], by Benabou and Tirole (2002) and Brunnermeier and Parker (2005). We use the work of these authors and apply it in the information structure of an endogenous entry decision game.

Benabou and Tirole (2002) propose a general economic model to explain why people value their self-image and how they seek to forget or preserve it through a variety of seemingly irrational behaviors. The basic idea behind the model is that individuals can affect the probability of remembering information, particularly they have "costly" mechanisms that allow them to forget bad signals and recall good news, whenever that is optimal. The ideas of selective memory or awareness management were extended by Benabou (2009) to develop a general model of groupthinking. This model tries to understand how wishful thinking and reality denial spread through organizations and markets.

Brunnermeier and Parker (2005) propose a structural model where agents can choose their beliefs in order to maximize their happiness[6]. In particular, their model assumes that before choosing their actions, agents choose the beliefs that maximize

[5]See Kyle and Wang (1997), Benabou and Tirole (1999,2006), Odean (1999), Barber and Odean (2001), Scheinkman and Xiong (2003) and Grubb (2008)

[6]Brunnermeier and Parker (2005) define happiness as the sum of the actual and future flow utilities.

their lifetime happiness and these are the beliefs used to choose subsequent actions. Therefore, in this model agents have subjective beliefs that are optimal but might be incorrect. The model allows two predictions that are opposite to two classical assumptions in economic literature: the share of a common prior by agents and the rational expectations assumption.

The interaction between the previous topics was explored by Pires and Santos-Pinto (2008). This paper considered an endogenous timing model with two firms where one is optimistic about its costs. Pires and Santos-Pinto find that for moderate levels of optimism there is a unique cost-dependent equilibrium where the optimistic player has a higher ex-ante probability of being the leader than the rational player. In this equilibrium optimism reduces the profits of the rational player but can increase the profits of the optimistic player when cost asymmetries are small.

Also closely related to our paper, De Meza and Southey (1996) propose a model with optimistic entrepreneurs that is able to rationalize some of the stylized facts of small-scale businesses. In their model most of the facts characterizing small-scale businesses, including high failure rates, reliance on bank credit rather than equity finance and credit rationing, can be explained by a tendency for those who are excessively optimistic to dominate new entries.

The seminal literature on strategic delegation (Vickers, 1985) analyzed how firms can strategically distort their managers' compensation contract away from profit maximization. Recently, some authors have analyzed the benefits and losses of employing managers with irrational behavior or different beliefs about the state of the world. Eichberger, Kelsey and Schipper (2005) found that under ambiguity optimistic or pessimistic responses to ambiguity affect behavior. Englmaier (2007) showed that it may be optimal for a firm to employ an optimistic manager because that can serve as credible commitment to get a competitive edge over the competitors.

The remainder of the paper is organized as follows. In section 2 we present the basic model. Section 3 describes and analyzes the model with unknown demand and Bayesian firm. Sections 4 and 5 provide an explanation to the heterogeneous beliefs that firms might have when the entry game starts. In section 4 we offer an explanation to the heterogeneous beliefs by extending the model proposed by Brunnermeier and Parker (2005) while in section 5 we extend the model by assuming that firms receive a signal about demand but could forget it. Section 6 concludes the paper. All proofs are in the appendix.

4.2 Model

We consider two firms that produce an homogeneous good. Firms produce with zero marginal costs and have no fixed costs. The price of the good is given by $P(q_H + q_L) \equiv \max\{\theta^* - q_H - q_L, 0\}$ where $\theta^* \in \{1, \theta\}$ and $\theta > 1$. Firms maximize expected utility and are Bayesian according to Savage's (1954) axiomatic foundations, that is, they make their choice using subjective probability distributions. There are only two subjective probability distributions: one with all mass in 1 and the other with all mass in θ. We denote a firm with the latter subjective probability by High Belief firm (H) and a firm with the former subjective probability by Low Belief firm (L). Each type of firm is uniquely defined by their subjective probability distribution on θ^*, so the types of firms are $\tau_i \in \{H, L\}$.

We consider an endogenous timing model with action commitment as in Hamilton and Slutsky (1990). Firms decide with full commitment a quantity to be produced at one of two dates. In the first period firms decide simultaneously whether to produce or not. If a firm does not produce at date 1, it must produce at date 2. Finally, at date 3, given the production decisions, the market clears.

For a firm there is a clear trade-off between the timing decisions. By producing in the first period a firm gets the possibility of becoming the leader and, in that way, influence the other firm's decision. However, by choosing the first period to produce, a firm cannot observe the other firm's decision and obtain information from that decision. Furthermore, there is the risk that the opponent also produces at date 1. On the other hand, by waiting until the second period a firm cannot influence the opponent's decision but it will have more information when it takes its decision since it can observe the quantity chosen by the rival, or the rival's decision to wait.

Summing up,the timing of the model is:

Period -1: Nature draws θ

Period 0: Each firm receives information about θ and takes an action with effect in the posterior beliefs

Period 1: Firms decide whether to commit to a particular quantity or to wait

Period 2: A firm which has not produced at date 1 decides its quantity at date 2.

Period 3: Market clears

In this section we consider that the game starts in period 1. Firms make their choice under "complete ignorance", that is, they only know that $\theta^* \in \{1, \theta\}$. We assume that firms maximize expected utility and they are Bayesian, i.e., they have a subjective probability distribution F_i on θ^*. There are only two subjective probability distributions: one with all mass in 1 and the other with all mass in θ. We assume firms might "agree to disagree", that is, firms might have different subjective beliefs and they are aware of that. Both types of firm have a common prior about the type of the other firm. Let λ be the common prior each firm has that the other firm is a High Belief type.

In this section a pure strategy for each firm is a choice of a production period $\bar{t}_i \in \{1, 2\}$ and a set of functions $\vartheta_i : \{(\bar{t}_i = 1, \bar{t}_j = 1), (1, 2), (2, 1) \times \mathbb{R}^+, (2, 2)\} \to \mathbb{R}^+$ which is firm's quantity choice as a function of both the production periods and of q_j when it is a Stackelberg follower. We assume that given the decisions to produce in period 1 or 2, firms will not mix over outputs[7]. Let $\mu_i^t(\tau_j | \bar{t}_j, \vartheta_j, \theta^*)$ be the belief of firm i in period t about the other firm's type conditional on θ^* and observed variables. The possible states of the world over which beliefs must be formed include the demand state and the competitor's type. The beliefs in period 1 are defined as $\nu_i^1(\tau_j, \theta^* | \bar{t}_j, \vartheta_j) = F_i(\theta^*) \times \mu_i^1(\tau_j | \bar{t}_j, \vartheta_j, \theta^*)$ while the beliefs in period 2 for a player that chooses to wait in period 1 are $\nu_i^2(\tau_j, \theta^* | \bar{t}_j = 2, \vartheta_j) = F_i(\theta^*) \times \mu_i^2(\tau_j | \bar{t}_j = 2, \vartheta_j, \theta^*)$ if $\bar{t}_j = 2$ and

[7] Given this definition a mixed strategy is a randomization over production periods.

$\nu_i^2(\tau_j, \theta^*|\bar{t}_j = 1, \vartheta_j) = F_i(\theta^*)$ if $\bar{t}_j = 1$. According to our definitions

$$\begin{aligned} F_i(\theta^* = 1) &= 1 \text{ if } \tau_i = L \\ F_i(\theta^* = \theta) &= 1 \text{ if } \tau_i = H \end{aligned}$$

In proposition 1, we show that there exists a perfect Bayesian equilibrium where a High Belief firm produces in the first period while a Low Belief firm produces in the second period. Next, in proposition 2 it is shown that the previous equilibrium is the unique perfect bayesian equilibrium in pure strategies.

Proposition 4..1 *If $\lambda \geq \lambda^*(\theta)$, then there exists an equilibrium with beliefs*

$$\mu_i^1\left(\tau_j = H|\bar{t}_j = 1, \vartheta_j = \frac{\theta - \frac{1-\lambda}{2}}{2\lambda+1}\right) = \lambda,$$

$$\mu_i^1\left(\tau_j = L|\bar{t}_j = 2, \vartheta_j = \frac{1-q_{1i}}{2}\right) = 1-\lambda$$

and

$$\mu_i^2\left(\tau_j = L|\bar{t}_j = 2, \vartheta_j = \frac{1}{3}\right) = 1$$

for $i, j = 1, 2$ in which the firms will have the following strategies:

1. *If firm is* High Belief *then* :

   (a) It produces $q_H = \frac{\theta - \frac{1-\lambda}{2}}{2\lambda+1}$ at date 1;

   (b) If it did not produce at date 1 and the other firm produced $\bar{q}$ at date 1, it would produce at date 2 according to $q_H = \frac{\theta - \bar{q}}{2}$;

   (c) If both firms did not produce at date 1, then it would produce $q_H = \frac{3\theta-1}{6}$ at date 2;

2. *If firm is* Low Belief *then*

(a) It produces at date 2;

(b) If it were to produce at date 1, it would produce $q_L = \frac{1+4\lambda-2\lambda\theta+\lambda^2}{6\lambda+4\lambda^2+2}$*;*

(c) If the other firm produces $\bar{q}$ *at date 1, it will produce at date 2 according to* $q_L = \frac{1-\bar{q}}{2}$*;*

(d) If both firms do not produce at date 1, then it will produce $q_L = \frac{1}{3}$ *at date 2;*

Proposition 4..2 *The equilibrium described in proposition 1 is the unique Perfect Bayesian equilibrium in pure strategies*

Propositions 1 and 2 imply that, in a Perfect Bayesian equilibrium, when λ is sufficiently high a High Belief firm produces in the first period while a Low Belief firm produces in the second period. These results are consistent with the evidence found by some experiences that more optimistic players move first. High Belief firms produce in the first period because their expected gain from the first-mover advantage is higher than the expected loss of a Stackelberg war. On the other hand, for Low Belief firms the inverse applies. This difference is explained by the different priors about market size. Note that the first-mover gain and the loss from Stackelberg war are proportional to the market size. Thus, differences in beliefs about the market size lead to different expected gains and losses.

Results described in proposition 1 and 2 are conditional on a sufficiently high value of λ. The intuition is the following. Suppose λ is low, that is, for each firm the belief that the other firm is a Low Belief firm is high. In this case, a Low Belief firm has a large incentive to deviate from the strategy proposed in proposition 1, because if it deviates, there is a high probability of becoming a Stackelberg leader instead of playing a Cournot game. Thus, there exists a high probability of achieving higher profits through the first mover advantage. On the other hand, the possible loss associated with a Stackelberg war has a low probability. Therefore, with a low λ the gains from deviating are higher than the losses.

As shown in the appendix the cutoff value of λ that enables the proposed strategy to be an equilibrium is increasing with θ. The basic idea is that the quantity produced by a High Belief firm in the first period is increasing in θ. Therefore, if a Low Belief firm deviates and produces in the first period, the cost of a Stackelberg war will be higher for it. Thus, a higher θ reduces the Low Belief firms' incentive to deviate. The table in the appendix shows that θ does not need to be very large in order to obtain an equilibrium for a large range of values of λ. For example, if the priors of both types of firms differ in 25% then there exists an equilibrium for values of λ higher than 0.17. So, we think that the prediction that an High Belief firm moves first is robust. In the appendix we also show that this result is generalizable for other non linear demand specifications.

Corollary 4..1 *If $\lambda \geq \lambda^* (\theta)$ and firms have different beliefs, then a firm with high beliefs becomes Stackelberg leader.*

All in all, the results in proposition 1 and 2 suggest that in a world where managers have different beliefs about the value of demand, firms with managers with a more positive view of the world will enter earlier in the market, as long as the market players have a sufficiently high belief that the other competitor has a positive view of the world. Firms with a manager with a high belief about demand produce a larger quantity than firms with a manager with a low belief about demand. The larger production is explained by two reasons: (i) the expectation of a greater demand and (ii) the earlier entry and consequent Stackelberg leadership advantage. These results are in line with some of the findings in Malmendier and Tate (2008) and Glaser, Schafers and Weber (2008) where it is shown that more optimistic managers invest more.

Robustness of results to beliefs specification In this subsection we analyze the sensitivity of the results to beliefs specification.

Let a Low Belief firm be defined in the same way. The previous results can be extended to a model where the High Belief firm is defined as a firm with a positive prior of state θ, but not necessarily with all mass at θ. For example, let p be the subjective probability that state is θ and $1-p$ the subjective probability that state is 1, with $p>0$. So, the mean belief of a High Belief firm is $\bar{\theta}_H = E_O\left(\theta^*\right) = p\left(\theta - 1\right) + 1$. Replacing θ by $\bar{\theta}_H$ in Proposition 1 and 2 and respective proofs, they can be applied to this new framework.

The last example shows that our results about the timing of entry can be extended to more generic models where firms have beliefs that are nondegenerate distributions on θ^*. For example, another possible generalization of the standard framework is to define the Low Belief type as the type with the lowest mean belief and the High Belief type as the type with the highest mean belief. From this generalization we obtain the following corollary from proposition 1 and 2:

Corollary 4..2 *If the distribution characterizing the beliefs of one type of firm first* *order stochastically dominates the distribution of the other type of firm, then in a Perfect Bayesian Equilibrium the latter type produces in the second period while the former produces in the first period.*

Finally, let us relax the assumption that the two types of firms are Bayesian. Suppose, one type is Bayesian with a prior $\{1-p, p\}$ on $\{1, \theta\}$ where $\theta > 1$. The other type follows a maximin criterion. It is easy to check that the problem of the type that follows a maximin criterion is equal to the problem of a Bayesian firm with a subjective probability with all mass in 1. So, one more time the results of proposition 1 and 2 can be extended to this framework (given the definition of High Belief firm we only need to replace θ by $\bar{\theta}$) where the Low Belief firm is the firm following the maximin criterion.

The previous examples show that the model might be extended to a more broad class of situations. Therefore, the model does not crucially depend on the extreme assumptions that we made on the beliefs.

Welfare Analysis Welfare analysis in this model is ambiguous, since we assume that the true state of demand and firms' beliefs are unknown. Nevertheless, some comments can be made.

Take the endogenous timing model with known demand of Hamilton and Slutsky (1990) as benchmark. It has three pure strategies equilibria: one Cournot equilibrium in the first period and two Stackelberg equilibria. The equilibrium outcome of our model with unknown demand depends on firms' beliefs and so we can have three different situations: one High Belief and one Low Belief firm, two High Belief firms, two Low Belief firms. We will present the welfare analysis for the three cases.

Case 1: one High Belief and one Low Belief firm

If the true value of θ^* is θ, then the High Belief firm is strictly better while the Low Belief firm is strictly worse than in the model with known demand. On the other hand, if the true value of θ^* is 1, then the Low Belief firm is also strictly worse but the result for the High Belief firm is ambiguous.

The intuition for these results is the following. When the Low Belief firm does not know the demand, it loses any possible first move advantage that it may have in the model with complete information. Furthermore, because the other firm is High Belief then it will overproduce and so the Low Belief firm has to decrease its production to avoid a large price decrease. Despite the reduction of the produced quantity by the Low Belief firm, the prices will decrease and so Low Belief's profits will be lower.

Concerning the High Belief firm we have that profits can increase or decrease. On one hand, in the model with unknown demand the High Belief firm has surely a first mover advantage. Furthermore, when the true value of θ^* is θ, the Stackelberg leader gain of a High Belief firm is higher than with known demand, because the follower firm produces less in the second period due to the less favorable view of the world. Thus, when the true value of θ^* is θ, the High Belief firm is better with unknown demand. On the other hand, when θ^* is 1, the High Belief firm does an optimization mistake which implies losses due to the bias in judgement. In this case the effect of unknown demand on welfare is ambiguous, because with a small bias the High Belief firm may remain better if the strategic advantage of moving first is higher than the loss due to the optimization mistake.

Case 2: two Low Belief firms

When the two firms are Low Belief and they play according to the strategies described in proposition 1, the differences in firms' profits depend on the equilibrium outcome in the standard case and the true demand parameter. If $\theta^* = 1$, then profits with unknown demand are equal to Cournot profits in the standard case. If competition is *à la* Stackelberg with complete information, then the leader (with complete information) is worse with unknown demand and the follower is better.

When $\theta^* = \theta$ and there is Cournot competition with complete information, then firms are better with unknown demand if and only if $\theta \in [1, 2]$. With unknown demand firms do a judgement bias and produce less than in the Cournot game with complete information. If firm is a Stackelberg leader with known demand, then it is always worse with incomplete information while if it is a Stackelberg follower with known demand, it is better if $\theta \in \left(1, \frac{4\sqrt{2}+8}{3}\right)$

Case 3: two High Belief firms

When firms play according to the strategies described in proposition 1 and the two firms are High Belief, the welfare analysis is as follows. When $\theta^* = \theta$, firms are better with unknown demand if and only if they are a Stackelberg follower in the complete information equilibrium and $\lambda < 0.671\,57$ and $\theta \in \left(1, \bar{\theta}\right]$ or $\lambda > 0.671\,57$ and $\theta > \bar{\theta}$ where

$$\bar{\theta} = \frac{1}{4\lambda^2 - 28\lambda + 17}\left(2\sqrt{2}\sqrt{\left(-2\lambda^2 + \lambda + 1\right)^2} - 20\lambda + 8\lambda^2 + 12\right)$$

Otherwise, firms are worse with unknown demand.

When $\theta^* = 1$, firms are better with unknown demand if and only if they are Stackelberg follower in the complete information equilibrium and $\theta \in \left(1, \frac{1}{8}\left((2\lambda + 1)\sqrt{2} + 6\right)\right)$.

Otherwise, firms are worse with unknown demand.

4.4 Optimal Expectations

The goal of the next two sections is to provide an explanation for the heterogeneous beliefs that firms may have when the entry game starts. In this section we propose an extension of the model proposed by Brunnermeier and Parker (2005), which provides a motivation for firms´ heterogeneous beliefs.

Suppose in period 0 both firms have a common prior ρ on $\theta^* \in \{1, \theta\}$ with

$$\theta < \frac{1}{18\lambda}\bar{\theta}_R \left(24\lambda + 3 + 9\lambda^2 + (1+2\lambda)\sqrt{3}\sqrt{2\lambda + 7\lambda^2 + 3}\right)$$

where $\bar{\theta}_R$ is the expectation of θ^* given the common prior, i.e.,

$$\bar{\theta}_R = \rho(1) + \theta\rho(\theta)$$

Assume that there are two types of firms. One type maximizes expected utility using the objective beliefs, the rational type (R). The second type maximizes expected utility using the "optimal"[8] beliefs that maximize its well being[9], the Optimal Expectations type (OE). The subjective beliefs that a OE firm can choose are restricted to the objective beliefs and the beliefs that give probability 1 to state $\theta^* = \theta$ [10], that is, $\hat{\pi} \in \{\rho, \bar{\pi}\}$ where $\bar{\pi}(\theta) = 1$.

[8] We consider the definition of "optimal" beliefs proposed in Brunnermeier and Parker (2005).

[9] As in Brunnermeier and Parker (2005), we define well being as the expected time-average of the happiness of the firm/manager.

[10] Notice that if we allow the subjective beliefs to be chosen over a continuous choice set, then our problem does not have solution. However, we could relax the assumption of a choice set only with two elements, because that assumption only simplifies the algebra.

A possible motivation to our assumption is to suppose that players know that only two possible distributions of states could exist

We assume each firm does not know the type of the other, and firms assign a probability λ to the possibility that the other firm is of OE type. We also assume that a OE firm knows that the beliefs of a R firm are the objective beliefs π. Thus, if a OE firm chooses subjective beliefs that differ from the objective beliefs, then firms will "agree to disagree".

In this model a pure strategy for each firm is a choice of a production period $\bar{t}_i \in \{1, 2\}$ and a set of functions $\vartheta_i : \{(\bar{t}_i = 1, \bar{t}_j = 1), (1, 2), (2, 1) \times \mathbb{R}^+, (2, 2)\} \rightarrow \mathbb{R}^+$ which is firm's quantity choice as a function of both the production periods and of q_j if it is a Stackelberg follower. The pure strategy of a OE firm also includes the choice of the subjective beliefs given the objective beliefs.

We assume that given the decisions to produce in period 1 or 2, firms will not mix over outputs[11]. The beliefs in period 1 are defined as $\nu_i^1 (\tau_j, \theta^* | \bar{t}_j, \vartheta_j) = F_i (\theta^*) \times \mu_i^1 (\tau_j | \bar{t}_j, \vartheta_j, \theta^*)$while the beliefs in period 2 for a player that chooses to wait in period 1 are $\nu_i^2 (\tau_j, \theta^* | \bar{t}_j = 2, \vartheta_j) = F_i (\theta^*) \times \mu_i^2 (\tau_j | \bar{t}_j = 2, \vartheta_j, \theta^*)$ if $\bar{t}_j = 2$ and $\nu_i^2 (\tau_j, \theta^* | \bar{t}_j = 1, \vartheta_j) = F_i (\theta^*)$ if $\bar{t}_j = 1$.

In the following proposition we show that there exists a perfect Bayesian equilibrium where a OE firm chooses to be optimistic and produces in the first period while a R firm produces in the second period.

Proposition 4..3 *If $\lambda \geq \lambda^* (\theta)$, then there exists a subgame perfect equilibrium with beliefs*

$$\mu_i^1 \left(\tau_j = OE | \bar{t}_j = 1, \vartheta_j = \frac{\theta - \frac{1-\lambda}{2}}{2\lambda + 1} \right) = \lambda,$$

[11]Given this definition a mixed strategy is a randomization over production periods.

$$\mu_i^1\left(\tau_j = R|\bar{t}_j = 2, \vartheta_j = \frac{1-q_{1i}}{2}\right) = 1-\lambda$$

and

$$\mu_i^2\left(\tau_j = OE|\bar{t}_j = 2, \vartheta_j = \frac{1}{3}\right) = 1$$

for $i, j = 1, 2$ in which firms have the following strategies:

1. *If firm is OE :*

   (a) *It chooses $\hat{\pi} = \bar{\pi}$*

   (b) *It produces $q_{OE} = \frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}$ at date 1;*

   (c) *If it did not produce at date 1 and the other firm produced $\bar{q}$ at date 1, it would produce at date 2 according to $q_{OE} = \frac{\theta - \bar{q}}{2}$;*

   (d) *If both firms did not produce at date 1, then it would produce $q_{OE} = \frac{3\theta - \bar{\theta}_R}{6}$ at date 2;*

2. *If firm is R*

   (a) *It produces at date 2;*

   (b) *If it were to produce at date 1, it would produce $q_R = \frac{\bar{\theta}_R\left(4\lambda+\lambda^2+1\right)-2\theta\lambda}{6\lambda+4\lambda^2+2}$;*

   (c) *If the other firm produces $\bar{q}$ at date 1, it will produce at date 2 according to $q_R = \frac{\bar{\theta}_R - q}{2}$, at date 2;*

   (d) *If both firms do not produce at date 1, then it will produce $q_R = \frac{\bar{\theta}_R}{3}$ at date 2.*

Proposition 3 is analog to proposition 1. The main message behind this proposition is that the equilibrium derived in section 3 fits under the suitable assumptions with an equilibrium derived from an Optimal Expectation Model. This extension to the initial model also has the attractive feature of allowing to identify a high belief firm with an optimist firm. Here optimism is a consequence of the possibility of choosing optimal beliefs instead of objective beliefs.

Finally, notice that Proposition 3 only claims existence. We are not claiming uniqueness.

4.5 Endogeneizing overconfidence

In this section we extend the model by assuming that in period 0 each firm receives a signal about the demand parameter θ^*. The signal may be either high or low, which can be interpreted, respectively, as no news and bad news. Bad news, $\sigma^i = \sigma^i_L$, are received with probability $1-p$ and no news at all, $\sigma^i = \sigma^i_H$,with probability p. We assume that the signal completely reveals the true value of demand, that is,

$$\begin{aligned} \Pr\left(\theta^* = 1 | \sigma^i = \sigma^i_L\right) &= 1 \\ \Pr\left(\theta^* = \theta | \sigma^i = \sigma^i_H\right) &= 1 \end{aligned}$$

As in Benabou and Tirole (2002), we consider that firms have access to a costly mechanism which enables to forget a signal. Let $\delta \equiv \Pr\left[\hat{\sigma} = \sigma_L | \sigma = \sigma_L\right]$ denote the probability that bad news are remembered accurately and $M(\delta)$ denote the memory cost. We will assume that the mechanism is such that $\delta \in \{0, 1\}$, i.e. firms can choose for bad news to be perfectly recalled, or completely forgotten. Furthermore, $M(0) > 0$ and $M(1) = 0$. We assume that there are two types of mechanisms, particularly, $M(\delta) \in \left\{M^L(\delta), M^H(\delta)\right\}$ where $M^L(0) < M^H(0) = \infty$ and $M^L(0) < \left(\frac{2[p\theta+(1-p)]-2\theta^*}{3}\right)^2 - \frac{\theta^*}{16}$. So, with a mechanism of type $M^H(\delta)$ a firm cannot forget. The upper bound in the mechanism for M^L guarantees that the threat of using the mechanism is always credible.

The existence of the two mechanisms can be interpreted as the existence of different managers. There are firms with optimistic managers, and these have a low cost of forgetting, and there are firms with rational managers, for whom it is impossible to forget. In the former case ignoring bad news is an option, whereas in the latter it is not.

We are going to assume that each firm's mechanism is public information. Furthermore, given our goal, we consider that each firm has a different type of mechanism and so we denote the firm with high mechanism by HM and the other firm by LM. Thus, in contrast with the previous section, here each firm knows the type of the other firm and, in each game there are always two different types of firms.

One can interpret the low mechanism firm's problem in period 0, when δ is chosen, as the choice of the game to play. In this particular case, one of the games is the standard game proposed by Hamilton and Slutsky (1990) while the other is the game with one uniformed firm and one informed firm.

In period 1 firms should take into account the reliability of their information. Therefore, when they do not recall any bad signal they should take into account the possibility that they could have forgotten it. Using Bayes rule the reliability of a "no recollection" signal is given by

$$r^* \equiv \Pr\left[\sigma = \sigma_H | \hat{\sigma} = \sigma_H, \delta^*\right] = \frac{p}{p + (1-p)(1-\delta^*)}$$

We assume that firms that choose to forget can learn the true demand in the second period if the informed firm choice of period of production or quantity produced in first period is state dependent.

In this game a pure strategy for firm HM is a function $\chi_{HM} : \theta \times \delta \rightarrow \{1, 2\}$ which is the choice of a production period as function of the signal received in period 0 and the memory awareness of the other firm and a set of functions $\vartheta_{HM} : \{(\bar{t}_{HM} = 1, \bar{t}_{LM} = 1), (1, 2), (2, 1) \times \mathbb{R}^{+}, (2, 2)\} \times \theta \times \delta \rightarrow \mathbb{R}^{+}$ which is firm's quantity choice as a function of production periods, and q_{LM} if it is a Stackelberg follower. On the other hand a pure strategy for firm LM is a function $\Delta : \theta \rightarrow \{0, 1\}$ which is firm's choice of δ as function of the signal received in period 0, a function $\chi_{LM} : \{(\delta = 0), (\delta = 1) \times \theta\} \rightarrow \{1, 2\}$ which is the choice of a production period as function of degree of memory awareness and of the signal received in period 0 if $\delta = 1$ and a set of functions $\vartheta_{LM} : \{(\bar{t}_{LM} = 1, \bar{t}_{HM} = 1), (1, 2), (2, 1) \times \mathbb{R}^{+}, (2, 2)\} \times \theta \times \delta \rightarrow \mathbb{R}^{+}$ which is firm's quantity choice as a function of production periods and of q_H when it is a Stackelberg follower.

In this application we use Subgame Perfect Nash Equilibrium as the equilibrium concept. We are also going to use the Forward Induction refinement[12] and the equilibrium refinement D1 (Cho and Kreps, 1987).

In the subgame originated by $\delta^{*} = 1$ we are in the standard game proposed by Hamilton and Slutsky (1990) and so the possible outcomes of the game are playing Cournot in the first period, or having either of the firms as a Stackelberg leader. In order to derive the equilibrium we need to understand what happens in the subgame created by the action $\delta^{*} = 0$.

[12] For a discussion of this refinement see Govindan and Wilson (2009)

Suppose the firm with low mechanism chooses $\delta^* = 0$ in period 0. Hence, in period 1 a firm with low mechanism only recalls a high signal and so its beliefs are given by

$$\begin{aligned} \Pr_{LM}\left[\sigma = \sigma_H | \hat{\sigma}_i = \sigma_H, \delta_i^*\right] &= p \\ \Pr_{LM}\left[\sigma = \sigma_L | \hat{\sigma}_i = \sigma_H, \delta_i^*\right] &= (1-p) \,. \end{aligned}$$

On the other hand, for a firm with a high mechanism the beliefs are

$$\begin{aligned} \Pr_{HM}\left[\hat{\sigma}_j = \sigma_H, \sigma = \sigma_L | \hat{\sigma}_i = \sigma_L, \delta_i^*\right] &= 1 \\ \Pr_{HM}\left[\sigma = \sigma_H | \hat{\sigma}_i = \sigma_H, \delta_i^*\right] &= 1. \end{aligned}$$

In the game that follows from $\delta^* = 0$, firms with high mechanism can be seen as informed firms while the firms with low mechanism can be seen as uniformed firms. Therefore, we have the same framework of Normann (2002) and by lemma 1 in that paper we know that in a pure-strategy equilibrium that satisfies the equilibrium refinement D1 (Cho and Kreps, 1987), all types of informed firm choose the same production period. Thus, we are going to focus our attention in pure strategy equilibria where the order of move is given by the type of mechanism. In the next two lemmas we will restrict the number of equilibria. In proposition 4 we describe the equilibrium outcomes for the complete game (a complete description of the strategy profiles that are SPNE is given in the appendix)

Lemma 4..1 *(Adapted from Normann, 2002) If $\delta^* = 0$ then both firms choosing period 2 is not an equilibrium*

Lemma 4..2 *If* $\delta^* = 0$*, then in equilibrium the firm with a low mechanism never chooses to produce in period 2. Thus, there is no separating equilibrium with* $\delta^* = 0$ *where firms with a low mechanism choose to produce in period 2.*

Proof. Suppose that there exists a separating equilibrium with $\delta^* = 0$ where firms with low mechanism choose to produce in period 2. Now, suppose a firm with low mechanism deviates and chooses $\delta^* = 1$. Thus, the firm will be informed about the value of demand. If the firm keeps producing in period 2, then the quantity produced by both firms remains the same, but the firm with low mechanism reduces its cost by not using the mechanism to forget. So, there is a gain from deviating and so we found a profitable deviation. Therefore, there is not any separating equilibrium where firms with low mechanism choose to produce in period 2 ■

In Proposition 4, rather than presenting the equilibrium strategies, we present the equilibrium outcomes. We denote the roles that firms take. C stands for Cournot competitor, L for Stackelberg leader and F for Stackelberg follower. We also specify which would be the equilibrium outcome if a different forgetting action would have been taken. In this way we can highlight the alternative for the LM firm.

Proposition 4..4 *The subgame perfect Nash equilibria of the game are characterized by the following outcomes*

$\delta = 0$, (C, C) *if* $\delta = 0$ *and* (C, C) *if* $\delta = 1$

$\delta = 0$, (C, C) *if* $\delta = 0$ *and* (F, L) *if* $\delta = 1$

$\delta = 0$, (L, F) *if* $\delta = 0$ *and* (C, C) *if* $\delta = 1$

$\delta = 0$, (L, F) *if* $\delta = 0$ *and* (F, L) *if* $\delta = 1$

$\delta = 1$, (L, F) *if* $\delta = 0$ *and* (L, F) *if* $\delta = 1$

$\delta = 1$, (C, C) *if* $\delta = 0$ *and* (L, F) *if* $\delta = 1$

$\delta = 1$, (F, L) *if* $\delta = 0$ *and* (L, F) *if* $\delta = 1$

$\delta = 1$, (F, L) *if* $\delta = 0$ *and* (C, C) *if* $\delta = 1$

$\delta = 1$, (F, L) *if* $\delta = 0$ *and* (F, L) *if* $\delta = 1$

The previous proposition shows that in this game there are several Subgame Perfect Nash Equilibria. In some of these equilibria the firm with low mechanism chooses to forget the signal.

Given the multiplicity of equilibria might be useful to understand which equilibria survive when we introduce some refinements. We use the Forward Induction refinement. Proposition 5 characterizes the outcomes of the Subgame Perfect Nash Equilibria that survive Forward Induction.

Proposition 4..5 *The outcomes*

$$\{\delta = 1, (L, F) \;\; \textit{if}\; \delta = 0 \;\; \textit{and} \;\; (L, F) \;\; \textit{if}\; \delta = 1\}$$

and

$$\{\delta = 1, (C, C) \;\; \textit{if}\; \delta = 0 \;\; \textit{and} \;\; (L, F) \;\; \textit{if}\; \delta = 1\}$$

are the unique SPNE that satisfy the forward induction refinement.

Proposition 5 implies that outcomes satisfying forward induction are such that in equilibrium the low mechanism firm does not use the mechanism, that is, the firm that could be optimist chooses not to be and does not forget any bad news. Nevertheless, the low mechanism firm always becomes a Stackelberg leader. Therefore, the low mechanism firm is weakly better by having the mechanism, despite not using it. Thus, the main insight of this result is that firms do not want to be optimist but they want to have the possibility to be.

The intuition for the result is the following. The memory awareness mechanism is a kind of signalling mechanism since it enables the firm with this mechanism to be leader. Since the other firm does not want to start a Stackelberg war, it chooses to be a follower because it knows that the firm with the mechanism will be the leader.

One interpretation of the latter results is that a firm with an optimistic manager who forgets bad news gets a strategic advantage. Even if this firm does not delegate anything to the optimistic manager, the possibility of doing so creates a competitive advantage relative to the firms without that possibility. Hence, the existence of this choice weakly improves the outcome of the firm.

4.6 Conclusion

In this paper we propose a framework that rationalizes the earlier entry of firms with more optimistic managers.

To do that we extend the endogenous timing model proposed by Hamilton and Slutsky (1990). We depart from the basic model by assuming imperfect information about demand and heterogenous beliefs about the true value of demand. In the unique Perfect Bayesian equilibria of the model, firms with more optimistic beliefs produce in the first period while firms with more pessimistic beliefs only produce in the second period. Therefore, if we interpret firms beliefs as the beliefs of their managers, we get that firms with more optimistic managers enter earlier.

In an extension of our model, we find that if one firm has an optimistic manager and the other firm has a rational manager, the firm with the optimistic manager does not delegate anything to the optimistic manager but this possibility generates a competitive advantage for the firm.

4.7 Appendix

Proof of Proposition 1:

Our proof's strategy is based in Branco (2008) and Pires and Santos-Pinto (2008).

The first step of the proof is to show that strategies are sequential rational given beliefs. Thus, we start by finding the optimal level of production for each type of firm in each contingency, assuming that each firm takes the strategy of the other firm as given.

1) Consider first the problem of a High Belief firm

i) If it produces in the first period, it may be that the other firm also has a high belief and will also produce at first period or else it will produce at the second period, if it is a low belief firm. In this case the problem of an High Belief firm is

$$\max_{q_H} \lambda \left(\theta - q_H - \frac{\theta - \frac{1-\lambda}{2}}{2\lambda + 1} \right) q_H + (1-\lambda)\left(\theta - q_H - \frac{1}{2} + \frac{1}{2} q_H \right) q_H$$

The solution to the problem is

$$q_H = \frac{\theta - (1-\lambda)\frac{1}{2}}{2\lambda + 1}$$

ii) If it produces in the second period and the other firm produced a quantity q in the first period, then it must choose the quantity that solves the problem

$$\max_{q_H} (\theta - q_H - q)\, q_H$$

Therefore the choice of the firm is

$$q_H = \frac{\theta - q}{2}$$

iii) If it produces at period 2, knowing that the other firm has not produced yet: then it infers that the other firm is a Low Belief firm and so it will produce 1/3 ; thus the High Belief firm must produce a quantity that solves the following problem:

$$\max_{q_H} \left(\theta - q_H - \frac{1}{3}\right) q_H,$$

which leads to production of:

$$q_H = \frac{3\theta - 1}{6}.$$

2. Consider the problem of a Low Belief firm

i) If it produces in the first period, it may be that the other firm is an High Belief firm and will also produce at first period or else it will produce at the second period, if it is a Low Belief firm. In this case the problem is

$$\max_{q_L} \lambda \left(1 - q_L - \frac{\theta - (1-\lambda)\frac{1}{2}}{2\lambda + 1}\right) q_L + (1-\lambda)\left(1 - q_L - \frac{1 - q_L}{2}\right) q_L.$$

The solution to this problem is:

$$\begin{aligned} q_L &= \frac{4\lambda + \lambda^2 - 2\theta\lambda + 1}{6\lambda + 4\lambda^2 + 2} \\ &= \frac{(1 + \lambda(2 - \theta))\frac{1}{(1+\lambda)} - (1-\lambda)\frac{1}{2}}{2\lambda + 1} \end{aligned}$$

(ii) If it produces at date 2, knowing that the other firm has produced the quantity q at period 1: then it must produce the quantity that solves the following problem:

$$\max_{q_L} (1 - q - q_L)\, q_L,$$

which leads to the production of:

$$q_L = \frac{1 - q}{2}.$$

(iii) If it produces at period 2, knowing that the other firm has not produced at date 1: then it infers that the other firm is also a Low Belief firm and so it will produce 1/3 at date 2; thus, it must produce a quantity that solves the following problem:

$$\max_{q_L}\left(1-q_L-\frac{1}{3}\right)q_L,$$

which leads to the production of:

$$q_L=\frac{1}{3}.$$

The optimal moment for production is determined by looking at the associated expected profits:

1. Consider the problem of an High Belief firm

a) If the High Belief firm produces in the first period its expected profit will be:

$$\begin{aligned}E\left(\pi_H^1\right) &= \lambda\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\left(\theta-\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)-\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\right)\\ &\quad+(1-\lambda)\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\left(\theta-\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)-\left(\frac{1}{2}-\frac{1}{2}\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\right)\right)\\ &= \frac{1+\lambda}{2}\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)^2\end{aligned}$$

b) If the High Belief firm produces in the second period its expected profit will be

$$\begin{aligned}E\left(\pi_H^2\right) &= \lambda\left(\frac{\theta}{2}-\frac{1}{2}\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\right)\left(\theta-\left(\frac{\theta}{2}-\frac{1}{2}\left(\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\right)-\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\\ &\quad+(1-\lambda)\left(\frac{3\theta-1}{6}\right)\left(\theta-\frac{3\theta-1}{6}-\frac{1}{3}\right)\\ &= \frac{\lambda}{4}\left(\theta-\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)^2+(1-\lambda)\left(\frac{3\theta-1}{6}\right)^2\end{aligned}$$

Thus, the difference between the payoffs is

$$E\left(\pi_H^1\right)-E\left(\pi_H^2\right)=\frac{1}{144}\frac{1-\lambda}{(2\lambda+1)^2}\left((6\theta-4)^2+\lambda(\lambda+1)(24\theta-25)-2\right)$$

and so the payoff of following the strategy and produce in period 1 is higher for all values of λ and θ.

2. Consider now the Low Belief firm problem

a) If the Low Belief firm produces in the first period its expected profit will be:

$$\begin{aligned} E\left(\pi_L^1\right) &= \lambda\frac{4\lambda+\lambda^2-2\theta\lambda+1}{6\lambda+4\lambda^2+2}\left(1-\frac{4\lambda+\lambda^2-2\theta\lambda+1}{6\lambda+4\lambda^2+2}-\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right) \\ &\quad +(1-\lambda)\frac{4\lambda+\lambda^2-2\theta\lambda+1}{6\lambda+4\lambda^2+2}\left(1-\frac{4\lambda+\lambda^2-2\theta\lambda+1}{6\lambda+4\lambda^2+2}-\left(\frac{1}{2}-\frac{1}{2}\frac{4\lambda+\lambda^2-2\theta\lambda+1}{6\lambda+4\lambda^2+2}\right)\right) \\ &= \frac{1}{8(\lambda+1)(2\lambda+1)^2}\left(4\lambda+\lambda^2-2\theta\lambda+1\right)^2 \end{aligned}$$

b) If it produces in the second period its expected profit will be

$$\begin{aligned} E\left(\pi_L^2\right) &= \lambda\left(\frac{1}{2}-\frac{1}{2}\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\left(1-\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}-\left(\frac{1}{2}-\frac{1}{2}\frac{\theta-(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)\right) \\ &\quad +(1-\lambda)\frac{1}{3}\left(1-\frac{1}{3}-\frac{1}{3}\right) \\ &= \frac{\lambda}{4(2\lambda+1)^2}\left(\frac{3\lambda+3-2\theta}{2}\right)^2+\frac{1-\lambda}{9} \end{aligned}$$

So, the difference between the payoffs is

$$E\left(\pi_L^1\right)-E\left(\pi_L^2\right)=\frac{1}{144}\frac{\lambda-1}{(\lambda+1)(2\lambda+1)^2}\Delta(\lambda)$$

where

$$\Delta(\lambda,\theta)=36\theta\lambda(\theta+\lambda-1)+\lambda\left(\lambda^2-34\lambda-1\right)-2$$

Let $\lambda^*(\theta)\equiv\inf\{\lambda\in[0,1]:\Delta(\lambda,\theta)\geq 0\}$. We can show that $\lambda^*(\theta)$ is non-empty and the payoff of following the strategy and produce in the second period is higher if

$\lambda \geq \lambda^* (\theta)$. $\lambda^* (\theta)$ is such that $\lambda^* (1) = 1$ and $\lambda^{*\prime} (\theta) < 0$ for the relevant range.[13] Given that $\lambda^* (1) = 1$ it is important to have some idea of the magnitude of the decrease of $\lambda^* (\theta)$ when θ increases, that is given by the following table

[13]Notice that we can write the difference between the payoffs as

$$\frac{1}{144} \frac{\lambda - 1}{(\lambda + 1) (2\lambda + 1)^2} \Delta (\lambda)$$

where

$$\Delta (\lambda, \theta) = 36\theta\lambda (\theta + \lambda - 1) + \lambda \left(\lambda^2 - 34\lambda - 1\right) - 2$$

Thus, the payoff of following the strategy and produce in the second period is higher if $\Delta (\lambda, \theta) > 0$. The goal is to show that there exists a $\lambda^* (\theta)$ such that for all $\lambda \geq \lambda^* (\theta)$ we have $\Delta (\lambda, \theta) > 0$.

First, define $\lambda_o (\theta) \equiv \inf \left\{\lambda \in [0, 1] : \frac{\partial \Delta}{\partial \lambda} \geq 0\right\}$. Since $\frac{\partial \Delta}{\partial \lambda}$ is continuous and

$$\frac{\partial \Delta}{\partial \lambda} (1, \theta) = 36\theta (\theta + 1) - 66 > 0$$

then $\lambda_o (\theta)$ is non-empty. Moreover, since

$$\frac{\partial^2 \Delta}{\partial \lambda^2} (\lambda, \theta) = 72\theta + 6\lambda - 68 > 0$$

then $\Delta (\lambda, \theta)$ is strictly convex in λ, which implies that $\frac{\partial \Delta}{\partial \lambda} > 0$ for $\lambda > \lambda_o (\theta)$.

Now define $\lambda^* (\theta) \equiv \inf \{\lambda \in [0, 1] : \Delta (\lambda, \theta) \geq 0\}$. Notice that

$$\Delta (1, \theta) = 36\theta^2 - 36 > 0$$

and so $\lambda^* (\theta)$ is always non-empty.

Furthermore, since $\frac{\partial \Delta}{\partial \lambda} > \Delta (\lambda, \theta)$ for $\lambda \in [0, 1]$, then $\lambda^* (\theta) > \lambda_o (\theta)$ for $\theta \geq 1$. Therefore, we know that for all $\lambda \geq \lambda^* (\theta)$ we have $\Delta (\lambda) > 0$.

Analitically is almost impossible to find this value. However, we can derive the relevant properties. With this goal define

$$F (\theta, \lambda^* (\theta)) = 36\theta\lambda (\theta + \lambda - 1) + \lambda \left(\lambda^2 - 34\lambda - 1\right) - 2 = 0$$

So,

$$\frac{\frac{\partial F}{\partial \theta}}{\frac{\partial F}{\partial \lambda}} = \frac{\partial \lambda}{\partial \theta} = -\frac{36\lambda (2\theta + \lambda - 1)}{36\theta (\theta + 2\lambda - 1) + 3\lambda^2 - 68\lambda - 1} < 0$$

for $\theta > 1$ and $\lambda \in [\lambda_o (\theta), 1]$. Thus, $\lambda^{*\prime} (\theta) < 0$ for the relevant range.

θ	$\lambda^*(\theta)$
1.01	0.88866
1.05	0.58164
1.1	0.38153
1.25	0.16534
5	0.00278

Now to wrap up the proof we only need to verify that given the strategies the beliefs proposed can be, whenever possible, updated by Bayes Rule. Particularly

$$\begin{aligned}
\mu_i^1\left(\bar{t}_j = 1, \vartheta_j = \frac{\theta - \frac{1-\lambda}{2}}{2\lambda+1}\right) &= P(\tau_i = H) = \lambda \\
\mu_i^1\left(\bar{t}_j = 2, \vartheta_j = \frac{1-q_{1i}}{2}\right) &= P(\tau_i = L) = 1-\lambda \\
\mu_i^2\left(\vartheta_j = \frac{1}{3}|\bar{t}_j = 2\right) &= \frac{P\left(\vartheta_j = \frac{1}{3}\bar{t}_j = 2\right)}{P(\bar{t}_j = 2)} = 1
\end{aligned}$$

Proof Proposition 2:

In order to prove proposition 2 we enumerate all possible strategy profiles that could be considered and explain why there cannot exist equilibria with such profiles.

Case 1: *A Low Belief firm produces at period 1 and an High Belief firm produces at period 2:*

Suppose such equilibrium exists. In that case each firm produces according to the following rule

1. *If firm is a Low Belief firm then* :

   (a) *It produces at date 1;*

(b) It produces $q_L = \frac{2-\theta\lambda}{6-4\lambda}$*;*

(c) If it did not produce at date 1 and the other firm produced $\bar{q}$ *at date 1, it would produce at date 2 according*

to $q_L = \frac{1-\bar{q}}{2}$*, at date 2;*

(d) If both firms did not produce at date 1, then it would produce $q_L = \frac{3-\theta}{6}$ *at date 2;*

2. If firm is a High Belief firm

(a) It produces at date 2;

(b) If it were to produce at date 1, it would produce $q_H = \frac{6\theta-2+2\lambda-6\theta\lambda+\theta\lambda^2}{4\lambda^2-14\lambda+12} = \frac{\theta\left(6-6\lambda+\lambda^2\right)+2(\lambda-1)}{(2-\lambda)(6-4\lambda)}$*;*

(c) If the other firm produces $\bar{q}$ *at date 1, it will produce at date 2 according to* $q_H = \frac{\theta-\bar{q}}{2}$*, at date 2;*

(d) If both firms do not produce at date 1, then it will produce $q_H = \frac{\theta}{3}$ *at date 2;*

The expected profit that the High Belief firm obtains from following this strategy profile is

$$\lambda\left(\frac{\theta}{3}\right)^2 + \frac{(1-\lambda)}{4}\left(\frac{6\theta-3\lambda\theta-2}{6-4\lambda}\right)^2$$

Now, suppose the High Belief firm deviates and produces in the first period. In that case its expected profit is $\frac{\left(6\theta+2\lambda-6\theta\lambda+\theta\lambda^2-2\right)^2}{8(2-\lambda)(2\lambda-3)^2}$. Notice that

$$E\left(\pi_H^1\right) = \frac{\left(6\theta+2\lambda-6\theta\lambda+\theta\lambda^2-2\right)^2}{8\left(2-\lambda\right)\left(2\lambda-3\right)^2} > \lambda\left(\frac{\theta}{3}\right)^2 + \frac{(1-\lambda)}{4}\left(\frac{6\theta-3\lambda\theta-2}{6-4\lambda}\right)^2 = E\left(\pi_H^2\right)$$

for all $\lambda \in [0,1]$ and $\theta > 1$ and so the High Beliefs firms always have incentives to deviate[14]

Case 2: *Both players produce at period 1, regardless of their beliefs:*

In this case an High Belief firm produces $\frac{3\theta+\lambda(1-\theta)-1}{6}$ while a Low Belief firm produces $\frac{2+\lambda(1-\theta)}{6}$. So, the expected profit of following the strategy for a Low Belief firm is $\frac{1}{36}\left(\lambda-\theta\lambda+2\right)^2$. Now suppose a Low Belief firm deviates and produces in the second period according to the following rule $\frac{1-q}{2}$, where q is the quantity produced by the other firm. So, the expected profits are

$$\begin{aligned} E\left(\pi\right) &= \lambda\left(\frac{1-\frac{3\theta+\lambda(1-\theta)-1}{6}}{2}\right)\left(1-\frac{1-\frac{3\theta+\lambda(1-\theta)-1}{6}}{2}-\frac{3\theta+\lambda\left(1-\theta\right)-1}{6}\right) \\ &\quad +\left(1-\lambda\right)\left(\frac{1-\frac{2+\lambda(1-\theta)}{6}}{2}\right)\left(1-\frac{1-\frac{2+\lambda(1-\theta)}{6}}{2}-\frac{2+\lambda\left(1-\theta\right)}{6}\right) \\ &= \frac{\lambda}{4}\left(\frac{7-3\theta-\lambda\left(1-\theta\right)}{6}\right)^2+\frac{1-\lambda}{4}\left(\frac{4-\lambda\left(1-\theta\right)}{6}\right)^2 \end{aligned}$$

---

[14] Notice that we can write the difference between the profits as

$$E\left(\pi_H^1\right)-E\left(\pi_H^2\right)=\frac{1}{144}\frac{\lambda\Delta\left(\lambda\right)}{\left(2-\lambda\right)\left(2\lambda-3\right)^2}$$

where

$$\Delta\left(\lambda,\theta\right)=\theta^2\left(\lambda^3+31\lambda^2-66\lambda+36\right)+36\left(1-\lambda\right)\left(\lambda\theta-1\right)$$

Since

$$\frac{\partial}{\partial\theta}\Delta\left(\lambda,\theta\right)=2\theta\left(\lambda^3+31\lambda^2-66\lambda+36\right)+36\left(1-\lambda\right)\lambda>0$$

for all $\lambda\in[0,1]$, the minimum value of $\Delta\left(\lambda,\theta\right)$ is achieved at $\theta=1$. Furthermore

$$\frac{\partial}{\partial\lambda}\Delta\left(\lambda,1\right)=3\lambda^2-10\lambda+6$$

and so $\Delta\left(\lambda,1\right)$ is increasing until $\lambda\simeq 0.78$ and so it decreases after $\lambda\simeq 0.78$. Therefore, for $\lambda\in[0,1]$ and $\theta>1$, the minimum value of $\Delta\left(\lambda,\theta\right)$ is achived at $(0,1)$ or $(1,1)$. Since

$$\begin{aligned} \Delta\left(0,1\right) &= 0 \\ \Delta\left(1,1\right) &= 2 \end{aligned}$$

then $\Delta\left(\lambda\right)>0$ for $\lambda\in[0,1]$ and $\theta>1$. This implies that $E\left(\pi_H^1\right)-E\left(\pi_H^2\right)>0$

Notice

$$\frac{\lambda}{4}\left(\frac{7-3\theta-\lambda\left(1-\theta\right)}{6}\right)^{2}+\frac{1-\lambda}{4}\left(\frac{4-\lambda\left(1-\theta\right)}{6}\right)^{2}>\frac{1}{36}\left(\lambda-\theta\lambda+2\right)^{2}$$

and so a Low Belief firm always have incentives to deviate and so the strategy proposed cannot be an equilibrium

Case 3: *Both players produce at period 2, regardless of their beliefs:* This cannot be an equilibrium because if both players wait regardless of their type, then they have no information gain by waiting. If they deviate by committing to a quantity at date 1 they have a first-mover advantage gain.

Proof of Proposition 3:

The first step of this proof is to show that strategies are sequential rational given beliefs. We start by setting the subjective beliefs of a OE firm equal to $\bar{\pi}$. Thus, the beliefs for a rational firm coincide with the objective beliefs while the beliefs of a OE firm are such that it gives probability 1 to state $\theta^{*}=\theta$.

We find the optimal level of production for each type of firm in each contingency, assuming that each firm take as given the strategy of the other firm. In a second stage we will show that the beliefs $\bar{\pi}$ are really the optimal beliefs.

1) Consider the problem of a OE firm

i) If it produces in the first period: it may be that the other firm is also a OE firm and will produce at first period or else it will produce at the second period, if it is a rational firm. In this case the problem of OE firm is

$$\max_{q_O}\lambda\left(\theta-q_{OE}-\frac{\theta-\bar{\theta}_R\left(1-\lambda\right)\frac{1}{2}}{2\lambda+1}\right)q_{OE}+\left(1-\lambda\right)\left(\theta-q_{OE}-\frac{\bar{\theta}_R-q_{OE}}{2}\right)q_{OE}$$

The solution to the problem is

$$q_{OE} = \frac{\theta - \bar{\theta}_R (1 - \lambda) \frac{1}{2}}{2\lambda + 1}$$

where $\bar{\theta}_R$ is the expected value of θ^* given the objective beliefs about θ^*, that is,

$$\bar{\theta}_R = \rho(1) + \rho(\theta)\,\theta$$

ii) If it produces in the second period and the other firm produced a quantity q in the first period, then it must choose the quantity that solves the problem

$$\max_{q_{OE}} (\theta - q_{OE} - q)\, q_{OE}$$

Therefore the choice of the OE firm following this strategy is

$$q_{OE} = \frac{\theta - q}{2}$$

iii) If it produces at period 2, knowing that the other firm has not produced yet: then it infers that the other firm is rational and that it will produce $\bar{\theta}_R/3$; thus the OE firm must produce a quantity that solves the following problem:

$$\max_{q_{OE}} \left(\theta - q_{OE} - \frac{\bar{\theta}_R}{3}\right) q_{OE},$$

which leads to production of:

$$q_O = \frac{3\theta - \bar{\theta}_R}{6}.$$

2. Consider the problem of R firm

i) If it produces at the first period: it may be that the other firm is OE and will produce at first period or else it will produce at the second period, if it is a R firm. In this case the problem is

$$\max_{q_R} \lambda \left(\bar{\theta}_R - q_R - \frac{\theta - \bar{\theta}_R (1-\lambda) \frac{1}{2}}{2\lambda + 1} \right) q_R + (1-\lambda) \left(\bar{\theta}_R - q_R - \frac{\bar{\theta}_R - q_R}{2} \right) q_R.$$

The solution to this problem is:

$$q_R = \frac{\bar{\theta}_R \left(4\lambda + \lambda^2 + 1\right) - 2\theta\lambda}{6\lambda + 4\lambda^2 + 2}$$

(ii) If it produces at date 2, knowing that the other firm has produced the quantity q at period 1: then it must produce the quantity that solves the following problem:

$$\max_{q_R} \left(\bar{\theta}_R - q - q_R\right) q_R,$$

which leads to production of:

$$q_R = \frac{\bar{\theta}_R - q}{2}.$$

(iii) If it produces at period 2, knowing that the other firm has not produced at date 1: then it infers that the other firm is also rational and that he will produce $\bar{\theta}_R/3$ at date 2; thus must produce a quantity that solves the following problem:

$$\max_{q_R} \left(\bar{\theta}_R - q_R - \frac{\bar{\theta}_R}{3} \right) q_R,$$

which leads to production of:

$$q_R = \frac{\bar{\theta}_R}{3}.$$

The optimal moment for production is determined by looking at the associated expected profits:

1. Consider the OE firm's problem

a) If a OE firm produces in the first period its expected profit will be:

$$\begin{aligned} E\left(\Pi_{OE}^{1}\right) &= \lambda\left(\frac{\theta-\bar{\theta}_R\left(1-\lambda\right)\frac{1}{2}}{2\lambda+1}\right)\left(\theta-\frac{\theta-\bar{\theta}_R\left(1-\lambda\right)\frac{1}{2}}{2\lambda+1}-\frac{\theta-\bar{\theta}_R\left(1-\lambda\right)\frac{1}{2}}{2\lambda+1}\right) \\ &+\left(1-\lambda\right)\left(\frac{\theta-\bar{\theta}_R\left(1-\lambda\right)\frac{1}{2}}{2\lambda+1}\right)\left(\theta-\frac{\theta-\bar{\theta}_R\left(1-\lambda\right)\frac{1}{2}}{2\lambda+1}-\frac{\bar{\theta}_R-\frac{\theta-\bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}}{2}\right) \\ &= \frac{\lambda+1}{2}\left(\frac{\theta-\bar{\theta}_R\left(1-\lambda\right)\frac{1}{2}}{2\lambda+1}\right)^2 \end{aligned}$$

b) If a OE firm produces in the second period its expected profit will be

$$\begin{aligned} E\left(\Pi_{OE}^{2}\right) &= \lambda\left(\frac{\theta-\frac{\theta-\bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}}{2}\right)\left(\theta-\frac{\theta-\frac{\theta-\bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}}{2}-\frac{\theta-\bar{\theta}_R\left(1-\lambda\right)\frac{1}{2}}{2\lambda+1}\right) \\ &+\left(1-\lambda\right)\left(\frac{3\theta-\bar{\theta}_R}{6}\right)\left(\theta-\frac{3\theta-\bar{\theta}_R}{6}-\frac{\bar{\theta}_R}{3}\right) \\ &= \lambda\left(\frac{\theta-\frac{\theta-\bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}}{2}\right)^2+\left(1-\lambda\right)\left(\frac{3\theta-\bar{\theta}_R}{6}\right)^2 \end{aligned}$$

So, the difference between the payoffs is

$$E\left(\Pi_{OE}^{1}\right)-E\left(\Pi_{OE}^{2}\right)=\frac{1}{144}\frac{1-\lambda}{\left(2\lambda+1\right)^2}\left(\begin{array}{c} 14\bar{\theta}_R^2-25\bar{\theta}_R^2\lambda^2-25\bar{\theta}_R^2\lambda+ \\ 24\bar{\theta}_R\theta\lambda^2+24\bar{\theta}_R\theta\lambda-48\bar{\theta}_R\theta+36\theta^2 \end{array}\right)$$

and so the payoff of follow the strategy and produce in period 1 is higher for all values of λ and θ.

2. Consider the R firm's problem

a) If a R firm produces in the first period its expected profit will be:

$$
\begin{aligned}
E\left(\Pi_R^1\right) &= \lambda \left(\begin{array}{c} \bar{\theta}_R - \frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1} \\ -\frac{\bar{\theta}_R\left(4\lambda+\lambda^2+1\right)-2\theta\lambda}{6\lambda+4\lambda^2+2} \end{array}\right) \left(\frac{\bar{\theta}_R\left(4\lambda+\lambda^2+1\right)-2\theta\lambda}{6\lambda+4\lambda^2+2}\right) \\
&\quad + (1-\lambda) \left(\begin{array}{c} \bar{\theta}_R - \frac{\bar{\theta}_R\left(4\lambda+\lambda^2+1\right)-2\theta\lambda}{6\lambda+4\lambda^2+2} \\ -\frac{\bar{\theta}_R - \frac{\bar{\theta}_R\left(4\lambda+\lambda^2+1\right)-2\theta\lambda}{6\lambda+4\lambda^2+2}}{2} \end{array}\right) \left(\frac{\bar{\theta}_R\left(4\lambda+\lambda^2+1\right)-2\theta\lambda}{6\lambda+4\lambda^2+2}\right) \\
&= \frac{1}{8(\lambda+1)(2\lambda+1)^2}\left(\theta_R - 2\theta\lambda + 4\lambda\theta_R + \lambda^2\theta_R\right)^2
\end{aligned}
$$

b)If it produces in the second period its expected profit will be

$$
\begin{aligned}
E\left(\Pi_R^2\right) &= \lambda \left(\bar{\theta}_R - \frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1} - \frac{\bar{\theta}_R - \frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}}{2}\right) \left(\frac{\bar{\theta}_R - \frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}}{2}\right) \\
&\quad + (1-\lambda)\left(\bar{\theta}_R - \frac{\bar{\theta}_R}{3} - \frac{\bar{\theta}_R}{3}\right)\frac{\bar{\theta}_R}{3} \\
&= \frac{\lambda}{4}\left(\bar{\theta}_R - \frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)^2 + (1-\lambda)\left(\frac{\bar{\theta}_R}{3}\right)^2
\end{aligned}
$$

So, the difference between the payoffs

$$
E\left(\Pi_R^1\right) - E\left(\Pi_R^2\right) = \frac{1}{144}\frac{\lambda-1}{(\lambda+1)(2\lambda+1)^2}\Delta(\lambda,\theta)
$$

$$
\Delta(\lambda,\theta) = 36\theta\lambda\left(\theta + \lambda\theta_R - \theta_R\right) + \lambda\theta_R^2\left(\lambda^2 - 34\lambda - 1\right) - 2\theta_R^2
$$

Let $\lambda^*(\theta) \equiv \inf\{\lambda \in [0,1] : \Delta(\lambda,\theta) \geq 0\}$. We can show that $\lambda^*(\theta)$ is non-empty and the payoff of following the strategy and produce in the second period is higher if $\lambda \geq \lambda^*(\theta)$. $\lambda^*(\theta)$ is such that $\lambda^*(1) = 1$ and $\lambda^{*\prime}(\theta) < 0$ for the relevant range.[15]

[15]Notice that we can write the difference between the payoffs as

$$
\frac{1}{144}\frac{\lambda-1}{(\lambda+1)(2\lambda+1)^2}\Delta(\lambda)
$$

The next step of this proof is to show that the beliefs $\bar{\pi}$ are really the optimal beliefs. The well being of OE firm if it chooses $\hat{\pi} = \bar{\pi}$ is

$$\begin{aligned} W(\bar{\pi}) &= \frac{\lambda+1}{4}\left(\frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)^2 - \\ &\quad - \frac{\left(12\theta^2\lambda + 4\theta^2 + 4\theta\lambda^2\theta_R - 20\theta\lambda\theta_R - 8\theta\theta_R - \lambda^3\theta_R^2 - 7\lambda^2\theta_R^2 + 5\lambda\theta_R^2 + 3\theta_R^2\right)}{16(2\lambda+1)^2} \\ &= \frac{1}{8(2\lambda+1)^2}\left(2\theta\theta_R - 4\theta^2\lambda + 10\theta\lambda\theta_R + \lambda^3\theta_R^2 + 3\lambda^2\theta_R^2 - 3\lambda\theta_R^2 - \theta_R^2\right) \end{aligned}$$

On the other hand if it chooses $\hat{\pi} = \pi$ then the well being is

$$W(\pi) = \frac{\lambda}{4}\left(\bar{\theta}_R - \frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}\right)^2 + (1-\lambda)\left(\frac{\bar{\theta}_R}{3}\right)^2$$

---

where

$$\Delta(\lambda,\theta) = 36\theta\lambda(\theta + \lambda\theta_R - \theta_R) + \lambda\theta_R^2\left(\lambda^2 - 34\lambda - 1\right) - 2\theta_R^2$$

Thus, the payoff of following the strategy and produce in the second period is higher if $\Delta(\lambda,\theta) > 0$. The goal is to show that there exists a $\lambda^*(\theta)$ such that for all $\lambda \geq \lambda^*(\theta)$ we have $\Delta(\lambda,\theta) > 0$.

First, define $\lambda_o(\theta) \equiv \inf\left\{\lambda \in [0,1] : \frac{\partial\Delta}{\partial\lambda} \geq 0\right\}$. Since $\frac{\partial\Delta}{\partial\lambda}$ is continuous and

$$\frac{\partial\Delta}{\partial\lambda}(1,\theta) = 36\theta\left(\theta + \theta_R^2\right) - 66 > 0$$

then $\lambda_o(\theta)$ is non-empty. Moreover, since

$$\frac{\partial^2\Delta}{\partial\lambda^2}(\lambda,\theta) = 2\theta_R(36\theta - 34\theta_R + 3\lambda\theta_R) > 0$$

then $\Delta(\lambda,\theta)$ is strictly convex in λ, which implies that $\frac{\partial\Delta}{\partial\lambda} > 0$ for $\lambda > \lambda_o(\theta)$.

Now define $\lambda^*(\theta) \equiv \inf\{\lambda \in [0,1] : \Delta(\lambda,\theta) \geq 0\}$. Notice that

$$\Delta(1,\theta) = 36\left(\theta^2 - \theta_R^2\right) > 0$$

and so $\lambda^*(\theta)$ is always non-empty.

Furthermore, since $\frac{\partial\Delta}{\partial\lambda} > \Delta(\lambda,\theta)$ for $\lambda \in [0,1]$, then $\lambda^*(\theta) > \lambda_o(\theta)$ for $\theta \geq 1$. Therefore, we know that for all $\lambda \geq \lambda^*(\theta)$ we have $\Delta(\lambda) > 0$.

Analitically is almost impossible to find this value. However, we can derive the relevant properties. With this goal define

$$F(\theta,\lambda^*(\theta)) = 36\theta\lambda(\theta + \lambda\theta_R - \theta_R) + \lambda\theta_R^2\left(\lambda^2 - 34\lambda - 1\right) - 2\theta_R^2 = 0$$

So,

$$\frac{\frac{\partial F}{\partial\theta}}{\frac{\partial F}{\partial\lambda}} = \frac{\partial\lambda}{\partial\theta} = -\frac{36\lambda(2\theta - \theta_R + \lambda\theta_R)}{36\theta(\theta + 2\lambda\theta_R - \theta_R) + \theta_R^2\left(3\lambda^2 - 68\lambda - 1\right)} < 0$$

for $\theta > 1$ and $\lambda \in [\lambda_o(\theta), 1]$. Thus, $\lambda^{*\prime}(\theta) < 0$ for the relevant range.

So, $W(\bar{\pi}) > W(\pi)$ if

$$\theta < \frac{1}{18\lambda}\theta_R\left(24\lambda + 3 + 9\lambda^2 + (1+2\lambda)\sqrt{3}\sqrt{2\lambda + 7\lambda^2 + 3}\right) = \bar{\theta}$$

Therefore, for $\theta < \bar{\theta}$ we have that $\hat{\pi} = \bar{\pi}$ are the "optimal" beliefs

Now to finish the proof we need to verify that given the strategies the beliefs proposed can be, whenever possible, updated by Bayes Rule. Particularly

$$\begin{aligned}
\mu_i^1\left(\bar{t}_j = 1, \vartheta_j = \frac{\theta - \bar{\theta}_R(1-\lambda)\frac{1}{2}}{2\lambda+1}\right) &= \lambda \\
\mu_i^1\left(\bar{t}_j = 2, \vartheta_j = \frac{\bar{\theta}_R - q_{1i}}{2}\right) &= 1-\lambda \\
\mu_i^2\left(\vartheta_j = \frac{\bar{\theta}_R}{3} | \bar{t}_j = 2\right) &= 1
\end{aligned}$$

Proof of Proposition 4:

In order to find the Subgame Perfect Nash Equilibrium we solve the game by backward induction.

Consider the game generated by choosing $\delta = 0$. In this case, we have a game with an informed firm and with an uniformed firm. Using the results in Normann (2002) we know that we have three pure strategies equilibria in that subgame: both firms play Cournot in the first period, Stackelberg equilibrium with uniformed firm as leader and Stackelberg equilibrium with informed firm as leader.

Now, consider the game generated by choosing $\delta = 1$. This is the standard game proposed by Hamilton and Slutsky (1990) with three pure strategies equilibria: one Cournot equilibrium in the first period and two Stackelberg equilibria.

By Lemma 2 we know that any strategy profile where firm with low mechanism chooses $\delta = 0$ and to produce in period 2 is not a Subgame Perfect Nash Equilibrium.

Thus, we can exclude these strategies from the possible equilibria candidates

Suppose we have a Cournot outcome when $\delta = 0$. Therefore, low mechanism firm's profits are higher by choosing $\delta = 0$ if when $\delta = 1$ firms either compete *à la* Cournot or low mechanism firm is Stackelberg follower. In a similar way, if low mechanism firm is Stackelberg leader when $\delta = 0$, its profits are higher by choosing $\delta = 0$ if when $\delta = 1$ firms either compete *à la* Cournot or low mechanism firm is Stackelberg follower.

Suppose low mechanism firm is Stackelberg leader when $\delta = 1$, its profits are always higher by choosing $\delta = 1$. On the other hand, if when $\delta = 1$ firms either compete *à la* Cournot or low mechanism firm is Stackelberg follower, low mechanism firm's profits are higher by choosing $\delta = 1$ if low mechanism firm is Stackelberg follower when $\delta = 1$

DESCRIPTION OF STRATEGY PROFILES THAT GENERATE THE OUTCOMES OF PROPOSITION 4

1. $\Delta = 0, \chi_{HM}(\theta^*, 0) = 1,\ \chi_{HM}(\theta^*, 1) = 1,\ \chi_{LM}(\theta^*, 0) = 1,\ \chi_{LM}(\theta^*, 1) = 1,$ $\vartheta_{HM}(1, 1, \theta^*, 0) = \frac{2\theta^* - p\theta - (1-p)}{3}, \vartheta_{LM}(1, 1, \theta^*, 0) = \frac{2[p\theta + (1-p)] - 2\theta^*}{3}, \vartheta_{HM}(1, 1, \theta^*, 1) = \vartheta_{LM}(1, 1, \theta^*, 1) = \frac{\theta^*}{3}$

2. $\Delta = 0,\ \chi_{HM}(\theta^*, 0) = 1,\ \chi_{HM}(\theta^*, 1) = 1,\ \chi_{LM}(\theta^*, 0) = 1,\ \chi_{LM}(\theta^*, 1) = 2,$ $\vartheta_{HM}(1, 1, \theta^*, 0) = \frac{2\theta^* - p\theta - (1-p)}{3}, \vartheta_{LM}(1, 1, \theta^*, 0) = \frac{2[p\theta + (1-p)] - 2\theta^*}{3}, \vartheta_{HM}(2, 1, \theta^*, 1) = \frac{\theta^*}{2}, \vartheta_{LM}(2, 1, \theta^*, 1) = \frac{\theta^*}{2} - \frac{q_{HM}}{2}$

3. $\Delta = 0,\ \chi_{HM}(\theta^*, 0) = 2,\ \chi_{HM}(\theta^*, 1) = 1,\ \chi_{LM}(\theta^*, 0) = 1,\ \chi_{LM}(\theta^*, 1) = 1,$

$\vartheta_{HM}(1,2,\theta^*,0) = \frac{\theta^*}{2} - \frac{q_{LM}}{2}$, $\vartheta_{LM}(1,2,\theta^*,0) = \frac{p\theta+(1-p)}{2}$,

$\vartheta_{HM}(1,1,\theta^*,1) = \vartheta_{LM}(1,1,\theta^*,1) = \frac{\theta^*}{3}$

4. $\Delta = 0$, $\chi_{HM}(\theta^*,0) = 2$, $\chi_{HM}(\theta^*,1) = 1$, $\chi_{LM}(\theta^*,0) = 1$, $\chi_{LM}(\theta^*,1) = 2$, $\vartheta_{HM}(1,2,\theta^*,0) = \frac{\theta^*}{2} - \frac{q_{LM}}{2}$, $\vartheta_{LM}(1,2,\theta^*,0) = \frac{p\theta+(1-p)}{2}$, $\vartheta_{HM}(2,1,\theta^*,1) = \frac{\theta^*}{2}$, $\vartheta_{LM}(2,1,\theta^*,1) = \frac{\theta^*}{2} - \frac{q_{HM}}{2}$

5. $\Delta = 1$, $\chi_{HM}(\theta^*,0) = 2$, $\chi_{HM}(\theta^*,1) = 2$, $\chi_{LM}(\theta^*,0) = 1$, $\chi_{LM}(\theta^*,1) = 1$, $\vartheta_{HM}(1,2,\theta^*,0) = \frac{\theta^*}{2} - \frac{q_{LM}}{2}$, $\vartheta_{LM}(1,2,\theta^*,0) = \frac{p\theta+(1-p)}{2}$, $\vartheta_{HM}(1,2,\theta^*,1) = \frac{\theta^*}{2} - \frac{q_{LM}}{2}$, $\vartheta_{LM}(1,2,\theta^*,1) = \frac{\theta^*}{2}$

6. $\Delta = 1$, $\chi_{HM}(\theta^*,0) = 1$, $\chi_{HM}(\theta^*,1) = 2$, $\chi_{LM}(\theta^*,0) = 1$, $\chi_{LM}(\theta^*,1) = 1$, $\vartheta_{HM}(1,1,\theta^*,0) = \frac{2\theta^*-p\theta-(1-p)}{3}$, $\vartheta_{LM}(1,1,\theta^*,0) = \frac{2[p\theta+(1-p)]-2\theta^*}{3}$, $\vartheta_{HM}(1,2,\theta^*,1) = \frac{\theta^*}{2} - \frac{q_{LM}}{2}$, $\vartheta_{LM}(1,2,\theta^*,1) = \frac{\theta^*}{2}$

7. $\Delta = 1$, $\chi_{HM}(\theta^*,0) = 1$, $\chi_{HM}(\theta^*,1) = 2$, $\chi_{LM}(\theta^*,0) = 2$, $\chi_{LM}(\theta^*,1) = 1$, $\vartheta_{HM}(2,1,\theta^*,0) = \frac{\theta^*}{2}$, $\vartheta_{LM}(2,1,\theta^*,0) = \frac{\theta^*}{2} - \frac{q_{HM}}{2}$, $\vartheta_{HM}(1,2,\theta^*,1) = \frac{\theta^*}{2} - \frac{q_{LM}}{2}$, $\vartheta_{LM}(1,2,\theta^*,1) = \frac{\theta^*}{2}$

8. $\Delta = 1$, $\chi_{HM}(\theta^*,0) = 1$, $\chi_{HM}(\theta^*,1) = 1$, $\chi_{LM}(\theta^*,0) = 2$, $\chi_{LM}(\theta^*,1) = 1$, $\vartheta_{HM}(2,1,\theta^*,0) = \frac{\theta^*}{2}$, $\vartheta_{LM}(2,1,\theta^*,0) = \frac{\theta^*}{2} - \frac{q_{HM}}{2}$,

   $\vartheta_{HM}(1,1,\theta^*,1) = \vartheta_{LM}(1,1,\theta^*,1) = \frac{\theta^*}{3}$

9. $\Delta = 1$, $\chi_{HM}(\theta^*,0) = 1$, $\chi_{HM}(\theta^*,1) = 1$, $\chi_{LM}(\theta^*,0) = 2$, $\chi_{LM}(\theta^*,1) = 2$, $\vartheta_{HM}(2,1,\theta^*,0) = \frac{\theta^*}{2}$, $\vartheta_{LM}(2,1,\theta^*,0) = \frac{\theta^*}{2} - \frac{q_{HM}}{2}$, $\vartheta_{HM}(2,1,\theta^*,1) = \frac{\theta^*}{2}$,

   $\vartheta_{LM}(2,1,\theta^*,1) = \frac{\theta^*}{2} - \frac{q_{HM}}{2}$

Proof of Proposition 5:

Fix outside option $\delta = 0$. Notice that $\pi^F_{LM}(\delta = 1) < \pi^C_{LM}(\delta = 0) < \pi^L_{LM}(\delta = 0)$, that is, if firm with low mechanism chooses $\delta = 1$ and $t = 2$, then its profit is always lower than if it had chosen $\delta = 0$. Therefore, if firm with low mechanism firm chooses $\delta = 1$, then firm with high mechanism anticipates that low mechanism firm will not choose to be a follower. So, by forward induction we can rule out the SPNE's

$$\{\delta = 0,\ (L, F)\ \textit{if}\ \delta = 0\ \textit{and}\ (F, L)\ \textit{if}\ \delta = 1\}$$

$$\{\delta = 0,\ (C, C)\ \textit{if}\ \delta = 0\ \textit{and}\ (F, L)\ \textit{if}\ \delta = 1\}$$

Fix outside option $\delta = 1$. If firm with low mechanism chooses $\delta = 0$, then it never chooses $t = 2$, because for the remain SPNE's the equilibrium profits of choose $\delta = 0$ and $t = 2$ are always lower than the equilibrium profits associated with choose $\delta = 1$. So, by Forward Induction the following SPNE can also be ruled out

$$\{\delta = 1,\ (F, L)\ \textit{if}\ \delta = 0\ \textit{and}\ (L, F)\ \textit{if}\ \delta = 1\}$$

$$\{\delta = 1,\ (F, L)\ \textit{if}\ \delta = 0\ \textit{and}\ (C, C)\ \textit{if}\ \delta = 1\}$$

$$\{\delta = 1,\ (F, L)\ \textit{if}\ \delta = 0\ \textit{and}\ (F, L)\ \textit{if}\ \delta = 1\}$$

For the remain SPNE's we have that the profits when the equilibrium outcomes is such that $\delta = 1$ are higher than the profits when the equilibrium outcomes is such that $\delta = 0$. So by forward induction we can rule out the SPNE's

$$\{\delta = 0,\ (C, C)\ \textit{if}\ \delta = 0\ \textit{and}\ (C, C)\ \textit{if}\ \delta = 1\}$$

$$\{\delta = 0,\ (L, F)\ \textit{if}\ \delta = 0\ \textit{and}\ (C, C)\ \textit{if}\ \delta = 1\}$$

Therefore, the unique SPNE's outcomes that survive the elimination of strategies that not satisfy the forward induction refinement are the outcomes

$$\{\delta = 1, (L, F) \text{ if } \delta = 0 \text{ and } (L, F) \text{ if } \delta = 1\}$$

and

$$\{\delta = 1, (C, C) \text{ if } \delta = 0 \text{ and } (L, F) \text{ if } \delta = 1\}$$

REFERENCES CITED

1. Armstrong, M. and S. Huck (2010), Behavioral Economics as Applied to Firms: A Primer, *CESIFO WORKING PAPER NO. 2937*

2. Barber, B. and T.Odean (2001), Boys will be boys: Gender, overconfidence and common stock investment, *Quarterly Journal of Economics 116, 261-292*

3. Benabou R. and J. Tirole (1999), Self-Confidence: Intrapersonal Strategies, *mimeo, Princeton University*

4. Benabou R. and J. Tirole (2002), Self-Confidence and Personal Motivation, *Quarterly Journal of Economics 470*, 871-915

5. Benabou R. and J. Tirole (2006), Incentives and Prosocial Behavior, *American Economic Review 96*, 1652-1678

6. Benabou R. (2009), Groupthinking: Collective Delusions in Organizations and Markets, *mimeo, Princeton University*

7. Branco, F.(2008), Endogenous Timing in a Quantity Setting Duopoly, *mimeo, Universidade Católica Portuguesa*

8. Brocas, I. and J. Carrillo (1999), Entry Mistakes, Entrepreneurial Boldness and Optimism, *CEPR D.P. 2213*

9. Brocas, I. and J. Carrillo (2007), Systematic errors in decision making, *mimeo,*

10. Brunnermeier, M. and J.Parker (2005), Optimal expectations, *The American Economic Review 95, 1092-1118*

11. Camerer C. and D. Lovallo (1999), Overconfidence and Excess Entry: An Experimental Approach, *American Economic Review 89(1)*, 306-318

12. Cho,I. and D.Kreps (1987), Signaling games and stable games, *Quarterly Journal of Economics 102*, 179-221

13. De Meza, D. and C.Southey (1996), The Borrower's Curse: Optimism, Finance and Entrepreneurship, *The Economic Journal 106, 375-386*

14. Eichberger, J. D.Kelsey, B.Schipper (2005), Ambiguity and Social Interaction, *Working Paper UC Davis*

15. Englmaier, F. (2007) A strategic rationale of having overoptimistic managers, *mimeo*

16. Gal-Or, E. (1987), First Mover Disadvantages with Private Information, *Review of Economic Studies 54, 279-292*

17. Gilboa, I and D.Schmeidler (1993), Updating Ambiguous Beliefs, *Journal of Economic Theory 59, 33-49*

18. Govindan, S. and R.Wilson (2009), On Forward Induction, *Econometrica 77, 1-28*

19. Grubb, M. (2009), Selling to Overconfident Consumers, *The American Economic Review (forthcoming)*

20. Hamilton, J. and S. Slutsky (1990), Endogenous Timing in Duopoly Games: Stackelberg or Cournot Equilibria, *Games and Economic Behavior 2*, 29-46

21. Heifetz, A., C.Shannon and Y. Spiegel (2000), The Dynamic Evolution of Preferences, *Economic Theory 32, 251–286*

22. Huck, S., W.Muller and H.Normann (2002), To Commit or Not to Commit: Endogenous Timing in Experimental Duopoly Markets, *Games and Economic Behavior 38, 240-264*

23. Langer, E. (1975), The Illusion of Control, *Journal of Personality and Social Psychology 32, 311-328*

24. Larwood, L. and W.Whittaker (1977), Managerial Myopia: Self-serving Biases in Organizational Planning, *Journal of Applied Psychology 62, 194-198*

25. Klemperer, P. and M.Meyer (1986), Price competition vs quantity competition: the role of uncertainty, *The RAND Journal of Economics 17, 618-638*

26. Kreps, D. and R. Wilson (1982), Sequential Equilibria, *Econometrica 50, 863-894*

27. Kyle, A. and A.Wang (1997), Speculation Duopoly with agreement to disagree: Can overconfidence survive the market test, *The Journal of Finance 52, 2073-2090*

28. Mailath, G (1993), Endogenous Sequencing of Firm Decisions, *Journal of Economic Theory 59, 169-182*

29. Malmendier, U. and G.Tate (2003), Who Makes Acquisition? CEO Overconfidence and the Market's Reaction, *Stanford GSB Research Paper #1978*

30. Malmendier, U. and G.Tate (2005), CEO Overconfidence and Corporate Investment, *Journal of Finance 60, 2661-2700*

31. Malmendier, U. and G.Tate (2005), Does Overconfidence Affect Corporate Investment? CEO Overconfidence Measures Revisited, *European Financial Management 11, 649-659*

32. March, J. and Z.Schapira (1987), Managerial Perspectives on Risk and Risk Taking, *Management Science 33, 1404-1418*

33. Matsumura, T. (1999), Quantity-setting oligopoly with Endogenous Sequencing, International Journal of Industrial Organization 17, 289-296

34. Normann, H (1997), Endogenous Stackelberg Equilibria with Incomplete Information, *Journal of Economics 66(2), 177-187*

35. Normann, H (2002), Endogenous Timing with Incomplete Information and with Observable Delay, *Games and Economic Behavior 39, 282-291*

36. Odean, T (1999), Do Investors Trade Too Much, *The American Economic Review 89, 1279-1299*

37. Ottaviani, M. and P.Sorensen (2006), The Timing of Parimutuel Bets, *mimeo*

38. Pires, T and L.Santos-Pinto (2008), Self-Confidence and Timing of Entry, *Cahiers de Recherches Economiques du Département d'Econométrie et d'Economie politique (DEEP) 09.05*

39. Rasmunsen, E. and Y. Yoon (2007), First versus Second-Mover Advantage with Information Asymmetry about the Size of New Markets, *CAEPR Working Paper 2007-017*

40. Savage, L. (1954), The Foundations of Statistics, *New York: Wiley*

41. Scheinkman, J. and W.Xiong (2003), Overconfidence and Speculative Bubbles, *Journal of Political Economy 111, 1183-1219*

42. van Damme, E. and S.Hurkens (1999), Endogenous Stackelberg Leadership, *Games and Economic Behavior 28, 105-129*

43. van Damme, E. and S.Hurkens (2004), Endogenous Price Leadership, *Games and Economic Behavior 47, 404-420*

44. Vickers, J. (1985), Delegation and the Theory of the Firm, *Economic Journal Supplement 95, 138-147*

45. Weinstein, N. (1980), Unrealistic Optimism about Failure Life Events, *Journal of Personality and Social Psychology 39, 806-820*