Spring 5-15-2018

Essays in Knowledge Transfer

Leonardo Mayer Kluppel
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Olin Business School

Dissertation Examination Committee:
Anne Marie Knott, Chair
Nicholas S. Argyres
Lamar Pierce
Michael D. Ryall
Ulya Tsolmon

Essays in Knowledge Transfer

by

Leonardo Mayer Kluppel

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2018
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

This dissertation would not be possible without the direct and indirect support of many people. First, I would like to thank Anne Marie Knott for her guidance and support. She was key in my development as a scholar. I want to extend an especial thanks to Lamar Pierce, who helped me immensely during all my academic journey. I am also indebted to my dissertation committee for their comments, time, and patience: Nick Argyres, Michael Ryall, and Ulya Tsolmon.

I am grateful for all the knowledge spillovers resulting from being surrounded by great scholars, in particular the strategy department: Dan Elfenbein, Jackson Nickerson, Phech Colatat, and Tarek Ghani. Todd Zenger and Adina Sterling also left a lasting impression on my academic career.

The support from friends is impossible to understate. Special thanks to those that were closer to me during my PhD: Trey, Sarah, Bala, and Srinidhi. Finally, I need to thank my family. Even though they have no idea what I am doing, they have always supported me.

Leonardo Mayer Kluppel

Washington University in St. Louis

May 2018

Dedicated to my family.

ABSTRACT OF THE DISSERTATION

Essays in Knowledge Transfer

by

Leonardo Mayer Kluppel

Doctor of Philosophy in Business Administration

Washington University in St. Louis, 2018

Professor Anne Marie Knott, Chair

This dissertation advances our understanding of when firms are able to get information from outside agents and the impact of that information on innovation. In chapter one, I set up a theoretical model that illuminates how vertical integration changes the incentives for producers to share information with suppliers. Both transaction costs and bargain ability play major roles in information sharing. One result from chapter one is that producers share less information with vertically integrated suppliers. The second chapter tests this prediction, finding robust evidence that once a supplier vertically integrates, it gains fewer spillovers from competitor producers. Finally, the last chapter looks at how information brought by an external CEO can change the firm's innovation direction. The third chapter finds that external CEOs tend to shift the direction of innovation less often that internal hires.

Preface

Knowledge is one of the most critical assets a firm owns. For this reason, it is imperative for strategy scholars to understand how a firm acquires, produces, and uses knowledge. Knowledge is especially vital because firms live in dynamic environments that demand the continual production and use of new knowledge. To create new knowledge, the firm needs to process information (Nonaka 1994) that is generated both inside and outside its boundaries. The information generated inside the firm is shared between individuals and groups within the firm in an environment of shared principles (Kogut and Zander 1992) and language (Monteverde 1995), where individuals share an identity that fosters collaboration (Kogut and Zander 1996). The information sharing within the firm allows for finding the solution to complex problems (Nickerson and Zenger 2004), and is thus vital for firm performance.

The information generated outside the firm does not enjoy the same protective environment. Transfer of information from an outside agent to the firm is harder to achieve. One reason for this difficulty is the fundamental problem with information sharing described by Arrow (1974). The value of information is not known before the information is revealed. However, upon revelation, the buyer could use the information without payment because information is a nonrival good. This creates an incentive for the buyer to renege the contract after being exposed to the information.

The difficulty of getting knowledge from outside agents does not undermine the usefulness of information transfer. Information transfer is essential for technological innovation and product development (Ahuja 2000, Ahuja and Katila 2001, Hall et al. 2010, Lee and Kim 2014, Makri et al. 2010). Thus, firms should be careful when selecting information sources to maximize the possibility that new knowledge will be created. There are many ways to acquire information from outside agents, such as supplying firms that contain useful knowledge (Alcácer and Oxley 2014), buying firms with relevant knowledge (Hitt et al. 1991, 1996), hiring experienced employees (Crossland et al. 2014), physically locating near information sources (Alcácer and Chung 2007, 2014), occupying the same market as more knowleadgeable firms (Bloom et al. 2013, Knott et al. 2009), learning from investors (Dushnitsky and Lenox 2005, Dushnitsky and Shaver 2009, Katila et al. 2008), buying R&D from other firms (Knott 2017), engaging in alliances and joint ventures (Dyer and Hatch 2006, Mowery et al. 1998, 2014, Sampson 2007), getting information from consumers (von Hippel 2001, Jeppesen and Molin 2003), and third party connections (Maggie and Wan 2017, Pahnke et al. 2015, Reagans et al. 2015).

This dissertation studies two ways a firm can acquire knowledge from outside agents. The first two chapters analyze the information a supplier gets from producers. Both chapters look at how vertical integration changes the incentives for producers to share information with suppliers. The third chapter focuses on how external CEOs shape the direction of innovation of the firm.

The first chapter advances a theoretical model aimed at explaining how vertical integration between one producer and one supplier changes the incentives for other producers in the downstream market to share information with the supplier. To achieve that objective, the chapter uses a value capture model (Brandenburger and Stuart 1996, 2007, Gans and Ryall 2017, Lippman and Rumelt 2003, Ryall and MacDonald 2004) to show how producers' decisions to share information with suppliers change when there is vertical integration in the market. The value capture model allows the understanding of how markets' competitive structures and players' bargaining abilities shape the willingness for firms to share information.

In the model, nonintegrated suppliers will not be able to get all the useful information that producers have. One reason for the inability of suppliers to extract all information from producers is the costs to adopt and innovate. Both of these costs are intrinsically hard to contract upon (Hall and Khan 2003, Howells et al. 2008), opening the door to opportunistic behavior and creating transaction costs (Williamson 1979). The existence of transaction costs is one factor for the decrease of information transfer from producer to supplier, but it is not the only one. High consumer bargaining ability exacerbates the problem created by the inability to contract.

Although the transaction costs inherent to adoption and innovation inhibit the transfer of information from one producer to the supplier, when the same costs exist between the supplier and other producers, they help information transfer. Thus, the chapter shows how it is not enough to look at the transaction costs between two firms to determine information transfer. We need to account for the transaction costs between all firms to have a complete picture of the situation.

Finally, the first chapter shows that vertical integration eliminates the transaction costs between the integrated firms, as is common in the transaction costs economics literature (Coase 1937, Williamson 1991). This allows for the unimpeded flow of information between the integrated firms. The downside of integration is that, by decreasing transaction costs between a producer and a supplier, vertical integration decreases the incentives for the other producers to share information with the integrated supplier.

Chapter two is fundamentally interested in testing the theoretical prediction that vertically integrated suppliers get less information from the other producers in the market. This is important since integration trades off more information from one source (the integrated producer) for less diversity of sources (information from other firms). If this is indeed the result of vertical integration, managers have to understand which producers are the critical sources of information for future input innovation. Because different managers will likely have different perceptions about the critical sources of information, this can explain why firms have different vertical structures (Kapoor 2013, Leiblein and Miller 2003).

To test empirically the hypothesis that integrated suppliers receive less information from other producers, chapter two uses the citation lag of suppliers' patents to producers' patents before and after vertical integration. The idea is that citation lag proxies for the amount of information the producers makes available to the supplier. The less information the producer offer to a supplier, the longest it takes for the supplier to find, understand, and use the information contained in the producer's patents. The results from chapter two indicate that citations to other producers' patents made by suppliers' patents increase by about 200 days on average.

Finally, chapter three looks at how CEOs hired from outside the firm shift the direction of innovation in their current firm. Besides having a large impact on firm performance (Mackey 2008), the CEO sits at the top of the organizational chart, a privileged position to spot resource complementarities across all business units and to take care of long-term planning (Chandler 1962, Hambrick 1994). For that reason, the CEO position is uniquely suited for the delineation of the firm's innovation direction.

A firm's innovation path depends heavily on the CEO's experience, knowledge, and beliefs about the world (Hambrick and Mason 1984). Those characteristics can be impacted by a multitude of experiences in the life of the CEO, from living through economic depressions (Malmendier and Nagel 2011) to his/her functional experiences (Crossland et al. 2014). Past employment can also change the way CEOs process and respond to new information. It is known that experiences beyond the focal firm and industry can change the CEOs commitment to the firm's current strategy (Hambrick et al. 1993), and strategic distinctiveness from the other industry competitors (Crossland et al., 2014). Although those results point to differences in how external CEOs behave when making long-term decisions about the firm's overall direction, they do not answer the question of how external CEOs steer the firm's innovation. Chapter three tackles this issue by computing the change in the direction of the vector composed of the technology classification of the firms' patents.

Chapter three finds that, surprisingly, outside CEOs tend to change innovation direction less than insiders. Also, this result is stronger for centralized firms with low R&D intensity, where CEOs should have more control over the direction of innovation. Furthermore, CEOs coming from far away firms in term of technological vectors have a stronger impact on the direction of innovation. One possible explanation for those results is that outside CEOs lack the expertise to change the direction of innovation optimally. Those results are in line with the findings of Cummings and Knott (2018), who show that firms with outside CEOs respond less to CEO incentive compensation.

In summary, this dissertation makes inroads in our understanding of when firms are able to get information from outside agents and the impact of that information on innovation. The first two chapters are dedicated to the impact of vertical integration on information transfer from producers to suppliers, while the last chapter looks at how information brought by an external CEO can change the firm's innovation direction.

Statement of contribution

The first two chapters of this dissertation are solo-authored, even though they benefited immensely from comments and suggestions from many colleagues. The third chapter is co-authored with Trey Cummings. My contributions included formulating the initial question, gathering and cleaning part of the data, performing the statistical analysis, and writing the chapter. Trey was fundamental in helping frame the question, gathering the outside CEO data, thinking about potential statistical pitfalls, and writing the chapter.

Chapter 1: Vertical Integration and Information Sharing

1.1 Introduction

Where should we draw the line delimiting a firm's boundaries? Few questions in strategy have more theoretical and practical relevance. One of the reasons that lead managers to internalize transactions that could otherwise be carried over by markets is to facilitate information transfer between the integrated firms (Demsetz 1988, Grant 1996, Kogut and Zander 1992, 1996, Nickerson and Zenger 2004). Indeed, empirical evidence suggests that integration increases knowledge transfer between units and innovation output of the integrated firm (Ahuja and Katila 2001, Cassiman et al. 2005, Makri et al. 2010). Underexplored in the literature is a study of how vertical integration impacts the integrated firm's capability to access knowledge produced outside the firm's boundaries. This chapter address this shortcoming by using a value capture model (Brandenburger and Stuart 1996, 2007, Gans and Ryall 2017, Lippman and Rumelt 2003, Ryall and MacDonald 2004) to show how producers' decisions to share information with suppliers change when there is vertical integration in the market. The value capture model allows the understanding of how markets' competitive structures and players' bargaining abilities shape the willingness for firms to share information.

The model focuses on situations where producers have information about how to improve the input manufactured by suppliers and shows that vertical integration can change how much information a supplier can extract from producers. This is a fundamental issue, since empirical evidence shows that firms, in general, rely heavily on knowledge spillovers from outside firms

(Hall et al. 2010), and suppliers, in particular, develop more technological capabilities by selling to knowledgeable producers (Alcácer and Oxley 2014). Thus, if managers are interested in improving the input quality over time, they need to fully understand the impact of vertical integration on their ability to access relevant information.

The model shows that nonintegrated suppliers cannot access all the information possessed by producers. One reason for the inability of suppliers to extract all information from all producers is the very characteristics of the costs to adopt new inputs and the costs to innovate. Both of those costs are intrinsically hard to contract upon, opening the door to opportunistic behavior. This is expected from a transactions cost economics (TCE) point of view since the lack of contractibility tends to induce caution when investing. From a pure TCE vantage point, the only outcomes that will be implemented in a disintegrated market are the outcomes that are ex-post beneficial to both producer and supplier.

This chapter shows that transaction costs are not the only force mediating information transfer from producers to suppliers. The bargaining ability of the firms and consumers play a key role in how much information the supplier expects to receive from producers. Different levels of bargaining ability result in more or less information sharing even if we maintain constant the contractibility issues intrinsic to the adoption and innovation costs.

This chapter also argues that it is not enough to consider only the contractibility issues between the supplier and a producer; we need to consider the costs to write contracts between the supplier and other producers. This dependency on the transaction cost between the supplier and the other producers arises because increased competition in the downstream market can enable the supplier to extract more profits from the producers. Thus, after the supplier innovates, it would want for all producers to adopt the new input in order to increase competitive pressure downstream. However, the transaction costs between the supplier and other producers limit the supplier's ability to achieve mass adoption through contracting with all producers in the downstream. Therefore, keeping the transaction costs between the supplier and a producer fixed, the existence of widespread transaction costs helps that producer to send out information.

As an example of the effect described above, suppose that producer A is ex-post marginally better off if it is the only producer to adopt the new input and worst off if it adopts the new input alongside another producer. Suppose that producer B is marginally worse off if it adopts the innovation both alone or with another producer. Finally, suppose that the supplier is better off by having the new input being adopted by one producer, and it is much better off if both producers adopt the innovation. If we only analyze the situation between producer A and the supplier, we could conclude that the information will be shared by A, even with the possibility of holdup caused by the innovation and adoption costs.

However, this conclusion depends on the possibility of the supplier to strike a contract with producer B. If there is no holdup problem between the supplier and producer B, the supplier could write a contract promising to compensate B after it adopts the new input. Knowing that, producer A prefers not to send the information since that would entail innovation by the supplier and adoption by producer B (via contract). However, if producer B and the supplier suffer from the same holdup problems as producer A and the supplier, then the supplier could not write such contract. Without the contract, producer B would not adopt the new input, and thus producer A would be willing to share the information with the supplier. Therefore, A's decision to make the information available to the supplier depends not only on the transaction costs between A and the supplier but also on the transaction cost between B and the supplier.

The difficulties faced by Flextronics (Huckman and Pisano 2010) illustrate how vertical integration can change the incentives of other agents on the market. Flextronics was an OEM cell phone manufacturer that deals with firms that compete in the final product market, such as Motorola, Nokia, Kyocera, and Sony-Ericsson, among others. In 2001, Flextronics decided to integrate cell phone design capabilities. Although Flextronics' in-house design would be just one more competitor in the design market, Flextronics' old customers felt uneasy about sharing information contained in their designs. This threatened Flextronics' existing business relationships, suggesting that there is something different about vertical integration versus simply supplying competitors.

Besides highlighting the importance of competitors' transaction costs in the decision to share information, the chapter helps explain why some industries have firms with different levels of vertical integration (Argyres and Bigelow 2010, Helfat and Campo-Rembado 2016, Kapoor 2013, Kapoor and Adner 2012). Vertical integration gives more access to the integrated producer because it solves the opportunism threat between the integrated parties, but it can come at a cost in information from other producers. In the example above, producer A stops sending information to the supplier after the transaction cost between the supplier and producer B disappears (vertical integration is a possible cause of the decrease in transaction cost between the supplier and producer B). Thus, when considering integration, the manager has to predict which producers contain relevant information to the supplier due to the informational tradeoff described. Those predictions are naturally heterogeneous between managers since individuals have different priors and access different signals. Those different predictions will lead to different prescriptions in terms of vertical integration.

The use of a value capture model also highlights that the decision for players to share information depends crucially on the bargaining ability of the players involved in the game. This is made possible thanks to the use of biform games (Brandenburger and Stuart 2007) as the modeling tool. Biform games allow us to separate the impact of firms' capabilities from the effect of managers' ability to extract value from pure bargaining situations. This can potentially be used to inform a more micro theory of information sharing and vertical integration since we know that the managers' cultural background can influence their ability to extract value in bargaining situations (Henrich 2000, Henrich et al. 2005, Hoffman et al. 1996).

1.2 Theoretical development

1.2.1 Transaction cost between suppliers and producers

This section shows important particularities of innovation and adoption caused by information sent by producers to suppliers. It delineates which intrinsic transaction costs between suppliers and producers are taken into consideration in this chapter and what causes them. Although the process of innovation and adoption is riddled with transaction costs, it does not mean that the only way to solve that problem is through vertical integration. Relational contracts are pointed to in the literature as a remedy that helps alleviate exchange hazards (Dyer and Singh 1998, Elfenbein and Zenger 2013, Greif 1993).

To see how a supplier is impacted by the ability of other producers to write a contract, we need to recognize the particularities of innovation caused by information sharing from producers. When suppliers get a hold of relevant information from producers, they need to invest in R&D to develop and efficiently produce the updated input that crystalizes the information. However, R&D is riddled with information asymmetries, making it hard for producers to write a contract with suppliers to undertake the investment. Evidence of the difficulty in contracting R&D with another firm is the lack of productivity of outsourced R&D (Knott 2017). Further evidence is a survey of research-based pharmaceutical companies operating in the United Kingdom conducted by Howells, Gagliardi, and Malik (2008) showing that the main concern of external innovation sourcing is the "Confidence in the contractor/partner's ability to deliver a relevant solution." This difficulty in contracting naturally imposes constraints on how often suppliers will innovate, and thus how often producers will share information with suppliers.

After suppliers innovate and make the new input available, producers have to invest in the adoption of the new input in their production line. Here again, we have a noncontractible cost,[1] since adoption costs are specific to the producer. As Hall and Kahn (2003) argue:

> The [adoption] costs, especially those of the non-pecuniary "learning" type, are typically incurred at the time of adoption and cannot be recovered. There may be an ongoing fee for using some types of new technology, but typically it is much less than the full initial cost. That is, ex-ante, a potential adopter weighs the fixed costs of adoption against the benefits he expects, but ex-post, these fixed costs are irrelevant because a great part of them have been sunk and cannot be recovered.

The noncontractibility of adoption costs together with the noncontractibility of the innovation cost limits the likelihood that a disintegrated supplier will see full disclosure of information from producers. Although disintegrated suppliers can experience decreases in information sharing depending on the bargaining ability of players, integration can change that. Integration solves the

[1] The noncontractibility of adoption costs is also present as an assumption in economic models, see Antras (2005) for one example.

transaction cost problem because they have special mechanisms to prevent opportunistic behavior, such as authority (Coase 1937, Williamson 1991) and social identity (Kogut and Zander 1992, 1996). The supplier can have access to all information from the integrated producer, independently of bargaining ability. However, by doing so, the integrated supplier decreases the incentives for other producers to share information with it. Without the transaction cost, the integrated firm has the ability to adopt the innovation on the integrated producer's final good, and thus extract more value via the integrated supplier. This gives the other producers less incentive to share information, avoiding the increase in competition caused by the wider adoption.

1.2.2 Vertical Integration

There are three main theories in strategy designed to explain when firms should vertically integrate. One of them is known as transaction cost economics (TCE), since it derives predictions by analyzing each transaction separately to determine if they should be internalized within the firm (Williamson 1975, 1985). Most of the TCE literature relies on firms' inability to design a complete contract in an uncertain environment. Some attributes of transactions, such as the need for firm-specific investments, open the possibility for one of the parts to act opportunistically after a contract is signed. Since it is impossible to prevent opportunistic behavior due to contract incompleteness, firms can be better off internalizing the transaction. Hierarchies solve the transaction problem because they have special mechanisms to prevent opportunistic behavior, such as authority (Coase 1937, Williamson 1991) and social identity (Kogut and Zander 1992, 1996). Although TCE provides a theoretical explanation for vertical integration, it is likely not a complete theory, as noted by Leiblein and Miller (2003). They point out that TCE predicts similar vertical integration decisions for firms facing transactions with comparable characteristics. In reality, however, we observe considerable variance in vertical integration among similar firms (Kapoor 2013, Leiblein and Miller 2003).

The TCE literature has, however, recognized the impact of other firms in the determination of the optimal vertical structure. Due to scale economies, outside suppliers are able to economize in cost if they can aggregate the demand of other firms in the market (Riordan and Williamson 1985). However, this is only true if the input sold by the supplier is not specific to a particular firm. Thus, production costs are not independent of asset specificity. Empirical evidence supports this theory, showing that firms that had integrated transactions after the emergence of a dominant design (so that inputs have lower firm specificity) tend to exit the market (Argyres and Bigelow 2007). Also, increase modularization of technologies is associated with vertical specialization (Argyres and Bigelow 2010). In this literature, the connection between vertical integration and other firms in the market is through economies of scale. The model in this chapter features no economies of scale. Furthermore, this chapter focuses on the impact of vertical integration on the supplier's ability to access informational spillovers. The critical object here is the producer's decision to share information given the firm's capabilities, bargaining abilities, and vertical structure.

One way to explain the heterogeneity in the vertical structure between similar firms is through the resource-based view (RBV). RBV poses that firms are heterogeneous regarding resources and capabilities (Wernerfelt 1984) and that some of those resources and capabilities are hard to imitate or acquire (Barney 1991, Rivkin 2000). In this view, firms integrate to take advantage of unique characteristics such as excess capacity in resources (Penrose 1959) or a unique bundle of previous governance decisions (Argyres and Liebeskind 1999). While RBV is well suited to explain firm heterogeneity in vertical structure, it is not particularly well equipped to explain changes in vertical integration over time within an industry. One example of a change in firms' structure during industry evolution is the evidence that nascent industries tend to be more vertically integrated than more mature industries (Helfat and Campo-Rembado 2016, Klepper 1997).

Besides TCE and RBV, the knowledge-based view (KBV) is another paradigm meant to explain firm boundaries. Although the KBV is less developed than the other two points of view, it has made headway in explaining the vertical structure of a firm. The KBV focuses on the integration, generation and flow of knowledge in a firm (Grant 1996, Kogut and Zander 1992, 1996, Zenger et al. 2011). KBV highlights the fact that information is a key component in drawing the boundaries of the firm. Some theories in the KBV tradition focus on how the problems are chosen to be solved by managers (Nickerson and Zenger 2004). Problems that can be decomposed into modular parts are better solved by the market, while complex problems are more suited to be solved by a hierarchy. Since managers have different worldviews, it is natural that they choose to solve different problems, and therefore create firm heterogeneity. Another explanation rooted in the KBV tradition is the role of integrative capabilities (Helfat and Campo-Rembado 2016). From this perspective, integration is costly, but integrated firms are better at systematic innovation. The implication is that firms should integrate depending on how much systematic innovation they expect to happen. The decision to integrate also depends on managers' expectations about the future and thus can explain firm heterogeneity. Helfat and Campo-Rembado can also explain dynamics if we imagine that managers' beliefs within an industry are somewhat correlated since they get to see similar information. KBV theories, however, do not usually take into consideration the impact of competitors on information exchange between vertically linked units.

Although those theories cover a broad spectrum, they fail in understanding how information transfer from producers to suppliers will be impacted by vertical integration. The producer's decision to share information with a supplier depends not only on the contracting hazards between that producer and the supplier but also on the contracting hazards between the supplier and all producers. The impact of competing firms' contracting hazards on information sharing is not caused only by economies of scale, but also by the supplier's desire to encourage widespread adoption of a new input through contracting. Furthermore, the theories of vertical integration do not study how the agent's bargaining ability influence information sharing.

Next, the chapter introduces the model assumptions. After that, the model is solved for markets with no vertical integration and then for markets with vertical integration.

1.3 Model

The model presented here helps us understand how vertical integration changes the incentives for producers to share information with suppliers. The model uses cooperative games (Brandenburger and Stuart 1996, 2007, Gans and Ryall 2017) to look at how much value each firm expects to gain when producers share information with suppliers. Differences in vertical integration change how much information all producers want to relay to suppliers.

For each market, there are five players: two suppliers (S1 and S2), two producers (P1 and P2), and one consumer (Co). Producers combine the input bought from one of the suppliers with one input produced internally to create the final product. The product created by the combination of input from firm Si with the input from firm Pj is referred to as the final product $FP_{i,j}$. Since this model focuses on information exchange and not on efficiency gains, it assumes that both suppliers and producers have no production cost so that the total value creation depends only on the consumer valuation of the final product. To generate positive value, a group of players needs to contain at least the consumer and two firms: one producer and one supplier.

The choice of using only one consumer makes this model more suitable to analysis characterized by a winner-takes-all[2] market. Examples of instances are VHS, Blu-ray, and Microsoft Office, among others. The assumption of winner-takes-all also has seen implementation in the economics and politics literature (Frank and Cook 1996, Kamien and Schwartz 1974, Sah and Stiglitz 1987).

Before firms start to manufacture the inputs, producers P1 and P2 obtain some information that can be used to improve the input provided by suppliers S1 and S2. That information arises from a better understanding of the final consumer's needs or from an insight gained while assembling inputs into the final product. The existence of useful information from the producer to the supplier is consistent with the fact that some producers expand their information set to intersect with knowledge areas belonging to the supplier (Brusoni et al. 2001, Kapoor and Adner 2012, Takeishi 2002). Besides the evidence that producers have knowledge in areas pertinent to suppliers, Alcacer and Oxley (2014) show that suppliers learn from supplying to more and better producers, suggesting that suppliers do make use of knowledge from producers. Together, those papers demonstrate that producers usually have information that is useful to suppliers.

[2] Some models of winner-takes-all depend on the existence of network externalities and the network characteristics (Lee et al. 2006, Lee Jeho et al. 2015).

Timeline. There are four stages in this game, as shown in Figure 1.1. In the informational stage, both producers decide whether to share the input enhancing information with the suppliers. Each producer decides to share information with each supplier individually. In the innovation stage, suppliers S1 and S2 observe the information sent and decide if they want to produce a new input using the information. If they decide to use the information, they invest in R&D to manufacture an improved input that incorporates the information sent. If no producer decides to send information, the supplier has no alternative but to offer the old input, without quality enhancements.

The third stage is the technology adoption. At this point, producers are aware of the selection of inputs made available by the suppliers and decide if they want to adopt the new input. If producers decide to buy the high-quality input, they need to pay an adoption cost to be able to use the new input on the production line. If they do not pay the adoption cost, their final products do not get the benefits from using the improved input.

Finally, in the bargaining stage, producers buy the input, manufacture the final product and sell it to the consumer. All players engage in a bargain to decide the allocation of created rents.

Valuation. Since the final product is composed of parts created by the supplier and the producer, consumer valuation will vary depending on which firm sold the input and what producer created the final product.

1.4 Equilibrium

To understand how competition influences the incentive for producers to share information with suppliers, we need to characterize the equilibrium value distribution for each of the five players. There are two forces that determine the equilibrium. One is competition, driven by the capabilities of the firms in the market as well as consumers' characteristics. Competition is a powerful force, putting limits on how much each player is expected to appropriate, but it does not generally pin down the equilibrium. The second force determining the equilibrium is the bargaining ability of each player.

Competition imposes two conditions on the value distribution: feasibility and stability (Gans and Ryall 2017, Ryall and MacDonald 2004). Feasibility requires that the sum of values distributed to players be no higher than the total value created. Stability means that, given the value distribution, there is no subset of players that can do better when they break from the main coalition and bargain by themselves. Those conditions guarantee that no player will be able to deviate from the proposed distribution of values.

More formally[3], define G to be any subset of players from the set $N=\{S1,S2,P1,P2,Co\}$. When the set of players G comes together, they can generate a maximum value of v_G. This value v_G is the fruit of the capabilities and needs of each player in the set considered. To illustrate the feasibility and stability requirements, it is useful to go over an illustration. Suppose that we have a ticket scalper with a ticket. The ticket scalper gives zero value to the ticket, and so the value generated by the ticket scalper alone is zero. Assume that the buyer, ticketless, arrive and has a maximum value of $5 for the ticket. When the buyer is alone, she also produces zero value. However, when the buyer and the ticket scalper are together, they can generate a total value of $5. Therefore, we expect to see the buyer and the scalper get together and generate $5 rather than both stay alone and generate zero each. If we assume that the addition of a player does not decrease the value generated, as is the case in the model in this chapter, it is sensible to assume

[3] This discussion follows (Gans and Ryall 2017) closely.

that all players get together to generate the most value possible, so that the equilibrium should feature v_G, the value produced when all players are in the coalition. This is the first thing we can say about the equilibrium of the bargaining.

We are also interested in knowing something about how much each of the players will appropriate from this encounter. Although both players can generate \$5 when they are together, this does not imply anything about how much each player will appropriate from the \$5. To understand that, we need to introduce a notation for how much each player gets in each of the possible coalitions. We do that by defining a function that assigns the value appropriated π_i for each player i. In the ticket scalper example, we could say that when the players are together, the ticket scalper gets \$3 and the buyer gets \$2. Notice that not all possibilities for the function π_i are credible. If we want to analyze value distributions that are likely to happen in reality, we can safely disregard distributions that are not feasible. Formally, we say that our prediction π_i should satisfy feasibility (14.1). This is the second thing we can say about the equilibrium.

$$\sum_{i \in N} \pi_i \leq v_N \tag{14.1}$$

However, feasibility is not the only feature of predicted value distributions. After all, a distribution of -\$2 for the scalper and \$7 for the buyer adds up to \$5, and thus it is feasible. We do not expect the scalper to pay \$2 and give the ticket to the buyer because she can leave the negotiation and get zero instead. In the scalper-buyer case, we do not expect that any of the players get less than zero since they always have the option to walk away from the deal. This property can be generalized with the addition of more players. Suppose that another scalper arrives in possession of a front row ticket, evaluated at \$10 by the buyer. If this new scalper also gets zero by herself, we require that any equilibrium attributes to all players at least zero. However, we also need to consider more possibilities. If a distribution gives less than $10 for the combined value appropriated by the buyer and the scalper with the front row ticket, they have an incentive to break from the group and negotiate alone. By breaking from the group, they can make at least one person strictly better off. More generally, if π_i is the equilibrium, there should be no subgroup that would do better by breaking from the main group. This new condition is called stability (14.2), since it requires that the equilibrium be stable when we consider all the other possible subgroups. This is the final condition competition will impose in the equilibrium.

$$\sum_{i \in G} \pi_i \geq v_G \text{ for all } G \subseteq N \tag{14.2}$$

The set of all distributions that satisfy both feasibility (14.1) and stability (14.2) is named the core. The core for each player is usually a closed interval, not a point. To make a more precise prediction about all players' payoff, we need to introduce another concept: players' bargaining ability. Each player's bargaining ability indicates how much value the player expects to obtain given the bounds imposed by feasibility and stability. A player with high bargaining ability expects to appropriate nearly all value available to her given the bounds imposed by competition. Brandenburger and Stuart (2007) introduce the confidence index to capture the bargaining ability idea, that is, that players have different expectations about how much value each player can appropriate in a pure bargaining situation.

Formally, the confidence index for player i is a number α_i between 0 and 1. The higher the α_i, the more certain player i is to appropriate the upper bound of the equilibrium interval of appropriable rents. Let $\overline{\pi}_i$ and $\underline{\pi}_i$ be the upper and lower limits of the equilibrium rents in player i's core. Then, the expected rents for player i are calculated by $\overline{\pi}_i \alpha_i + \underline{\pi}_i (1-\alpha_i)$.

After finding the expected value appropriated by each player, the game will be solved by backward induction, calculating players' optimal decisions at the adoption, innovation, and information stages.

Next, this chapter will describe the equilibrium when no firm is vertically integrated. Then it will solve the model for integration between S1 and P1. The final analysis will be concerned with the integration between firms S1 and P2.

The derivation of the generalized model is in the Appendix 1, but it is valuable to go over an example. This example will be used in the rest of the chapter. Suppose that the consumer values the different final products according to Table 1.1.

1.4.1 No vertical integration

Given how much the consumer values each final good, we can calculate the value created when each group of players gets together. Table 1.2 shows the value created by all possible groups of players when S1, S2, P1, and P2 are separate firms.

Bargaining. To find the equilibrium in the bargaining phase, we need to calculate the core given the value created by all possible groups. That is, we need to characterize all allocations of value that satisfy feasibility (14.1) and stability (14.2). From the previous discussion, we know that the equilibrium will feature all players $N = \{S1, S2, P1, P2, Co\}$, and will produce 10. This means that the consumer will buy a final product composed by the input from S1 and the input from P1.

Before proceeding, it is useful to define one more object: added value. Each player's added value is how much the player brings to the table. To calculate player $i's$ added value, we consider the difference in value created by all players minus the value created by all players except player i. Value added is then a measure of how much value player i adds to the game. It can be shown that feasibility and stability together imply that no player can get more than its added value[4] (Brandenburger and Stuart 1996). The added value of each player in the example above is shown in Table 1.3.

Since no player can obtain more than its added value, all equilibrium π_i have to satisfy conditions (14.3).

$$\begin{aligned} &\pi_{S1} \leq 6 \\ &\pi_{P1} \leq 2 \\ &\pi_{Co} \leq 10 \\ &\pi_{S2}, \pi_{P2} \leq 0 \end{aligned} \tag{14.3}$$

Stability requires that we take into consideration all possible coalitions. The first possibility is a collation of one player. The one player coalitions impose the conditions that $\pi_i \geq 0$ for all i, since all players can get zero by themselves. Next, we need to consider all coalitions composed of two players. Since no coalition of two players can generate positive value, this condition imposes that $\pi_i + \pi_j \geq 0$. Next, consider all coalitions of three players. There are 60 coalitions, but we do not need to worry with all of them. Any coalition that does not include a supplier, a producer, and the consumer will generate zero value. The relevant three-player coalitions impose the restrictions depicted in (14.4).

[4] See proof in Appendix.1

$$\begin{aligned}
\pi_{S1}+\pi_{P1}+\pi_{Co} &\geq 10 \text{ (value generated by S1,P1,Co)}\\
\pi_{S1}+\pi_{P2}+\pi_{Co} &\geq 8 \text{ (value generated by S1,P2,Co)}\\
\pi_{S2}+\pi_{P1}+\pi_{Co} &\geq 4 \text{ (value generated by S2,P1,Co)}\\
\pi_{S2}+\pi_{P2}+\pi_{Co} &\geq 2 \text{ (value generated by S2,P2,Co)}
\end{aligned} \tag{14.4}$$

Next, we need to consider all coalitions with four players. The relevant conditions are shown in (14.5).

$$\begin{aligned}
\pi_{S1}+\pi_{P1}+\pi_{P2}+\pi_{Co} &\geq 10 \text{ (value generated by S1,P1,P2,Co)}\\
\pi_{S1}+\pi_{P1}+\pi_{S2}+\pi_{Co} &\geq 10 \text{ (value generated by S1,P1,S2,Co)}\\
\pi_{S1}+\pi_{P2}+\pi_{S2}+\pi_{Co} &\geq 8 \text{ (value generated by S1,P2,S2,Co)}\\
\pi_{S2}+\pi_{P1}+\pi_{P2}+\pi_{Co} &\geq 4 \text{ (value generated by S2,P1,P2,Co)}
\end{aligned} \tag{14.5}$$

Finally, stability requires that the coalition of all players distribute at least 10, that is (14.6).

$$\pi_{S1}+\pi_{P1}+\pi_{S2}+\pi_{P2}+\pi_{Co} \geq 10 \tag{14.6}$$

The first conclusion from (14.6) and feasibility is that $\pi_{S1}+\pi_{P1}+\pi_{S2}+\pi_{P2}+\pi_{Co}=10$. Next, we can deduce that $\pi_{S2},\pi_{P2}=0$ by (14.3) and that $\pi_i \geq 0$ for all i. Substituting $\pi_{S2},\pi_{P2}=0$ on (14.4) we obtain (14.7).

$$\begin{aligned}
\pi_{S1}+\pi_{P1}+\pi_{Co} &\geq 10\\
\pi_{S1}+\pi_{Co} &\geq 8\\
\pi_{P1}+\pi_{Co} &\geq 4\\
\pi_{Co} &\geq 2
\end{aligned} \tag{14.7}$$

Direct substitution of $\pi_{S2},\pi_{P2}=0$ on (14.5) gives us (14.8), which were already covered by (14.7).

$$\begin{aligned} \pi_{S1} + \pi_{P1} + \pi_{Co} &\geq 10 \\ \pi_{S1} + \pi_{Co} &\geq 8 \\ \pi_{P1} + \pi_{Co} &\geq 4 \end{aligned} \tag{14.8}$$

Also, players' added value calculated in Table 3 is the smallest upper limit possible. To see that, consider the distributions shown in (14.9). They all belong to the core, and on each of them, one of the players is achieving their added value.

$$\begin{aligned} &\pi_{Co} = 10, \pi_{S1} = \pi_{P1} = \pi_{S2} = \pi_{P2} = 0 \ (\pi_{Co} = Co's \text{ added value}) \\ &\pi_{Co} = 4, \pi_{S1} = 6, \pi_{P1} = \pi_{S2} = \pi_{P2} = 0 \ (\pi_{S1} = S1's \text{ added value}) \\ &\pi_{Co} = 8, \pi_{P1} = 2, \pi_{S1} = \pi_{S2} = \pi_{P2} = 0 \ (\pi_{P1} = P1's \text{ added value}) \end{aligned} \tag{14.9}$$

Some restrictions are redundant. Reducing the set of restrictions to the essential restrictions gives us (14.10).

$$\begin{aligned} &\pi_{S1} + \pi_{P1} + \pi_{Co} = 10 \\ &8 - \pi_{Co} \leq \pi_{S1} \leq 6 \\ &4 - \pi_{Co} \leq \pi_{P1} \leq 2 \\ &2 \leq \pi_{Co} \leq 10 \\ &\pi_{S2}, \pi_{P2} = 0 \end{aligned} \tag{14.10}$$

All the conditions shown above are driven by the structure of competition on the market, i.e., the value generated by all possible coalitions. Note that feasibility and stability are capable of establishing a point prediction for S2 and P2 only; all other players have a range of possibilities. Competition by itself can determine that the consumer appropriate at least 2 and at most 10, but cannot pin down the exact number.

Given the consumer confidence index, we can determine how much of the rents are expected to be appropriated by the consumer. If α_{Co} is the consumer's confidence index, the expected value appropriated by the consumer is $\pi_{Co} = 10\,\alpha_{Co} + 2(1-\alpha_{Co})$, since the consumer's core is between 2 and 10. If, for example, we suppose that the consumer has confidence index $\alpha_{Co} = 0.2$ her expected payoff will be equal to $\pi_{Co} = 3.6$. The chapter will maintain the value $\alpha_{Co} = 0.2$ until the end of the analysis.

To completely describe the equilibrium, we need to take into consideration the other players' confidence index. One way to do that is by assigning a confidence index for each firm in the same fashion as the confidence index assigned to the consumer, that is, chose an exogenous number between 0 and 1. In this case, the confidence indices would be independent of each other, and there would be no guarantee that the sum of the equilibrium value of all players adds up to a feasible outcome.[5] This is a valid route, but it is not what this chapter does.

This chapter only considers confidence indices for S1 and P1 that are consistent, that is, that leads to a value distribution that is ex-post feasible. If the confidence indices are consistent, then the sum of each player's expected value is equal to the total value generated by all players. The corollary of this requirement is that confidence indices for S1 and P1 are linked – a high confidence index for S1 implies a low confidence index for P1. By setting up the confidence indices in this way, the model assumes that changes in the bargaining ability of firms do not change the amount appropriated by the consumer and thus focuses on the bargaining tensions between firms instead of bargaining tensions between firms and consumers. Formally, let α_{S1} be the confidence index for S1 and α_{P1} be the confidence index for P1. The consistency condition in

[5] For example, this approach allows all players to believe that they would appropriate all of the rents available to them (i.e., confidence index equal to 1). In this case, the consumer, S,1 and P1 would believe that they would get 10, 6 and 2, respectively. This distribution is obviously untenable ex post.

the example means that the sum of $\pi_{S1} + \pi_{P1} + \pi_{Co}$ has to be 10. Adding this restriction results in $\alpha_{P1} = (1 - \alpha_{S1})$, as shown in (14.11).

$$\begin{aligned}
\pi_{Co} + \overbrace{(8-\pi_{Co})(1-\alpha_{S1}) + 6\alpha_{S1}}^{\pi_{S1}} + \overbrace{(4-\pi_{Co})(1-\alpha_{P1}) + 2\alpha_{P1}}^{\pi_{P1}} = 10 \\
\pi_{Co}(\alpha_{S1} + \alpha_{P1}) - \pi_{Co} - 2(\alpha_{S1} + \alpha_{P1}) + 12 = 10 \\
(2 - \pi_{Co})(\alpha_{S1} + \alpha_{P1}) = 2 - \pi_{Co} \\
\therefore \alpha_{P1} = (1 - \alpha_{S1})
\end{aligned} \tag{14.11}$$

The confidence index for firm P1 will be $\alpha_{P1} = 1 - \alpha_{S1}$ as long as $\pi_{Co} \leq 4$. This is because if $\pi_{Co} > 4$, then $4 - \pi_{Co} < 0$ and the core for P1 will go from $4 - \pi_{Co} \leq \pi_{P1} \leq 2$ to $0 \leq \pi_{P1} \leq 2$. For the derivation of the full model, please see the Appendix. Substituting the confidence indices in our example gives us the equilibrium payoffs shown in (14.12).

$$\begin{aligned}
\pi_{S1} &= 6\alpha_{S1} + 4.4(1 - \alpha_{S1}) \\
\pi_{P1} &= 2(1 - \alpha_{S1}) + 0.4\alpha_{S1} \\
\pi_{Co} &= 3.6 \\
\pi_{S2}, \pi_{P2} &= 0
\end{aligned} \tag{14.12}$$

The zero value captured by firms S2 and P2 does not necessarily mean that those firms are actually producing in the market. Those firms can be out of the market, and the values associated with their presence is a mere reflection of the possible value they would generate had they decided to enter the market. The value table needs to consider all the credible entrants even if in equilibrium they decide not to enter the market.

Although the only product $FP_{1,1}$ is produced, the amount of value appropriated by firms does not depend only on it. Firm S1 is better off with a higher $FP_{1,2}$, while firm P1 prefers a higher $FP_{2,1}$.

This makes sense, since increments of $FP_{1,2}$ ($FP_{2,1}$) strengthen the bargaining position for firm S1 (P1).

Who wants innovation? Given the equilibrium above, we can ask which products each player wants to improve. For now, suppose that the product created by S1 and P1 ($FP_{1,1}$) increases its value from 10 to 11. Let $\hat{\pi}_i$ represent the new equilibrium value for player i. In this case, the new equilibrium distribution[6] of values is shown on (14.13).

$$\begin{aligned}
&\hat{\pi}_{S1} = 7\alpha_{S1} + 4.2(1-\alpha_{S1})\\
&\hat{\pi}_{P1} = 3(1-\alpha_{S1}) + 0.2\alpha_{S1}\\
&\hat{\pi}_{Co} = 3.8\\
&\hat{\pi}_{S2}, \hat{\pi}_{P2} = 0
\end{aligned} \tag{14.13}$$

Note that the new equilibrium is always better for the consumer, but this is not always the case for S1 and P1. Supplier S1 prefers the new equilibrium only when $\alpha_{S1} > {}^{1}/_{6}$, as demonstrated by (14.14).

$$\begin{aligned}
\hat{\pi}_{S1} &\geq \pi_{S1}\\
7\alpha_{S1} + 4.2(1-\alpha_{S1}) &\geq 6\alpha_{S1} + 4.4(1-\alpha_{S1})\\
\alpha_{S1} &\geq 0.2(1-\alpha_{S1})\\
\alpha_{S1} &\geq \frac{1}{6}
\end{aligned} \tag{14.14}$$

Likewise, producer P1 prefers the higher value $FP_{1,1}$ when $\alpha_{S1} < {}^{5}/_{6}$. If the confidence index α_{S1} is between ${}^{1}/_{6}$ and ${}^{5}/_{6}$, both P1 and S1 will be better off with the higher quality product. If

[6] See the Appendix for the solution with arbitrary values.

the bargaining ability is too skewed towards either the supplier or the producer, one of the firms would have preferred no product enhancement at all.

The cutoffs $^1/_6$ and $^5/_6$ depend on the value of α_{Co}. More generally, we substitute $\pi_{Co} = 10\,\alpha_{Co} + 2(1-\alpha_{Co})$ into the equilibrium values for S1 and P1 to get (14.15) and (14.16).

$$\begin{aligned}
\pi_{S1} &= (8 - [\overbrace{10\alpha_{Co} + 2(1-\alpha_{Co})}^{\pi_{Co}}])(1-\alpha_{S1}) + 6\alpha_{S1} \\
\pi_{S1} &= (6 - 8\alpha_{Co})(1-\alpha_{S1}) + 6\alpha_{S1} \\
\pi_{S1} &= 6 - 8\alpha_{Co} + 8\alpha_{Co}\alpha_{S1}
\end{aligned} \tag{14.15}$$

$$\begin{aligned}
\pi_{P1} &= (4 - [\overbrace{10\alpha_{Co} + 2(1-\alpha_{Co})}^{\pi_{Co}}])\alpha_{S1} + 2(1-\alpha_{S1}) \\
\pi_{P1} &= (2 - 8\alpha_{Co})\alpha_{S1} + 2 - 2\alpha_{S1} \\
\pi_{P1} &= 2 - 8\alpha_{Co}\alpha_{S1}
\end{aligned} \tag{14.16}$$

We can now replace $\hat{\pi}_{Co} = 11\,\alpha_{Co} + 2(1-\alpha_{Co})$ in the equilibrium to derive S1 and P1's equilibrium values with a better $FP_{1,1}$. This substitution yields $\hat{\pi}_{S1} = 6 - 9\alpha_{Co} + \alpha_{S1} + 9\alpha_{Co}\alpha_{S1}$ and $\hat{\pi}_{P1} = 3 + \alpha_{S1} - 9\alpha_{Co}\alpha_{S1}$. Then, S1 prefers the improved product if condition (14.17) is met.

$$\begin{gathered}
\hat{\pi}_{S1} > \pi_{S1} \\
6 - 9\alpha_{Co} + \alpha_{S1} + 9\alpha_{Co}\alpha_{S1} > 6 - 8\alpha_{Co} + 8\alpha_{Co}\alpha_{S1} \\
-\alpha_{Co} + \alpha_{S1} + \alpha_{Co}\alpha_{S1} > 0 \\
\alpha_{S1} > \frac{\alpha_{Co}}{1+\alpha_{Co}}
\end{gathered} \tag{14.17}$$

Figure 1.2 shows all the combinations of α_{Co} and α_{S1} where (14.17) is true, that is, all the points where S1 is better off with higher quality $FP_{1,1}$. The figure is cropped because we know that this equilibrium is valid only if $\pi_{Co} \leq 4$, as discussed previously.

We can calculate when P1 is better off with an enhanced $FP_{1,1}$ in the same fashion. By comparing $\hat{\pi}_{P1}$ with π_{P1}, we can see that P1 is better off with higher quality $FP_{1,1}$ when $\alpha_{S1} < 1/_{1+\alpha_{Co}}$. Figure 1.3 represents this result in the confidence index space.

Finally, Figure 1.4 combines Figure 1.2 and Figure 1.3. The important issue for the discussion here is that as the consumer appropriates more (that is, as α_{Co} increases), the interval of α_{S1} that sustains information transfer decreases.

It is important to notice that the efficient outcome is to have a producer manufacture the best final products possible. Therefore, efficiency will be achieved only when both P1 and P2 send all useful information[7] they have to improve the input provided by S1.

Adoption. Now that we know how innovation on inputs influences value appropriation by firms, we can ask in what condition would producers be willing to adopt the new input, if available. To use a new input to manufacture the final product, the producer has to pay a cost to adapt the production line to process the new input.

The cost incurred in the adaptation process has two special properties. First, the producer has no experience with the improved input, since the input is new. The literature in learn by doing shows that the cost curve tends to fall over time after the introduction of a new product (e.g.,

[7] Supposing that input quality is increasing giving that both P1 and P2 send their information.

Levitt et al. 2013), suggesting that there are inefficiencies and costs associated with introducing new products in the production line. This lack of experience with the new input means that it is very hard to estimate how much it costs to fix all issues on the production line once a new input is introduced.

Second, adoption cost is effectively sunk and nonredeployable (Hall and Khan 2003), raising contracting problems caused by the potential of holdup (Williamson 1979). Once a firm invest resources to learn how to process the new input optimally, it is hard to sell this capability to other firms. Some of the knowledge required to adjust the production line to the new inputs will be tacit (Conner and Prahalad 1996), while some of the know-how required to efficiently process the new input is specific to the bundle of capabilities and routines of a particular firm.

Even if the prerequisite knowledge for a successful adoption were not tacit and completely modular between firms, there is still a last barrier for a firm to sell this knowledge to other firms. Information is known for having an intrinsic contractability problem (Arrow 1962), since the value of information is not known before the knowledge is revealed. However, upon revelation, the buyer could use the information because it is a nonrival good. This creates an incentive for the buyer to renege the contract after being exposed to the information. For those reasons, this chapter assumes that the supplier cannot write a contract to pay the adoption cost to the producer. Firms trying to write such contract would find it hard to agree on the size of the adoption cost and, even if the contract existed, there would be incentives to renege the contract when the firms reach the bargaining stage.

Given the impossibility to write a credible contract, it is easy to see that P2 would not adopt any new input[8] since it would get zero payoffs in the bargaining phase. P1, on the other hand, adopts the new input if the increase in π_{P1} caused by the innovation is higher than the cost of adoption. To simplify the model, let the adoption cost be infinitesimal but positive. If the adoption cost is higher, it makes adoption harder still. Given the noncontractibility of the adoption costs, P1 wants to adopt the new input only when it increases P1's expected value capture in the bargaining stage. In the current example, this occurs when $\alpha_{S1} < 5/6$, but as we have seen before, this cutoff depends on consumer's bargaining ability as well.

This analysis tells us that the structure imposed by competition is not enough to understand the adoption decision by producers. We need to understand the bargaining capabilities of consumers and firms to know when disintegrated producers would be willing to adopt input innovations.

Innovation. The next step in solving the model is to analyze the innovation stage. On this stage, suppliers decide if they should improve the input they sell to producers after observing the information sent by the producers. To improve the input, suppliers need to observe the information and invest in R&D to create the new input. The innovation process is very uncertain (Girotra et al. 2004), and the implementation of the products of R&D into real solutions is fraught with the same contractibility problems as before. This is because the optimal implementation of the result of an R&D project usually depends on tacit knowledge acquired during the innovation phase. Evidence for the fact that R&D is hard to contract is the lack of productivity of outsourced R&D (Knott 2017). For this reason, this chapter assumes that

[8] As long as the new input does not make the product produced by P2 be the highest valued final product. In this respect, this paper focuses on inputs innovations that maintain the order between the final products produced by firms P1 and P2. Incremental innovation is more likely to have this property.

innovation is costly and noncontractible. For simplicity, the cost is assumed to be positive but infinitesimal.

With the innovation cost in mind, we can assert that S1 in our example will innovate only if $1/6 < \alpha_{S1} < 5/6$. If $\alpha_{S1} < 1/6$, S1 is better off by selling the old input. If $\alpha_{S1} < 5/6$, the supplier knows that P1 does not adopt the product, and thus S1 would not want to bear the innovation cost.

Information sharing. In the first stage of the game, producers decide when to share information with suppliers. Since sharing information is usually not costless, the model assumes that producers incur in an infinitesimal but positive cost to relay their information to a supplier.

In equilibrium, producer P1 shares information with supplier S1 if P1 benefits from it ($\alpha_{S1} < 5/6$) and if the information is going to be used by the supplier ($\alpha_{S1} > 1/6$). If the confidence index of supplier S1 is too high ($\alpha_{S1} > 5/6$), P1 will not adopt the innovation even after the new input exists. The lack of adoption is caused by the noncontractibility of the adoption cost, as discussed previously.

If the confidence index of supplier S1 is too low ($\alpha_{S1} < 1/6$, in the example above), then S1 would not want to innovate even if the information were provided for free. In this case, the problem is the noncontractibility of the innovation cost.

Producer P2 gets zero value in all possible equilibria, and therefore does not have an incentive to share information with the suppliers. Finally, no producer sends information to supplier S2, since it does not have an incentive to act upon the information.

Equilibrium. As discussed above, the equilibrium in the example features information sharing from P1 to S1 and technology adoption if $1/6 < \alpha_{S1} < 5/6$. Table 1.4 summarizes the equilibrium for the example and highlights what issues keep the firms from sharing, innovating and adopting.

From a welfare perspective, the efficient outcome would always be for both producers to share information, the supplier S1 to innovate, and P1 to adopt the improved input. This outcome would maximize the spillovers from the information contained in the downstream.

One of the driving forces of P1's decision not to share information when $\alpha_{S1} < 1/6$ or $\alpha_{S1} > 5/6$ is the noncontractibility of either the adoption or innovation costs between S1 and P1. This is the usual argument in transaction cost economics, and it partially explains why there is no efficiency in this market. However, this is not the only force shaping P1's decision to share information. To see why, suppose for a moment that P2 also adopts the innovation, increasing the value of $FP_{2,1}$ from 8 to 9. Denote the new equilibrium when both P1 and P2 adopt the innovation as $\tilde{\pi}_i$ for all players i. In that case, the equilibrium values for S1 and P1 would be characterized by $\tilde{\pi}_{S1} = 7\alpha_{S1} + 5.2(1-\alpha_{S1})$ and $\tilde{\pi}_{P1} = 2(1-\alpha_{S1}) + 0.2\alpha_{S1}$[9]. Given the example above, supplier S1 would be better off if all firms adopted since $\tilde{\pi}_{S1} > \hat{\pi}_{S1}$, and thus S1 would be willing to write a contract with P2 to cover its adoption cost. However, the same contractibility problem impacts the transaction between S1 and P2 – a transaction cost crucial for P1's decision. Since P1's equilibrium value when there is no innovation is bigger than the equilibrium when both P1 and

[9] Derivations for the general form for π_i are in the Appendix.

P2 adopt[10] ($\pi_{P1} > \hat{\pi}_{P1}$), P1 would be better off not sharing information if P2 were to adopt. So, if for any reason the transaction cost between S1 and P2 disappears, P1's decision to share information will change even though there was no change in the contractibility situation between P1 and S1.

This shows that a producer's decision to share information with a supplier depends not only on the transaction cost between that producer and the supplier but also on the transaction cost between the supplier and the other producers. It is the transaction cost between the supplier and P1's competitors that opens the door to information transfer from P1 to S1.

This example shows that decreasing transaction costs between a supplier and some producers can impact the other producer's decisions to share information. The transaction cost between S1 and P2 help alleviate the competition between P1 and P2. As we will see in the next sections, reducing the transaction costs between S1 and P2 via vertical integration can lead to increased competition for P1, decreasing its incentives to share information.

1.4.2 Vertically integrated firms

Now, suppose that some firms decide to integrate vertically. There are two scenarios to contemplate, one where S1 integrates with P1 and the other where S1 merges with P2. Each scenario will be analyzed separately.

Integration between S1 and P1

Suppose that S1 and P1 are now a single firm, named S1P1. The value table, in this case, is shown in Table 1.5.

[10] π_{P1} is also bigger than a scenario where S1 innovates and only P2 adopts. Also, S1 prefers to have P2 as the solo adopter if P1 decides not to adopt.

To find the core, we need to take the same steps as before. First, we compute the added value for all players, as shown in Table 1.6. Then, we calculate the maximum amount of value each player can appropriate, as shown in (14.18).

$$\begin{gathered}\pi_{S1P1} \leq 8 \\ \pi_{Co} \leq 10 \\ \pi_{S2}, \pi_{P2} \leq 0\end{gathered} \tag{14.18}$$

Now we need to consider all possible combinations of players and impose the stability constraint. The result of those considerations is in (14.19).

$$\begin{gathered}\pi_{S1P1} + \pi_{Co} \geq 10 \\ \pi_{Co} \geq 2 \\ \pi_{S1P1}, \pi_{S2}, \pi_{P2} \geq 0\end{gathered} \tag{14.19}$$

We need to make sure that players' added value is indeed the smallest possible upper bound on the core set. To accomplish that, note that the value distributions in (14.20) are in the core. This means that the core contains players' added value.

$$\begin{gathered}\pi_{Co} = 10, \pi_{S1P1} = \pi_{S2} = \pi_{P2} = 0 \ (\pi_{Co} = Co's \text{ added value}) \\ \pi_{Co} = 2, \pi_{S1P1} = 8, \pi_{S2} = \pi_{P2} = 0 \ (\pi_{S1} = S1P1's \text{ added value})\end{gathered} \tag{14.20}$$

Putting all conditions together, we reach the set of equations displayed in (14.21).

$$\begin{gathered}\pi_{S1P1} + \pi_{Co} = 10 \\ 0 \leq \pi_{S1P1} \leq 8 \\ 2 \leq \pi_{Co} \leq 10 \\ \pi_{S2}, \pi_{P2} = 0\end{gathered} \tag{14.21}$$

Note that $0 \le \pi_{S1P1} \le 8$ is irrelevant, and that α_{Co} pins down the equilibrium. The outcomes in the bargaining phase are $\pi_{Co} = 10\alpha_{Co} + 2(1-\alpha_{Co})$ and $\pi_{S1P1} = 8(1-\alpha_{Co})$. In our example, the payoffs are $\pi_{Co} = 3.6$ and $\pi_{S1P1} = 6.4$. An increase of $FP_{1,1}$ always increases the integrated firm's payoff while increasing $FP_{1,2}$ and $FP_{2,1}$ have no impact. The consumer's payoff is the same as the case without integration, and the sum of payoffs between S1 and P1 in the case without integration is the same as the payoff for S1P1.

With integration S1P1, information always flows internally from P1 to S1 – leading to an increase in $FP_{1,1}$ as long as there is useful information to be transmitted. The increase in information flow is caused by the transaction cost extinction brought by the integration, as it is usual in the transaction cost economics literature. The transaction cost between S1 and P2 does not play a role anymore on P1's decision to send information to S1. From an efficiency standpoint, although P1 would always send information to S1, P2 would not. Integration between the high-quality firms (S1 and P1) does nothing to unlock information from other firms.

Integration between S1 and P2

The value table for when S1 and P2 are integrated is represented in Table 1.7.

Again, the same steps are necessary to find the core. The upper limits on the core set imposed by players' added value are shown in (14.22).

$$\begin{aligned} \pi_{S1P2} &\le 6 \\ \pi_{P1} &\le 2 \\ \pi_{Co} &\le 10 \\ \pi_{S2} &\le 0 \end{aligned} \tag{14.22}$$

Now we need to consider all possible combinations of players and impose the stability constraint. The result of those considerations is in (14.23).

$$\begin{gathered}\pi_{S1P2} + \pi_{P1} + \pi_{Co} \geq 10 \\ \pi_{S1P2} + \pi_{Co} \geq 8 \\ \pi_{P1} + \pi_{Co} \geq 4 \\ \pi_{S2} \geq 0\end{gathered} \tag{14.23}$$

Next, the value distributions in (14.24) show that each player's added value is indeed part of one core solution.

$$\begin{gathered}\pi_{Co} = 10,\ \pi_{S1P2} = \pi_{P1} = \pi_{S2} = 0\ (\pi_{Co} = Co's \text{ added value}) \\ \pi_{Co} = 4,\ \pi_{S1P2} = 6,\ \pi_{P1} = \pi_{S2} = 0\ (\pi_{S1} = S1P2's \text{ added value}) \\ \pi_{Co} = 8,\ \pi_{P1} = 2,\ \pi_{S1P2} = \pi_{S2} = 0\ (\pi_{S1} = P1's \text{ added value})\end{gathered} \tag{14.24}$$

Finally, we can calculate the set of core solutions for each player as shown in (14.25).

$$\begin{gathered}8 - \pi_{Co} \leq \pi_{S1P2} \leq 6 \\ 4 - \pi_{Co} \leq \pi_{P1} \leq 2 \\ 2 \leq \pi_{Co} \leq 10 \\ \pi_{S2} = 0\end{gathered} \tag{14.25}$$

The equilibrium payoffs for the example are defined by (14.26). The general case is shown in the Appendix, as usual.

$$\begin{aligned}\pi_{S1P2} &= 6\alpha_{S1P2} + (8 - \pi_{Co})(1 - \alpha_{S1P2}) \\ \pi_{P1} &= 2(1 - \alpha_{S1P2}) + (4 - \pi_{Co})\alpha_{S1P2} \\ \pi_{Co} &= 10\alpha_{Co} + 2(1 - \alpha_{Co})\end{aligned} \tag{14.26}$$

Using $\alpha_{Co} = 0.2$ as the confidence index for the consumer, the payoffs are shown in (14.27).

$$\begin{aligned}
\pi_{S1P2} &= 6\alpha_{S1P2} + 4.4(1-\alpha_{S1P2}) \\
\pi_{P1} &= 2(1-\alpha_{S1P2}) + 0.4\alpha_{S1P2} \\
\pi_{Co} &= 3.6 \\
\pi_{S2} &= 0
\end{aligned} \tag{14.27}$$

An increase in $FP_{1,1}$ from 10 to 11 caused by the use of an improved input will benefit all firms (but not the consumer), resulting in equations (14.28).

$$\begin{aligned}
\hat{\pi}_{S1P2} &= 7\alpha_{S1P2} + 5(1-\alpha_{S1P2}) \\
\hat{\pi}_{P1} &= 3(1-\alpha_{S1P2}) + \alpha_{S1P2} \\
\hat{\pi}_{Co} &= 3 \\
\hat{\pi}_{S2} &= 0
\end{aligned} \tag{14.28}$$

If the information sent from P1 to S1P2 increases only $FP_{1,1}$ and not $FP_{1,2}$, then P1 would be better off sending the information and adopting the new input, while S1P2 would innovate upon receiving the information. This is because $\hat{\pi}_{S1P2} > \pi_{S1P2}$ and $\hat{\pi}_{P1} > \pi_{P1}$ as shown when comparing (14.27) and (14.28).

The story, however, does not stop there. When S1P2 integrates, the transaction cost between S1 and P2 disappears since they now have the same payoff function. Assume that the information given by P1 also makes it possible for $FP_{1,2}$ to increase its value by x units (going from 8 to $8+x$). By also adopting the new input (made possible by P1's information) in P2's final product, S1P2 increases its payoff from $\hat{\pi}_{S1P2} = 7\alpha_{S1P2} + 5(1-\alpha_{S1P2})$ to $\tilde{\pi}_{S1P2} = 7\alpha_{S1P2} + (5+0.2x)(1-\alpha_{S1P2})$. S1P2's payoff growth when $FP_{1,2}$ increases is independent of the values of x or α_{S1P2}.

So far, we know that, if the innovation exists, both P1 and P2 would adopt it. Now we need to know if P1 would decide to share the information in the first place. Before any adoption, P1 expected to get $\pi_{P1} = 2(1-\alpha_{S1P2}) + 0.4\alpha_{S1P2}$ in equilibrium. If P1 decides to send information, S1P2 will innovate and adopt the innovation in P2's final product. Given that there is innovation, the best for P1 is also to adopt and get a payoff of $\tilde{\pi}_{P1} = (3-x)(1-\alpha_{S1P2}) + (1-0.8x)\alpha_{S1P2}$. Therefore, P1 will release information only if $\pi_{P1} > \tilde{\pi}_{P1}$. Formally, we need condition (14.29) to be satisfied.

$$\alpha_{S1P2} \leq \frac{1-x}{0.2(2-x)} \tag{14.29}$$

While P1 stands to send information less often to the integrated firm S1P2, P2 does the opposite. We have seen that S1P2 is better off by adopting if P1 adopts. If S1P2 adopts alone, it will also benefit, getting a payoff of $6\alpha_{S1P2} + (5+0.2x)(1-\alpha_{S1P2})$. Therefore, S1P2 is willing to adopt the new input even if P1 decides not to. In terms of knowledge sharing, this means that P2 will always send information internally to S1P2. For the first time, the equilibrium allows for efficiency.

In Appendix 1, it is shown that producer P1 would want to send information to the integrated firm S1P2 if (14.30) is satisfied[11]. The parameter γ is the ratio between the change in $FP_{1,2}$ due to adoption and the change in $FP_{1,1}$.

$$\alpha_{S1P2} < \frac{1-\gamma}{\alpha_{Co}(2-\gamma)} \tag{14.30}$$

[11] See equation A.34 and A.35 in the Appendix.

Moreover, the Appendix 1 shows that if consumers view the input quality and producers' quality as substitutes, then producer P1 would refrain from sending information to the integrated firm S1P2.

Having an integrated supplier selling inputs to outsiders is very common, as shown by Atalay et al. (2014). In this same paper, Atalay and colleagues propose that one of the reasons for this behavior is the ability to transfer information. This is exactly what the current model suggests since integration is paramount to unlock information from P2, while opening the possibility that S1 gets information from P1 at the same time.

When to integrate? The model presented here has important lessons for a manager thinking about the informational consequences of vertical mergers. When considering integrating along vertical lines, the first thing a manager should understand is the competitive position of all players in the market and their bargain ability. Second, managers need to estimate how the information is distributed across producers. Is the relevant information concentrated in one producer or scattered across all producers? Integration increases the information flow within a firm by decreasing the costs of operating through the market. This decrease in transaction cost can inhibit other firms from sharing information with the integrated supplier, although this is not necessarily true. If the other producers have information that improves the input just for them, then vertical integration can be a way to extract information from all producers. Table 1.8 summarizes the willingness for producers to share information across different vertical structures for the example we have been using.

1.5 Discussion

This chapter uses a value capture model to show how vertical integration impacts information transmission among market participants. In doing that, it highlights the effect of competition and bargaining ability on the incentives for producers to share information with suppliers. The model demonstrates that analysis only considering transaction and firm characteristics between two firms misses an important determinant of the decision to share information: other firms in the market.

Although transaction costs between a producer and a supplier can decrease the amount of information sent, the same transaction costs between the supplier and the producer's competitors can soften competition. This isolation caused by third-party transaction costs incentivizes the producer to share information with the supplier. When third-party transactions costs are solved through integration, the producer might be less inclined to send information to the supplier, even though the transaction producer and supplier both have the same hazards when analyzed separately.

The model developed here can also help explain industry dynamics. The model shows that a market with no integration will see information sharing only if the bargaining ability of consumers is low and the difference in bargaining power of producers and suppliers is small. That observation is important because it can explain why industries tend to be more integrated in its early days (Klepper 1997). When a technological shock occurs, Afuah and Utterback (1997) argue that the bargaining ability of suppliers is low while the bargaining ability of consumers is high. This combination of bargaining ability distribution decreases the incentives for producers to share information. One way to unlock the information contained on producers is to integrate vertically.

Another contribution of this chapter is to understanding why firms fail to adopt new technologies (Dosi 1982, Henderson and Clark 1990, Utterback and Acee 2005). The choice of organizational structure is shown to influence the source and quantity of information received by the supplier. Failure to integrate can lead to too little information being directed at the supplier. On the other hand, integration can cut the flow of information from other producers and thus lead the integrated firms off course.

Finally, the model shows how internal organization can influence firm capabilities, responding to a call to integrate those two concepts (Argyres et al. 2012). The way a firm organizes changes the flow of information from market participants to the firm. This difference in information flow causes differences in firms' knowledge set and ends up shaping future firm capabilities.

1.6 References

Ahuja G, Katila R (2001) Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. *Strateg. Manag. J.* 22(3):197–220.

Alcácer J, Oxley J (2014) Learning by supplying. *Strateg. Manag. J.* 35(2):204–223.

Afuah AN, Utterback JM (1997) Responding to Structural Industry Changes: A Technological Evolution Perspective. *Ind. Corp. Change* 6(1):183–202.

Antras P (2005) Incomplete contracts and the product cycle. *Am. Econ. Rev.* 95(4):1054–1073.

Argyres N, Bigelow L (2007) Does Transaction Misalignment Matter for Firm Survival at All Stages of the Industry Life Cycle? *Manag. Sci.* 53(8):1332–1344.

Argyres N, Bigelow LS (2010) Innovation, Modularity, and Vertical Deintegration: Evidence from the Early U.S. Auto Industry. Organ. Sci. 21(4):842–853.

Argyres NS, Felin T, Foss N, Zenger T (2012) Organizational Economics of Capability and Heterogeneity. *Organ. Sci.* 23(5):1213–1226.

Argyres NS, Liebeskind JP (1999) Contractual commitments, bargaining power, and governance inseparability: Incorporating history into transaction cost theory. *Acad. Manage. Rev.* 24(1):49–63.

Arrow K (1962) Economic welfare and the allocation of resources for invention. *Rate Dir. Inven. Act. Econ. Soc. Factors*. (Nber), 609–626.

Atalay E, Hortaçsu A, Syverson C (2014) Vertical integration and input flows. *Am. Econ. Rev.* 104(4):1120–1148

Barney J (1991) Firm Resources and Sustained Competitive Advantage. *J. Manag.* 17(1):99–120.

Brandenburger A, Stuart H (2007) Biform Games. *Manag. Sci.* 53(4):537–549.

Brandenburger AM, Stuart HW (1996) Value-based Business Strategy. *J. Econ. Manag. Strategy* 5(1):5–24.

Brusoni S, Prencipe A, Pavitt K (2001) Knowledge Specialization, Organizational Coupling, and the Boundaries of the Firm: Why Do Firms Know More than They Make? *Adm. Sci. Q.* 46(4):597–621.

Cassiman B, Colombo MG, Garrone P, Veugelers R (2005) The impact of M&A on the R&D process. *Res. Policy* 34(2):195–220.

Coase RH (1937) The Nature of the Firm. *Economica* 4(16):386–405.

Conner KR, Prahalad CK (1996) A Resource-Based Theory of the Firm: Knowledge versus Opportunism. *Organ. Sci.* 7(5):477–501.

Demsetz H (1988) The Theory of the Firm Revisited. *J. Law Econ. Organ.* 4(1):141–161.

Dosi G (1982) Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Res. Policy* 11(3):147–162.

Dyer JH, Singh H (1998) The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage. *Acad. Manage. Rev.* 23(4):660–679.

Elfenbein DW, Zenger TR (2013) What Is a Relationship Worth? Repeated Exchange and the Development and Deployment of Relational Capital. *Organ. Sci.* 25(1):222–244.

Frank RH, Cook PJ (1996) *The Winner-Take-All Society: Why the Few at the Top Get So Much More Than the Rest of Us* Reprint edition. (Penguin Books, New York, NY).

Gans J, Ryall MD (2017) Value capture theory: A strategic management review. *Strateg. Manag. J.* 38(1):17–41.

Girotra K, Terwiesch C, Ulrich KT (2004) New drug development at Merck & Co. *Case Study Whart. Sch.*

Grant RM (1996) Toward a knowledge-based theory of the firm. *Strateg. Manag. J.* 17(S2):109–122.

Greif A (1993) Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *Am. Econ. Rev.* 83(3):525–548.

Hall BH, Khan B (2003) Adoption of New Technology. Jones DC, ed. *New Econ. Handb.* (Academic Press).

Hal BH, Mairesse J, Mohnen P (2010) Measuring the Returns to R&D. *Handb. Econ. Innov.* 2:1033–1082.

Hefat CE, Campo-Rembado MA (2016) Integrative Capabilities, Vertical Integration, and Innovation Over Successive Technology Lifecycles. Organ. Sci. 27(2):249–264.

Hederson RM, Clark KB (1990) Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Adm. Sci. Q.* 35(1):9–30.

Henrich J (2000) Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *Am. Econ. Rev.* 90(4):973–979.

Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H, McElreath R, et al. (2005) Models of decision-making and the coevolution of social preferences. *Behav. Brain Sci.* 28(06):838–855.

Hoffman E, McCabe K, Smith VL (1996) Social Distance and Other-Regarding Behavior in Dictator Games. *Am. Econ. Rev.* 86(3):653–660.

Howells J, Gagliardi D, Malik K (2008) The growth and management of R&D outsourcing: evidence from UK pharmaceuticals. *RD Manag.* 38(2):205–219.

Huckman RS, Pisano GP (2010) Flextronics International, Ltd.

Kamien MI, Schwartz NL (1974) Patent Life and R and D Rivalry. *Am. Econ. Rev.* 64(1):183–187.

Kapoor R (2013) Persistence of Integration in the Face of Specialization: How Firms Navigated the Winds of Disintegration and Shaped the Architecture of the Semiconductor Industry. Organ. Sci. 24(4):1195–1213.

Kapoor R, Adner R (2012) What Firms Make vs. What They Know: How Firms' Production and Knowledge Boundaries Affect Competitive Advantage in the Face of Technological Change. Organ. Sci. 23(5):1227–1248.

Klepper S (1997) Industry Life Cycles. *Ind. Corp. Change* 6(1):145–182.

Knott AM (2017) *The Record on R&D Outsourcing* (Social Science Research Network, Rochester, NY).

Kogut B, Zander U (1992) Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organ. Sci.* 3(3):383–397.

Kogut B, Zander U (1996) What Firms Do? Coordination, Identity, and Learning. *Organ. Sci.* 7(5):502–518.

Lee E, Lee Jeho, Lee Jongseok (2006) Reconsideration of the Winner-Take-All Hypothesis: Complex Networks and Local Bias. *Manag. Sci.* 52(12):1838–1848.

Lee Jeho, Song Jaeyong, Yang Jae-Suk (2015) Network structure effects on incumbency advantage. *Strateg. Manag. J.* 37(8):1632–1648.

Leiblein MJ, Miller DJ (2003) An empirical examination of transaction- and firm-level influences on the vertical boundaries of the firm. *Strateg. Manag. J.* 24(9):839–859.

Levitt SD, List JA, Syverson C (2013) Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant. *J. Polit. Econ.* 121(4):643–681.

Lippman SA, Rumelt RP (2003) A Bargaining Perspective on Resource Advantage. *Strateg. Manag. J.* 24(11):1069–1086.

Makri M, Hitt MA, Lane PJ (2010) Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. *Strateg. Manag. J.* 31(6):602–628.

Nickerson JA, Zenger TR (2004) A Knowledge-Based Theory of the Firm: The Problem-Solving Perspective. *Organ. Sci.* 15(6):617–632.

Penrose E (1959) *The theory of the growth of the firm* (Oxford University Press).

Riordan MH, Williamson OE (1985) Asset specificity and economic organization. *Int. J. Ind. Organ.* 3(4):365–378.

Rivkin JW (2000) Imitation of Complex Strategies. *Manag. Sci.* 46(6):824–844.

Ryall MD, MacDonald G (2004) How Do Value Creation and Competition Determine Whether a Firm Appropriates Value? *Manag. Sci.* 50(10):1319–1333.

Sah RK, Stiglitz JE (1987) The Invariance of Market Innovation to the Number of Firms. *RAND J. Econ.* 18(1):98–108.

Takeishi A (2002) Knowledge Partitioning in the Interfirm Division of Labor: The Case of Automotive Product Development. *Organ. Sci.* 13(3):321–338.

Tsui E, Waters R (2009) Asustek to spin off motherboard arm. *Financial Times* (December 11) https://www.ft.com/content/699416d2-e67a-11de-98b1-00144feab49a.

Utterback JM, Acee HJ (2005) Disruptive technologies: an expanded view. *Int. J. Innov. Manag.* 9(01):1–17.

Wernerfelt B (1984) A resource-based view of the firm. *Strateg. Manag. J.* 5(2):171–180.

Williamson OE (1975) Markets and hierarchies. *N. Y.* 2630.

Williamson OE (1979) Transaction-Cost Economics: The Governance of Contractual Relations. *J. Law Econ.* 22(2):233–261.

Williamson OE (1985) The economic institutions of capitalism. *Simon Schuster*.

Williamson OE (1991) Comparative Economic Organization: The Analysis of Discrete Structural Alternatives. *Adm. Sci. Q.* 36(2):269–296.

Zenger TR, Felin T, Bigelow L (2011) Theories of the Firm–Market Boundary. *Acad. Manag. Ann.* 5(1):89–133.

1.7 Figures and Tables

Table 1.1: Product values

Product	Value
$FP_{1,1}$	10
$FP_{1,2}$	8
$FP_{2,1}$	4
$FP_{2,2}$	2

Table 1.2: Value table, no integration

Group	Value	Group	Value
S1, S2, P1, P2, Co	10	S1, P1, Co	10
S1, S2, P1, Co	10	S1, P2, Co	8
S1, S2, P2, Co	8	S2, P1, Co	4
S1, P1, P2, Co	10	S2, P2, Co	2
S2, P1, P2, Co	4	Other Groups	0

Table 1.3: Added values

Player i	(1) All players	(2) Value of group (1)	(3) All players except i	(4) Value of group (3)	(5) i's value added (2) – (4)
Co	S1, S2, P1, P2, Co	10	S1, S2, P1, P2	0	10
S1	S1, S2, P1, P2, Co	10	S2, P1, P2, Co	4	6
S2	S1, S2, P1, P2, Co	10	S1, P1, P2, Co	10	0
P1	S1, S2, P1, P2, Co	10	S1, S2, P2, Co	8	2
P2	S1, S2, P1, P2, Co	10	S1, S2, P1, Co	10	0

Table 1.4: No integration equilibrium summary

		Adoption (P1)	Innovation (S1)	Information sharing (P1)
$\alpha_{S1} \leq 1/6$	Action	Adoption	No innovation	No sharing
	Issue	–	Non-contractible innovation cost	Information sharing cost
$1/6 < \alpha_{S1} < 5/6$	Action	Adoption	Innovation	Sharing
	Issue	–	–	–
$\alpha_{S1} \geq 5/6$	Action	No adoption	No innovation	No sharing
	Issue	Non-contractible adoption cost	Innovation cost	Information sharing cost

Table 1.5: Value table, integrated S1 and P1

Group	Value	Group	Value
S1P1, S2, P2, Co	10	S2, P2, Co	2
S1P1, S2, Co	10	S1P1, Co	10
S1P1, P2, Co	10	Other Groups	0

Table 1.6: Added Values, S1 and P1 integrated

Player i	(1) All players	(2) Value of group (1)	(3) All players except i	(4) Value of group (3)	(5) i's value added (2) – (4)
Co	S1P1, S2, P2, Co	10	S1P1, S2, P2	0	10
S1P1	S1P1, S2, P2, Co	10	S2, P2, Co	2	8
S2	S1P1, S2, P2, Co	10	S1P1, P2, Co	10	0
P2	S1P1, S2, P2, Co	10	S1P1, S2, Co	10	0

Table 1.7: Value table, integrated S1 and P2

Group	Value	Group	Value
S1P2, S2, P1, Co	10	S2, P1, Co	4
S1P2, S2, Co	8	S1P2, Co	8
S1P2, P1, Co	10	Other Groups	0

Table 1.8: Conditions for P1 and P2 to send information to S1

	Not integrated	S1 and P1	S1 and P2
P1	$\frac{1}{6} < \alpha_{S1} < \frac{5}{6}$	Always	$\alpha_{S1P2} \leq \frac{1-x}{0.2(2-x)}$
P2	Never	Never	Always

Figure 1.1: Model timing

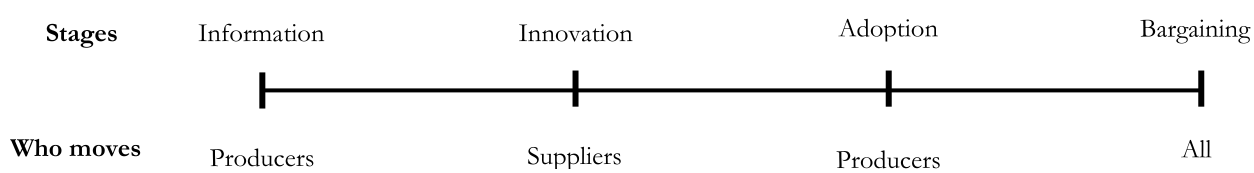


Figure 1.2: Points where S1 is better off with innovation

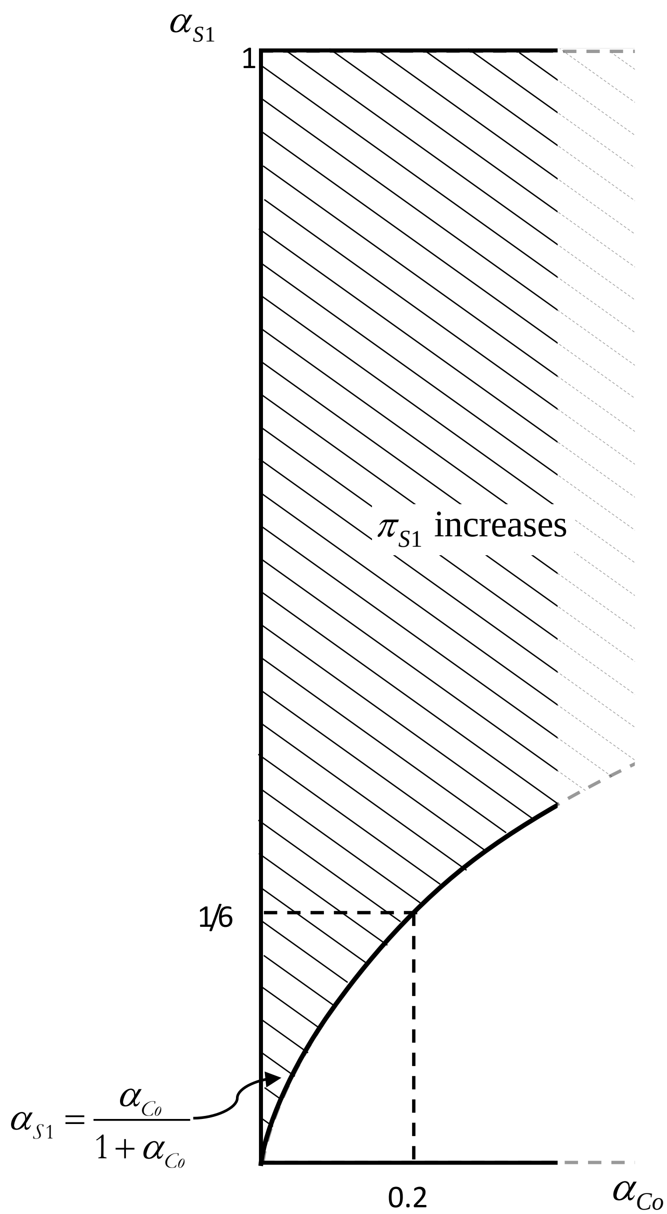


Figure 1.3: Points where P1 is better off with innovation:

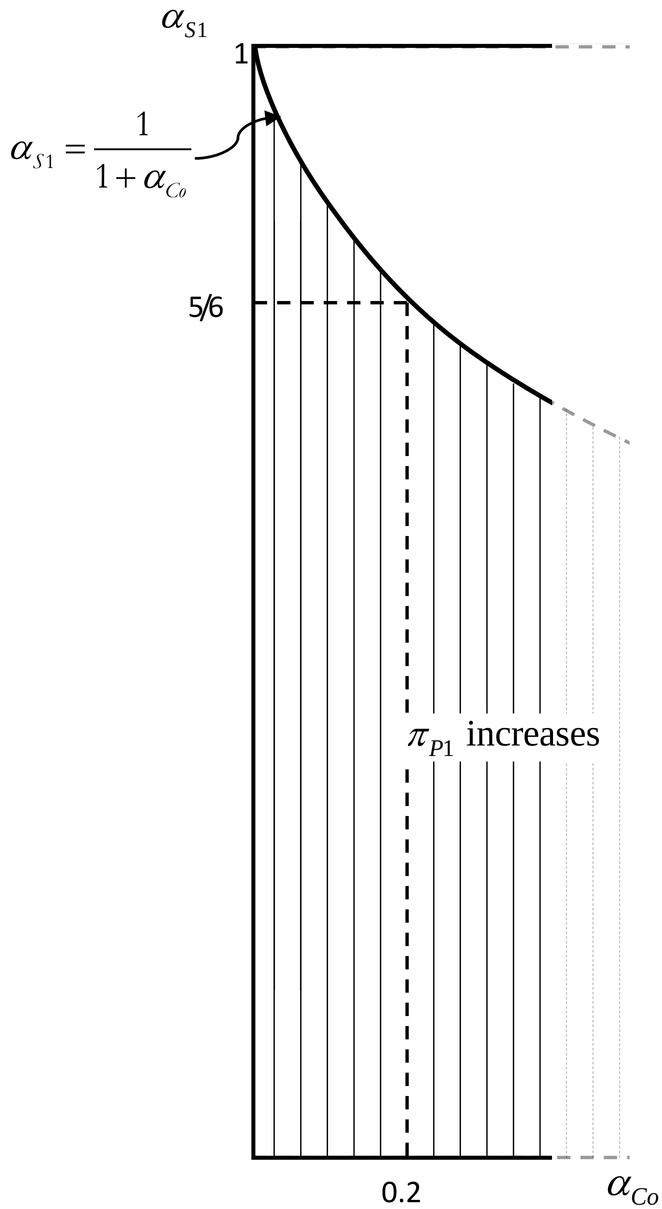
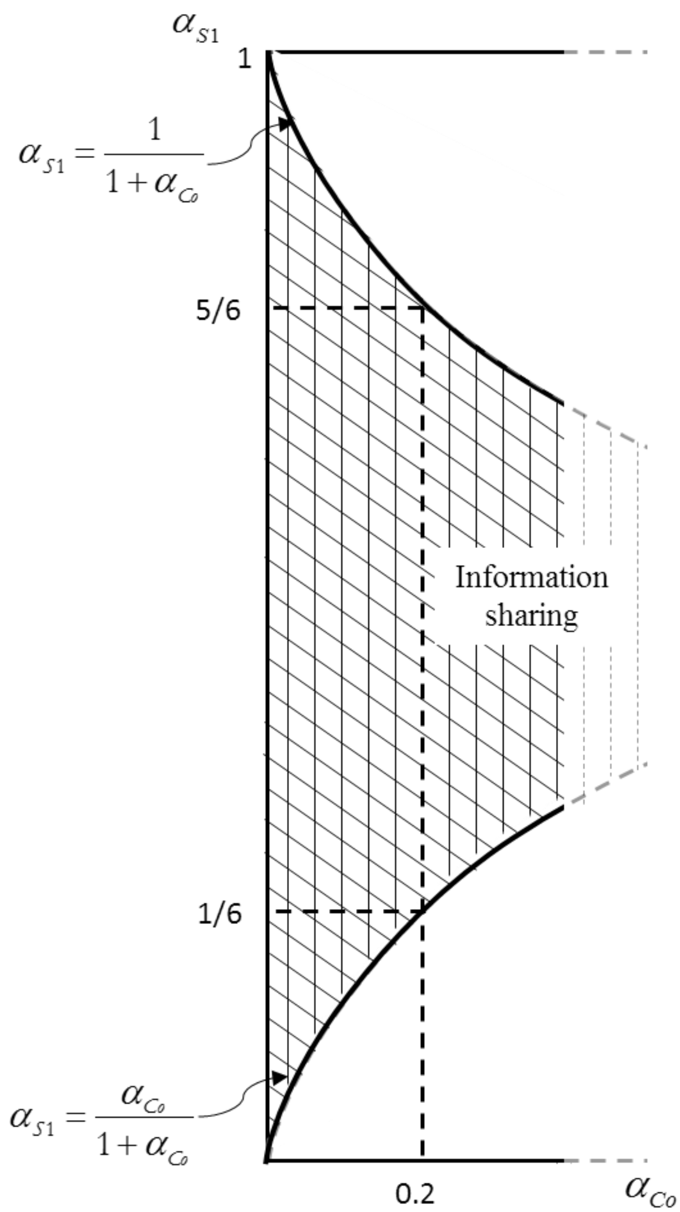


Figure 1.4: Regions with information sharing



Chapter 2: Informational Costs of Integration

Leonardo M. Klüppel
Doctoral Candidate
Washington University in St. Louis
lkluppel@go.wustl.edu

ABSTRACT

To understand how vertical integration impacts knowledge spillovers, we need a theory of the firm that takes into account the impact of integration on all market participants. This paper addresses that problem by proposing a theoretical model showing that markets stimulate sharing of information from all producers. Therefore, markets are a better organizational structure when the manager believes that key information is spread among competitors. Integration, on the other hand, is a better choice when the manager believes that information is concentrated in one producer. Empirical analysis of patent citation data suggests that integrated suppliers tend to receive less information from producers.

Keywords:

Vertical integration; firm boundaries; information sharing; competitive strategy; knowledge spillovers

2.1 Introduction

Knowledge spillovers are the dominant source of value creation from innovation (Hall, Mairesse, and Mohnen, 2010), making it critical for managers to understand what factors impact the firm's ability to get spillovers from other organizations. In line with that effort, the literature has identified many important determinants of the capacity one firm has to benefit from knowledge produced by other organizations. One set of determinants is linked with firms' characteristics such as absorptive capacity (Cohen and Levinthal, 1990), organizational routines (Dyer and Hatch, 2006), the relative amount of knowledge (Knott, Posen, and Wu, 2009), and resources (Katila, Rosenberger, and Eisenhardt, 2008). Other determinants are related to market and network features such as strength of intellectual property protection (Dushnitsky and Shaver, 2009), the composition of the knowledge networks in which the firm is located (Dyer and Singh, 1998; Pahnke et al., 2015), geographical agglomeration (Alcácer and Chung, 2007; Jaffe, Trajtenberg, and Henderson, 1993), and cooperation between firms such as joint ventures and alliances (Mesquita, Anand, and Brush, 2008; Schildt, Keil, and Maula, 2012; Vasudeva and Anand, 2011).

Despite the effort dedicated to understanding spillovers, researchers have yet to understand how vertical integration affects the firm's capability of accessing knowledge produced outside the firm's boundaries. Vertical integration can represent a fundamental change in firms' capabilities. By integrating transactions, firms solve coordination problems (Williamson, 1979), change their resource bundle (Barney, 1991; Wernerfelt, 1984), and impact how information is transmitted within the organization (Grant, 1996; Kogut and Zander, 1992; Nickerson and Zenger, 2004).

For this reason, the decision to share information with a firm can depend on whether that firm is vertically integrated or not.

The difficulties faced by Flextronics (Huckman and Pisano, 2010) illustrate how vertical integration can change the inclination of other agents in the market to share information with the integrated firm. As Huckman and Pisano (2010) describe, Flextronics was an OEM cell phones manufacturer who produced according to designs provided by its customers. Those designs contained customers' private information embedded in them, from future demand expectations to technological innovation. Although Flextronics customers were competitors for final consumers, they shared those designs with a common party. Some of that information could help Flextronics to improve its manufacturing process, an improvement that could be made available to all its customers.

In 2001, Flextronics decided to manufacture phones based on its own designs. Effectively, Flextronics integrated design capabilities into its manufacturing business. Although the experience of designing a phone was successful from a production point of view, the possibility of commercialization revealed a threat to existing business relationships. This case suggests that once Flextronics vertically integrated into the design space, existing customers felt uneasy about sharing information contained in the designs.

There are cases where firms chose to disaggregate in response to a change of customers' willingness to share information, as exemplified by the spinoff of Pegatron Corporation by Asustek Computer in 2010. Pegatron was a subsidiary of Asustek responsible for manufacturing motherboards for many big computer producers such as Dell and HP. In 2007, Asustek entered the portable computer market by launching the Eee PC, a small, lightweight, and low-cost notebook. This launch put Asustek in direct competition with Pegatron customers, making them wary of sharing information with Pegatron. According to a piece in the *Financial Times* (Tsui and Waters, 2009), “Computer makers who use Asustek’s customized motherboards inevitably impart proprietary technical know-how and business intelligence with each order and they fear that would give an unfair advantage to Asustek’s computer business.” To mitigate Pegatron customers’ fears and keep information flows, Asustek decided to spin off the motherboard manufacturer.

To shed light on the impact of vertical integration on spillovers, one needs to account for the impact of vertical integration on all market participants. Most of the theories of the firm[12] do not take into consideration the impact of integration on firms other than the merging organizations and thus present a partial equilibrium view. This paper takes a general equilibrium point of view, presenting a theoretical model that takes into account the impact of vertical integration on all firms in the market.

The model shows that markets with vertically disintegrated firms incentivize all producers to share information with suppliers. The market solution, however, does not usually provide incentives for any producer to share the optimal amount of information with the supplier. Relationships mediated through the market provide suppliers with a varied source of information, at the possible cost of some information hoarding by producers. Vertically integrated producers experience an increase in the incentive to share information within the firm, achieving optimal information flow. However, this increase in information sharing within the integrated firm is followed by the decrease of other producers’ incentives to share information

[12] Zenger et al. (2011) provides a literature review about theories of the firm.

with the integrated supplier. This is the informational cost of vertical integration that managers need to understand when deciding if they should integrate or not.

I then empirically test the model's implications by examining the impact of supplier acquisitions on the information flow from producers. The paper uses the time lag between the application for a patent by the supplier and the application date of cited patents as a proxy for the information flow between firms. More specifically, the paper looks at suppliers that get acquired by a producer and measures the age of citations to the producer's competitors before and after the acquisition. If competitors withhold information from the acquired supplier, we should observe integrated suppliers citing older patents on average. This is because, absent the informational flow, the integrated supplier will be slower to recognize and understand innovations from the downstream competitors. Results are consistent with the theory developed here.

The model presented also helps explain why some industries have vertically integrated firms coexisting with disintegrated firms (Argyres and Bigelow, 2010; Helfat and Campo-Rembado, 2016; Kapoor, 2013; Kapoor and Adner, 2012). Producers' choice to integrate depends critically on beliefs about the distribution of information in the marketplace. Markets stimulate sharing of information from all producers, being a better organizational structure when the manager believes that key information is spread among competitors. Integration, on the other hand, is better suited when the manager believes that information is focused in one producer. Therefore, managers with different beliefs about the distribution of information will make different choices about vertical structure.

Finally, this paper speaks to the link between organizational economics and firm heterogeneity. Argyres and colleagues (2012) argue that there is a need for theoretical work linking organizational economics and the development of firm capabilities. By choosing the vertical structure, the firm controls who will likely provide information to suppliers and helps determine the producer's capabilities. Vertically integrated suppliers get more information from the integrated producer, but less information from other producers. This change in sources of information can shape the future knowledge set of the firm, influencing its upcoming capabilities.

2.2 Theoretical background

The capacity of markets to aggregate information has been recognized since the founding work of Adam Smith. Markets generate incentives for individuals and firms to exchange many different types of information, from demand and supply expectations (Hayek, 1945) to problem solutions (Nickerson and Zenger, 2004; Nickerson, Yen, and Mahoney, 2012). On the other hand, we also know that firm organization can influence how information is transmitted within the firm (Demsetz, 1988; Kogut and Zander, 1996; Nickerson and Zenger, 2004). The incentives induced by the marketplace and by firm organization generate a pattern of information flows, molding the firm's knowledge set (Nonaka, 1994) and, thus, its future capabilities (Conner and Prahalad, 1996).

Those information flows can originate from many agents, both inside and outside the firm. A critically important informational flow for innovation is the information transmitted from producers of final goods to suppliers. Although producers and suppliers work in tandem to produce the same final good, they have different sources of information. Those differences in information come from the fact that both firms are situated in different markets, have different sets of competitors and customers, and hire different professionals. A survey conducted by

Arora, Cohen, and Walsh (2016) documents that almost half of innovators report using external sources (such as suppliers and customers) for their most important innovation. In Europe, the Community Innovation Survey shows that almost four-fifths of the innovative enterprises used information sent by suppliers, and one-fifth of those firms classified this informational source as being highly important.

The model in this paper assumes that producers have some private information that is useful for improving the quality of the input sold by the supplier. That information can arise from a better understanding of the final consumer's needs or from an insight gained while assembling inputs into the final product. The existence of useful information from the producer to the supplier is consistent with the fact that some producers expand their information set to intersect knowledge areas belonging to the supplier (Brusoni, Prencipe, and Pavitt, 2001; Kapoor and Adner, 2012; Takeishi, 2002). Besides the evidence that producers have knowledge in areas pertinent to suppliers, Alcacer and Oxley (2014) show that suppliers learn from supplying to more and better producers, suggesting that suppliers do make use of knowledge from producers. Together, those papers demonstrate that the set of knowledge intersecting producers and suppliers is not empty and that suppliers learn from interacting with producers in equilibrium.

Since information is private, there should be an incentive for the producer to share it with the supplier. On the one hand, sharing information enables the supplier to enhance input quality, which allows producers to create a final product with higher consumer valuation. However, once the information is used to improve the input, the supplier has the incentive to make the high-quality input available to all competitors. The expectation of information leakage to rivals decreases the incentives for producers to share information, a phenomenon studied in the context of supply chain (Anand and Goyal, 2009; Tan, Wong, and Chung, 2015), outsourcing (Baccara, 2007), venture capital (Pahnke, McDonald, Wang, and Hallen, 2015), and partnerships (Katila, Rosenberger, and Eisenhardt, 2008). Although information leakage has been recognized in the literature, changes in producers' incentives to share knowledge once a competitor becomes vertically integrated have not been studied. This topic is of crucial importance for managers deciding whether to integrate or not, especially in industries that depend heavily on knowledge. Even if integration realizes production complementarities, it can be detrimental to the firm's future if it endangers the quantity and quality of information available to the supplier. This can be harmful not only to the integrated firm but to all firms that depend on the information-starved supplier.

This paper is related to the literature that explains why, in some industries, vertically integrated firms coexist with disintegrated firms. The early literature on this topic suggested that firms would tend towards disintegration over time because of returns to scale (Klepper, 1997) and modularization (Langlois, 2003). However, empirical studies that found persistence of both structural forms within industries (Argyres and Bigelow, 2010; Kapoor and Adner, 2012) put limits on the theory of uniform tendency towards disintegration. Helfat and Campo-Rembado (2016) explain this heterogeneity by arguing that, although integration is costly, it helps firms to implement systemic innovations because of its superior integrative capabilities. In equilibrium, some firms choose to bear the cost of a vertically integrated structure to reap the benefits during systemic innovation events. The explanation developed here focuses on when information is shared between firms, allowing for the study of the market's behavior in response to vertical integration.

The model is also connected to the literature on firm boundaries (cf. Zenger, Felin, and Bigelow, 2011), since it outlines the costs and benefits of bringing the relationship between the supplier and producer inside the firm. It shows that managers should be careful about integrating the vertical chain since integrated firms can lose important sources of information. I demonstrate that the use of the market incentivizes firms to share information with suppliers, even when the supplier sells the improved input to all competitors. The market, however, does not give incentives for firms to share information efficiently. By integrating, the producer has incentives to share information efficiently with the supplier. The downside is the diminished incentives for other producers to share information with the integrated supplier.

The following section presents most of the model assumptions. I then solve the model for the nonintegrated case, showing full information revelation. Next, I solve the model for the integrated market structures, showing the effect on information sharing. Finally, I present patent citation data that corroborates the hypothesis that integrated suppliers get less information from market participants. The last section concludes the paper.

2.3 Model

The model shows that a market with a nonintegrated supplier incentivizes all producers to share information. However, no producer shares an efficient amount of information with the supplier, and the higher the expectation about the supplier's skills to appropriate rents, the higher the information distortion. In fact, if producers believe that the supplier is a great bargainer, they will not be incentivized to release information. As for the use of the shared information, the supplier uses all the information received, producing a high-quality input alongside the standard quality input. As for the adoption decision, producers only adopt high-quality inputs in the market where they are leaders.

In summary, incentives for information sharing in the nonintegrated market are homogeneous across firms and dependent on the expectations of the supplier's skills to appropriate rents. Given information sharing, the supplier will produce a varied mix of inputs, but adoption should occur only by producers that are leaders in their respective markets.

When a producer integrates with the supplier, it increases the incentive for that producer to share information. The integrated producer does not need to worry about the supplier's ability to extract rents during the bargaining phase, and thus relays information efficiently. While integration increases information sharing within the integrated firm, the willingness for the other producer to share information decreases. The reason is that the integrated firm has the opportunity to extract more rents from the nonintegrated producer by adopting the high-quality input in all markets, thus increasing competition in the market where the nonintegrated producer is a leader. This widespread adoption improves the bargaining position of the integrated firm[13] in the market where it is not the leader, allowing it to extract more rents from the solo producer. For this reason, the solo producer might be better off not sharing information, negating the integrated firm's chance to use the higher quality input as a way of extracting more rents.

In the integrated case, the incentives to share information are heterogeneous across producers. The integrated producer has high incentives to share information, while the incentives are lowered for the solo producer. Nonintegrated producers have to consider not only the supplier's

[13] More specifically, it increases the bargaining power of the supplier unit within the integrated firm. Before the integration, the non-market leader producer has no incentive to adopt the innovation in that market.

skills to appropriate rents but also the possibility that the supplier might adopt the high-quality input in all markets, increasing competition and decreasing rents for the solo producer.

Next, I make this intuition more accurate by laying down the assumptions of the formal model. Following that, I present the solution when the supplier is nonintegrated followed by the case when the supplier integrates with one of the producers.

2.3.1 Setup

Producers, supplier, and customers. Suppose that there are 2 downstream producers $i = 1, 2$ and one upstream supplier. Since the model focuses on information sharing, I assume that firms have no cost of producing either the input or the final products. The supplier sells an input that has two important attributes, and each of those attributes has associated to it a measure of quality, Q_k. Increases in total input quality, represented by the vector $\vec{Q} = (Q_1, Q_2)$, positively affect customers' willingness to pay. Although input's attribute quality is positively correlated to willingness to pay, the magnitude of the variation depends on customers' and producers' characteristics. Each producer i has a vector of characteristics τ_i that determine customers' willingness to pay for that producer's final product.

There are two markets $j = 1, 2$. Market j has homogeneous customers that display total willingness to pay for the product sold by producer i equal to $W_i^j(\vec{Q}) = W(\vec{Q}, \tau_i, j)$. Willingness to pay increases with input quality and, for a given input quality, customers in market j are better off buying from producer $i = j$. Expressed differently, my assumption is that $W_j^j(\vec{Q}) > W_{i \neq j}^j(\vec{Q})$. Finally, each market j has ρ_j customers in it.

Information. The supplier has the capacity to offer an input of standard quality $\vec{Q}_0 = (Q_1^0, Q_2^0)$. If the supplier gets hold of useful information, it can increase input quality through innovation. The increment in total willingness to pay for changes in attribute qualities is independent, that is, $\partial^2 W_i^j \vec{Q} \big/ \partial Q_1 \partial Q_2 = 0$.

Each producer firm might have information that is useful for increasing the quality of the input. For simplicity, I will assume that producer 1 only has information about input's attribute 1, and producer 2 only has useful information to increase the quality of attribute 2. This information can come from a better understanding of customers' desires, from learning-by-doing during the assembly of the final product or from haphazard insights. The literature lends support to the fact that producers extend their knowledge boundaries past their own products, gaining insight into how to improve inputs. Kapoor and Adner (2012) show that firms have knowledge sets that overlap with outsourced components. Likewise, the literature shows that suppliers do get useful information from buyers (Alcacer and Oxley, 2014; Arora et al., 2016; Hakansson, Havila, and Pedersen, 1999; Kotabe, Martin, and Domoto, 2003; Modi and Mabert, 2007).

Since producers can have useful information, we need to determine in what circumstances they will pass that to the suppliers. More importantly, we need to understand how market structure will influence the choice to share information with the supplier. Once producers choose to disclose or not the information, the supplier decides whether to use the information provided by the producers or not. If the supplier decides to use the information, it can sell either the improved input, the old input, or both simultaneously.

If a producer firm buys the input with higher quality to sell in a specific market, it needs to pay a one-time non-contractible cost k. This is a one-time cost to adapt the manufacturing line that serves a specific market to the new input. Changing inputs usually requires the producer to incur expenses that are hard to predict and hard to disentangle from normal expenditures. For example, new inputs require an optimization in the production line, demanding more attention from workers and managers. Introduction of new inputs might require new machines, changes in marketing strategies, hiring new workers, training, or the design of new contracts. Also, this adoption cost is effectively sunk and non-redeployable, raising the problem of hold-up (Williamson, 1979). For those reasons, this paper assumes that the supplier cannot credibly write a contract to pay the adoption cost to the producer. Parties in this contract would not agree on the size of the adoption cost since a third party would have difficulties enforcing such a contract and, even if the contract existed, there would be incentives to renege on the contract.

Timing. There are four stages in this game, as shown in Figure 2.2.1. In the informational stage, producers observe if they have quality-enhancing information and decide simultaneously to share or not the information with the supplier. This information is held privately, and thus only becomes known to the supplier if the producer decides to share it.

In the innovation stage, the supplier observes the information sent and decides what inputs to make available to the producers. If the supplier receives information from producers, it can choose to offer the standard input, the higher quality input, or both. If no producer decides to send information, the supplier has no alternative but to offer only the standard input.

The third stage is the technology adoption. At this point, producers are aware of the selection of inputs made available by the supplier and decide what input to buy for each of the markets they serve. If producers buy the high-quality input, they pay a non-contractible cost to be able to use the new input in the production line.

Finally, in the bargaining stage, producers decide which input to buy for each market. Producers buy the input, manufacture the product and sell to customers. All individuals engage in a bargain to decide the allocation of rents. It is assumed that customers do not have bargaining power so that all surplus is going to be divided among the two producers and the supplier.

The game is solved using backward induction, following the biform games structure developed by Brandenburger and Stuart (2007). The first three stages of the game (information, innovation, and adoption) are modeled as a traditional noncooperative sequential game. The bargaining stage, on the other hand, uses a cooperative game structure. One advantage of using a cooperative game structure is to make the results less reliant on the specific market institutions. Also, the cooperative game toolset has great instruments for understanding changes in rent allocation caused by changes in the bargaining power produced by the integration of two firms in the marketplace. Finally, using cooperative games to model the bargain between agents after noncooperative games is a technique also used on foundational work on property rights theory (Grossman and Hart, 1986; Hart and Moore, 1990).

The first step in the backward induction argument is to determine the equilibrium in the bargaining subgame. The strategy used to solve the cooperative part is the usual biform games solution. First, one determines the core allocation for each player for a given set of strategies played in the previous stages. Each player’s payoff is then calculated as a linear combination of the highest and lowest possible payoffs in the core. The weight used in the linear combination is an exogenous variable, and it measures how confident the player is in appropriating the surplus

(Brandenburger and Stuart, 2007). This linear combination represents the player's payoff in the bargaining stage for a given strategy profile. Finally, I find the Nash equilibrium using the usual methods for sequential games.

I first solve the equilibrium without integration. Subsequently, I analyze the game when one producer integrates with the supplier.

2.3.2 Solution – No Integration

***Bargaining stage*.** At this point in the game, the final product qualities are fixed at a certain level $\vec{Q}$, inducing willingness to pay W_i^j in market j for producer i's product. Customers in each market buy their favorite product based on the preferences defined above. I assume that producers are able to extract customers' willingness to pay completely, so we can focus on the rent distribution across all firms. This means that if producer i sells any product, then the price will be equal to the customers' willingness to pay.

The cooperative game between producers and the supplier has the form described in Figure 2.2.1. Consider first the market composed exclusively of type 1 customers (market 1). In this market, the value added by producer 1 to the game is equal to $W_1^1 - W_2^1$; the value added by producer 2 is zero; and the value added by the supplier is W_1^1.

For all players and all markets, the core has to be weakly greater than zero, since any firm can leave the market at cost zero. Also, no firm can appropriate more than W_i^i in market i, since this is the total surplus available in that market. In market 1, producer 2 cannot get more than zero income in equilibrium, since the coalition between the supplier and producer 1 can realize more value, splitting among them the portion appropriated by producer 2. Therefore, the core solution for producer 2 in market 1 contains only zero. Likewise, producer 1 cannot appropriate more than $W_1^1 - W_2^1$, since the coalition between the supplier and producer 2 can improve the situation of both. Consequentially, producer 1 has a core set that goes from zero to $W_1^1 - W_2^1$ in market 1. As for the supplier, its payoff cannot be lower than W_2^1; otherwise it would be advantageous to sell to producer 2 instead of selling the input to producer 1. The core solution for the supplier in market 1 is the set of numbers going from a minimum of W_2^1 to a maximum of W_1^1.

Finally, α represents the supplier's confidence index. This is the index that Brandenburger and Stuart (2007) use to calculate a player's payoff in the cooperative part of the game, indicating how skilled a player is in appropriating the rents from her own core set. The higher the α, the more certain the supplier is to appropriate all the rents available in her core set. I assume that the confidence index for the supplier is the same when bargaining with either producer and that the supplier and producers have a consistent confidence index. As a result, the bargain solution in market 1 is the set of payoffs for the supplier, producer 1 and producer 2 such that

$$\{\rho_1(\alpha W_1^1 + (1-\alpha)W_2^1),\ \rho_1(1-\alpha)(W_1^1 - W_2^1),\ 0\}, \alpha \in [0,1]. \tag{1}$$

The first term in this set represents the payoff for the supplier and is the result of the upper limit on the supplier's core (W_1^1) times the confidence index α summed with the lower limit on the core set (W_2^1) times one minus the confidence index $1-\alpha$. This payoff is multiplied by ρ_1, the number of people in market 1. The second term represents the payoff for producer 1. Since the confidence index for the supplier is α, the confidence index for producer 1 needs to be $1-\alpha$.

The upper limit on the core set for producer 1 in market 1 is $W_1^1 - W_2^1$, while the lower limit is zero. Finally, the third term in the (1) is zero because the core set for producer 2 only contains zero.

The situation is analogous for market 2. In that case, the bargaining solutions will be the payoffs represented by:

$$\{\rho_2(\alpha W_2^2 + (1-\alpha)W_1^2),\ 0,\ \rho_2(1-\alpha)(W_2^2 - W_1^2)\}, \alpha \in [0,1]. \tag{2}$$

With the assumptions made so far, producer 1 will have a positive payoff in market 1 and zero payoff in market 2. Conversely, producer 2 will have zero payoff in market 1 and a positive payoff in market 2. The payoffs in (1) and (2) will be valid no matter what quality of input producers decide to buy, as long as the brand preferences are such that customers in market i continue to prefer producer i's product.

It is important to note that those equilibria are valid only when all firms participate in all markets. If a producer j chooses not to participate in the market m, then the bargain solution will be the pair consisting of the supplier and producer i such that $\{\rho_m \alpha W_i^m,\ \rho_m(1-\alpha)W_i^m\}, \alpha \in [0,1]$. If the supplier leaves the market, all firms have zero payoffs.

Adoption of technology. At this point in the game, the supplier has already decided on the menu of inputs it will make available to producers. After learning about their options, producers need to decide on what inputs to buy and in which markets to employ them. If only the standard input $\vec{Q}_0 = (Q_1^0, Q_2^0)$ is offered, all firms will buy it, and the payoffs will be described by (1) and (2).

Suppose instead, that the supplier offers a high-quality input $\vec{Q}_h = (Q_1^h, Q_2^h)$, with attributes $Q_i^h \geq Q_i^0$. In this case, both producers need to decide if it is worthwhile to pay the cost k of adopting the new input. First, note that a producer's payoff is positive in one market and zero in the other, as indicated by the payoffs in (1) and (2). Let ΔW_i^j represent the increases in the willingness to pay of market j's customers for producer i's product when the high-quality input is used instead of the standard input, and let $\underline{W}_i^j = W_i^j(\vec{Q}_0)$ represent the willingness to pay for a final product with standard technology. Also, let $\bar{W}_i^j = \underline{W}_i^j + \Delta W_i^j$ be the total willingness to pay of market j's customers for producer i's product after the innovation. Firm i adopts the innovation if the expected increase in the bargain stage payoff is higher than the cost k of introducing the innovation. In other words, firm i will choose to adopt the innovation in market $j = i$ if condition (2) is satisfied.

$$\Delta W_i^i \geq \frac{k}{\rho_i(1-\alpha)} \tag{2}$$

In market $j \neq i$, firm i will choose to adopt the innovation if condition (2) is satisfied. This condition states that the innovation is beneficial enough that producer i would become a leader in market $j \neq i$, but the innovation is not beneficial enough for the present market leader j. The equilibrium, in this case, is that producer i adopts the innovation in market $j \neq i$, while firm j does not adopt.

$$\bar{W}_i^j > \bar{W}_j^j - \frac{k}{\rho_j(1-\alpha)} \text{ and } \bar{W}_i^j > \frac{k}{\rho_j(1-\alpha)} + \underline{W}_j^j \tag{2}$$

In the rest of the paper, I will assume that condition (2) is not satisfied. This amounts to ruling out innovations that would benefit the market follower so much and the market leader so little that it would be optimal for the follower, but not for the leader, to adopt the innovation in that market. In that sense, this paper deals with innovations that are not radical enough to subvert the present order of customers' willingness to pay.

Given the assumptions, either one producer adopts the innovation or no producer does. The reason is that in markets where a producer displays lower willingness to pay than the competitor, that producer will get zero profits. In this case, the producer would not be willing to pay any positive amount to adopt the innovation in that specific market. Since I assume that condition (2) does not hold, the equilibrium is that firm i adopts the high-quality input in market i if (2) holds true and does not adopt it in market $j \neq i$.

Decision to use knowledge. If the supplier does not get any information from the producers, there is nothing to be done. In this case, the supplier will offer the input with the standard quality $\vec{Q}_0$; all the producers will buy that input for all markets, and the supplier gets a total profit of (2) per market i.

$$\rho_i[\alpha \underline{W}_i^i + (1-\alpha)\underline{W}_j^i] \qquad (2)$$

Suppose instead that the supplier does get information from one of the producers. With that information, the supplier can produce input with a higher quality $\vec{Q}_h > \vec{Q}_0$. This higher quality product will allow firms to sell products to customers with higher willingness to pay.

The supplier needs to decide whether to offer only the standard input, only the improved input, or both. If only the standard input is on the market, the supplier's profit will be equal to (2). If the supplier decides to manufacture both the standard input and the improved input, its profit in market i will be either $\rho_i[\alpha \underline{W}_i^i + (1-\alpha)\underline{W}_j^i]$, if no producer buys the improved input, and $\rho_i[\alpha \overline{W}_i^i + (1-\alpha)\underline{W}_j^i]$ if producer i buys the improved input. Therefore, the supplier's profit will be weakly greater when it decides to offer both inputs.

Also, the higher the input quality, the better off the supplier is. Therefore, the supplier will use all information available to increase quality as much as possible. If the supplier gets information from both producers, it will use all of it to produce the input with the highest possible quality.

Finally, the supplier can decide to offer only the high-quality input. In this case, at the very least one producer will drop from each of the two markets. As explained before, the producer that makes zero profit will prefer to leave the market altogether rather than pay any positive cost k. In this situation, one possibility is that both producers decide to leave the market, leaving the supplier with zero profit in that market. The other possibility is that only one producer leaves market i, giving the supplier profits equal to $\rho_i \alpha \overline{W}_i^i$. In any case, the supplier gets higher profits when it provides both inputs in equilibrium.

One could imagine that the supplier could sell the high-quality input at a discount k to the producer that would drop from the market otherwise. This would allow the supplier to reap profits equal to $\rho_i[\alpha \overline{W}_i^i + (1-\alpha)\overline{W}_j^i] - k$, potentially higher than the profit obtained by selling both inputs. This scheme would not work. The equilibrium would not change even if the supplier promised to give the producer a discount equal to the adoption cost. The reason is the incentives for opportunistic behavior by the supplier (Williamson, 1979, 1985). After the producer incurred the sunk adoption cost, the bargaining phase would still be described as before, and payoffs would be described by (1) and (2). Since producers anticipate this opportunistic behavior, they would not be convinced to adopt in a market where they expected to get zero profit from the bargain.

Decision to share knowledge. At this stage, both producers have to decide to share information or not with the supplier. They know that the supplier will use any information shared and that there will be two types of inputs if any information is shared. Moreover, both producers know that only one of them will adopt the high-quality input in either market. To determine the incentives to share information with the supplier, consider the decision of producer i given the sharing decision made by producer j. Suppose that sharing this information would allow the supplier to increase quality to $Q_i^q > Q_i^0$, inducing a gain of $\Delta W_i^i = W_i^i(Q_i^q, Q_j) - W_i^i(Q_i^0, Q_j)$ for producer i.

First, suppose that the increase in quality permitted by sharing information does not allow for condition (2) to be true; that is, suppose that producer i will not adopt in market $j \neq i$. In this case, firms would gain by sharing information that would produce any innovation such that condition (2) is met. Since each producer i only has information about Q_i and the impact on input quality is independent across attributes, the information for all adopTable 2.innovation is shared to the supplier, leading to proposition 1.

Proposition 1: Under nonintegration, every producer i is willing to share information with the supplier as long as this information leads to an improved input that producer i would adopt.

Before showing the equilibrium with integration, it is important to note what would be the efficient outcome in this model. If we sum the payoffs of all producers and the supplier, we obtain $\rho_1 W_1^1 + \rho_1 W_2^2$. This aggregated payoff shows that the efficient outcome is achieved when information is shared and used if $\Delta W_i^i > \frac{k}{\rho_i}$. The deviations from the efficient amount of information are caused by supplier opportunism, and cannot be solved using a formal contract.

2.3.3 Solution – Integration

Now suppose that producer 1 integrates with the supplier. The game timing remains the same, but now there is an integrated firm playing the roles of both supplier and producer. Producer 2 will buy the input from the integrated supplier. This arrangement of integrated firms selling inputs to the outside market is very common, as shown by Atalay, Hortaçsu, and Syverson (2014)[14].

Bargaining stage. The integration between producer 1 and the supplier means that, in the bargaining stage, the merged firm will appropriate all of the surplus in market 1. In the market composed of type 2 customers, the profits of the integrated firm are equal to the supplier's profits in the nonintegrated case. The profit of the merged firm after the bargaining phase is now described by

$$\rho_1 W_1^1 + \rho_2 [\alpha W_2^2 + (1-\alpha) W_1^2]. \qquad (2)$$

In the bargaining stage, the profit for firm 2 will remain the same as before, that is

$$\rho_2 (1-\alpha)(W_2^2 - W_1^2). \qquad (2)$$

[14] Besides, if the integrated firm decides to stop supplying producer 2, the result would follow immediately. In that case, producer 2 would leave the market and never share information.

Adoption of technology. In market 1, the integrated firm will adopt the innovation if the increase in customers' willingness to pay caused by the high-quality input adoption satisfies condition (2). Note that the constraint for technology adoption is more relaxed than in the nonintegrated case (2). This means that the integrated firm will adopt innovations more often than the nonintegrated case.

$$\Delta W_1^1 > \frac{k}{\rho_1} \qquad (2)$$

The situation is different in market 2, since the integrated firm may have an incentive to adopt the innovation even when condition (2) is not satisfied. The reason is that by increasing the quality of the product sold by producer 1 in market 2, the integrated firm attains a better position in the bargaining stage. The bargaining stage payoff (2) indicates that the integrated producer 1 will adopt the innovation whenever condition (2) is satisfied. Different from the nonintegrated case, market 2 can experience technological innovation by producer 1.

$$\Delta W_1^2 > \frac{k}{\rho_2(1-\alpha)} \qquad (2)$$

Producer 2 will have the same strategy as the nonintegrated case in terms of technology adoption. More specifically, given the adoption strategy of producer 1, producer 2 will adopt the innovation if condition (2) holds.

$$\Delta W_2^2 > \frac{k}{\rho_2(1-\alpha)} \qquad (2)$$

Decision to use knowledge. The profit from the integrated firm (2) is increasing in W_i^j for all i and j. Consequently, all information available will be incorporated into its inputs. The supplier will produce only the highest possible quality input if conditions (2), (2), and (2) hold. If any of those conditions fails, the supplier will produce both types of inputs. This is because if condition (2) or (2) fails, the integrated firm would itself prefer to use the standard input. If condition (2) fails, producer 2 prefers to leave the market, decreasing the total amount of surplus exceeding W_1^2 to be divided.

Decision to share knowledge. Producer 1 decides to share knowledge for the same reasons explained in the nonintegrated case. In particular, producer 1 shares information whenever it would be profiTable 2.to adopt it, that is, under conditions (2) and (2). The nonintegrated producer 2, however, realizes that disseminating its knowledge can generate increased competition in market 2 caused by the adoption of the technology by the integrated producer 1. Any information that leads to an increment in W_1^2 such that producer 1 decides to adopt the resulting innovation will invariably increase competition in market 2.

If producer 2 could, it would give out knowledge with the condition that it could be applied only to the input sold to itself. This contract would be hard to enforce since the integrated supplier has the incentive to produce high-quality inputs for its own producer 1 as well. Contract and threat of legal action are not enough to stop firms from disseminating information about innovation. A survey of R&D managers (Cohen, Nelson, and Walsh, 2000) shows that patents are viewed at best as a flawed mechanism to protect intellectual property.

Therefore, producer 2 needs to be careful when it owns information that would induce an innovation adopTable 2.by the integrated firm in market 2. Producer 2 will share information if it generates an adopTable 2.input (2) and such that the benefit for firm 2 is big enough to overcome the loss from the increased competition caused by producer 1's adoption (2).

$$\Delta W_2^2 > \Delta W_1^2 \tag{2}$$

In contrast, the nonintegrated case only required condition (2) for disclosure. The area where information will not be shared with the supplier because of the vertical integration between the supplier and firm 1 is depicted in Figure 2.2.2. Area BCE in Figure 2.2 illustrates the region where producer 2 would have shared information in the nonintegrated case but does not share information in the integrated case. In contrast, area ACDF corresponds to events where producer 1 chooses to share information because of the integration.

Proposition 2: Under integration, the integrated producer i *is willing to share information with the supplier as long as this information leads to an input that the producer would choose to adopt—in other words, any information that would change customers' willingness to pay by at least* $\frac{k}{\rho_i}$ *in market* i *or* $\frac{k}{\rho_j(1-\alpha)}$ *in market* j.

The nonintegrated producer j *shares information only when this disclosure induces an innovation that changes customers' willingness to pay such that* $\Delta W_j^j > \frac{k}{\rho_j(1-\alpha)}$ *and* $\Delta W_j^j > \Delta W_i^j$.

Proposition 2 says that, under integration, the integrated producer shares information more often than under nonintegration. On the other hand, the nonintegrated producer shares information less often. The total impact on input innovation will depend on the information producers have and how it translates into changes in willingness to pay in each market.

2.4 Empirical evidence

The theory presented here makes two main predictions. First, the model shows that the integrated firm increases information transfer within the firm. Second, other producers in the market share less information with the supplier once it becomes vertically integrated. The first prediction is hard to test empirically, since observing information transfer within firms is hard. However, we can test the second prediction by using patents to proxy for information transfer between firms. With this purpose in mind, I use the time lag between the application date of the citing patents and the application date of the cited patents to measure the speed of information transfer between firms (Oxley and Wada, 2009).

While patents are public rather than private information, the speed of diffusion depends on private information transmitted by the patent holders. First, firms filing patents have incentives to disclose the minimum amount of information required to get the patent granted. Therefore, receiving extra information about the patent from the filing firm helps the citing firm to understand better the patent, a critical step in applying the information contained there to other patents. Second, there are many patents being granted at any point in time. Because of limits in attention, a firm might not be aware of all available patents, and thus information from the patent holder can alert the firm of that information. Finally, the measure of time lag uses application dates and not granted dates. Therefore, a firm that has information from the patent holder might be able to know about a patent that has not been released to the public yet. This prerelease information shortens the period for a firm to understand the patent and put the information contained there to use. For those three reasons, suppliers that receive information directly from producers will be faster at recognizing and using that information for their own innovations, decreasing the average lag of the citation. This leads to the following hypothesis:

Hypothesis: *Once the supplier integrates with a producer, the speed of technology transfer between the supplier and other producers decreases.*

2.4.1 Data

To test the hypothesis, I gathered information about buyer firms and the firms that were targets of M&A operations. The set of buyers and targets comes from the merger and acquisitions data from the S&P Capital IQ database. This dataset shows all mergers and acquisitions from companies and their subsidiaries since 1998[15]. The sample consists of large firms in Information, Capital Goods, and Chemicals[16]. The reason for the choice of industries is that those firms generally use patents to protect their innovation.

Since the dataset does not show actual suppliers, I look at target firms and classify them as belonging or not to an upstream industry of the buyer's industry. To achieve that, I link both buyer and target firms to the LexisNexis database to get each firm's primary and secondary[17] NAICS codes. I then use the Bureau of Economic Analysis Input-Output direct requirements Table 2.from 2007 to indicate if the target firm belongs in an upstream industry from the point of

---

[15] In North America. The database tracks European M&A since 2001, but selected deals are available before that.
[16] It includes the 1,200 biggest firms in Information, 729 biggest firms in Capital Goods, and 256 biggest firms in Chemicals.
[17] Up to five secondary NAICS codes were used.

view of the buyer firm. The direct requirements Table 2.measures the dollar amount of the input from each industry required to produce one dollar's worth of output.

To classify industries as belonging to an upstream industry or not, we need to agree on a threshold that separates upstream industries from non-upstream industries. Following the literature (Atalay et al., 2014), I considered all industries that supply at least 1 percent to the final product to be in the producer's upstream. Results with different thresholds are also reported.

The product of those three data sets identifies which targets are in the upstream of the buying companies and what the date of the integration is. To identify the buyer's competitors, I used the "quick competitors" list from S&P, which shows up to 10 pre-selected firms comparable to the buyer firm. Finally, in order to measure information flows, I aggregated information from the USPTO patent database.

All database linkages involved fuzzy matching of firms' names, with manual verification of the possible matches. I excluded design patents and patents that were applied for before or after 10 years of the acquisition date. The reason for the exclusion of data that is too distant from the acquisition is that firms can change their product mix with time, creating noise. In any case, regressions' results are robust to the inclusion of citations beyond the 10-year period. This procedure yielded 287 firms that acquired at least one of 572 targets, from 1998 to 2017. In total, those targets applied for 32,935 patents in the period from 1989 to 2016. Those patents cite 185,016 different patents. From those patents, 5,912 patents were issued by one of the buyer's competitors.

2.4.2 Method

The unit of observation is a citation i at date t by a target firm f that gets eventually acquired by buyer b. The equation for the basic model is:

$$TimeLag_{itfb} = \beta_t Acquired_t + \beta_i Cite\ Competitor_i + \beta_{it} Acquired_t * Cite\ Competitor_i$$
$$+Controls_{itfb} + \lambda_t + \gamma_f + \delta_b + \varepsilon_{itfb}$$

Dependent variable. The dependent variable is the time lag between the application date of the patent by the target of the M&A and the application date of the cited patents. If firms receive information from the owners of the cited patents, we should expect that it would take less time for the citing firm to identify and understand that information. This should lead target firms to use more recent information in their patents, decreasing the lag between the patent's application date and the cited patents' application date.

Independent variables. The dummy *Acquired* gets value equal to one if the citations are made by a patent whose application date is later that the announcement of the M&A. Note that I'm using the date of the M&A announcement as the date of acquisition. While not technically correct, the announcement date is the relevant date because, from that date on, all the competitors become aware of the likely integration and thus can act as if the target were already acquired. It is possible, however, that competitors know about the impending acquisition before the announcement date, since they all live in the same market, and probably share many connections through employees. This fact makes the determination of a hard "treatment" date impossible. Compounding the fuzziness problem, there is a nontrivial delay between the discovery of an innovation and its application.

When a firm applies for a patent, it needs to disclose all prior knowledge that is relevant to that patent application. Failure to do so can cause the patent to be deemed invalid, and so there is a strong incentive for firms to cite all the known previous knowledge accurately.

The variable *Cite Competitor* is a dummy that specifies when one of these citations refers to a patent originally filed by a competitor of the buying firm. Since the hypothesis investigates information transfers from the buyer's competitors to the acquired target, this paper focuses on those citations. In other words, the treatment variable is *Acquired* while the treated group is identified by the dummy *Cite Competitor*.

Control variables. One of the confounding factors is the change in the patenting and citing behavior of the target firm. For this reason, the regression needs to control for characteristics of citing and cited patents. For citation characteristics, *Num Claims* counts the number of claims of every cited patent, a measure of patent broadness. *Num Cit* is the count of citations that refer to that specific patent, a measure of the cited patent's innovativeness. *Mean Lag* is the average lag over all patents that also cite that specific patent, and it is meant to capture the speed that such citation usually diffuses. For the citing patents, the estimations control for *Num Claims*, the number of claims of the citing patent, and *Num Cit made* is the count of citations made by the patent. Finally, *daysafterbuy* and *SQRdaysafterbuy* are the difference between the announcement date of the M&A and the application date for the patent, and its square, respectively.

2.5 Results

The summary statistics are presented in Table 2.1. The mean time lag is over 10 years, showing that patents heavily cite old knowledge. About 37 percent of the citations in the data are made after the targets are acquired, and 5 percent of the citations refer to a buyer's competitor.

The raw data showing the time lag for competitors' citations is plotted using a lowess smoothing in Figure 2.3. The vertical axis represents the time lag for citations made to competitors' patents, while the horizontal axis shows the days after the announcement of the M&A. The red vertical line marks one year after the M&A announcement date. From this graph, we can see that the citations' lag difference between upstream firms versus non-upstream firms up to one year after the M&A announcement is small. After that, however, the time lag for upstream firms increases much faster than that for non-upstream firms. This pattern suggests that the hypothesis that competitors retain information from integrated upstream firms is true. The fact that the divergence in time lag occurs only after one year from the announcement date is not surprising since the impact of a lack of information should take time to be captured by patent applications.

The same conclusions can be drawn by Figure 2.4, which shows linear predictions and their 95 percent confidence intervals for averaged data before and after the announcement of the M&A. Each dark (light) point on this graph is built by dividing citations into 120-day buckets and averaging out the time lags of all upstream (non-upstream) firms.

Both pictures tell the same history: firms that are in the upstream of their buyers see the time lag of citations increase after the announcement of the M&A, even when we compare them to acquired firms that are not in the upstream. However, neither of these pictures controls for confounding factors such as firm-specific unobservable characteristics. To further our understanding about the impact of supplier acquisition on citation lag, I regressed time lag on *Acquisition* and *Cite Competitor*, controlling for target, year of citing patent, and buyer fixed effects. Also, standard errors are clustered at the citing firm level, to correct for correlation of the error terms among citations of the same firm (Bertrand, Duflo, and Mullainathan, 2004). The results of those regressions for the set of firms that are in the upstream of the buyer firm are in Table 2.2. Columns 1 through 8 correspond to different thresholds of vertical integration, ranging from 1 percent up to 8 percent of the M&A target's industry participation in the total final product.

Table 2.2 shows the effect of acquisition on the target's citations of competitors' patents. The positive and significant coefficient of the interaction between acquired and cite competitor suggests that, after acquisition, targets that are in the upstream cite older competitors' patents. This evidence corroborates the hypothesis and agrees with the visual inspection of the data, but it is not enough. It is possible that any acquisition leads to an increase in the time lag for citations of competitors' patents, while the hypothesis specifies that we should observe higher lags for targets in the buyer's upstream. To check this possibility, I ran the same specification for firms that get acquired but are not in the upstream of the buyer firm. Results for those regressions are in Table 2.3, where different columns show results for different upstream thresholds. Different from the results for upstream firms, the coefficient of the cross product between *Acquired* and *Cite Competitor* is not significant in all upstream definitions. Table 2.3 shows that mere acquisition is not enough to increase the time lag for citations of competitors' patents.

It is possible that, after acquisition, targets in the upstream will start to pursue different patents, affecting the characteristics of patent citations. To control for that, I add variables proxying for the characteristics of citing and cited patents. Results for those checks are shown in Tables 4 and 5. Table 2.4 contains the results for the subset of targets that are in the upstream. As before, firms in the upstream tend to cite older competitors' patents after getting acquired. In Table 2.4, citations to competitors' patents before acquisition are actually newer than other citations. The faster technology transfer from the downstream industry is expected since upstream firms should be paying close attention to technologies developed by their possible customers. As for control variables, citations that have narrower scope (proxied by the number of claims) and higher mean lag tend to be cited later on. The positive coefficient for *daysafterbuy* just means that as time goes by, older patents are cited. As before, Table 2.5 shows no significant delays for citations to competitors' patents after acquisition for the subset of targets that are not in the buyer's upstream. The coefficient for citations to competitors' patents before acquisition is now positive, although not very precise. The only other noteworthy change is that, in this subset, patents with narrower scope tend to cite older patents.

The direction and significance of the main results remain unchanged. Note that both Table 2.4 and Table 2.5 show relative stability of the coefficient values for the control variables across different upstream thresholds. This result gives confidence that the effects found are not dependent on the definition of upstream adopted.

Finally, I determine if the difference in citations' lag after acquisitions between upstream and non-upstream is statistically significant. In other words, I compute the triple difference coefficient, as shown in Table 2.6. All triple interactions are positive and significant, corroborating the hypothesis.

To further control for changes in patenting behavior, I run the triple differences adding fixed effects for the first International Patent Classification class that shows up in the citing patent. By using the patent class (*Main class*), I further control for the types of patents that the targets applied for. There are 411 unique patent classes in the dataset, suggesting that the patent classification is reasonably fine. The side effect of using patent classifications is that the universe of observation with class information drops to 498,167. The regressions incorporating classification fixed effects are presented in Table 2.7, showing results that support the paper's hypothesis.

A key assumption of differences in differences models is the parallel trends hypothesis. One way to test this hypothesis is to run a leads and lags model (Angrist and Pischke, 2008). The result of the leads and lags interaction in the triple difference model is displayed in Figure 2.6. Before the acquisition, differences in time lag between upstream and non-upstream firms are mostly not significant. Two years after acquisition, citation lags for upstream firms surpass lags for non-upstream firms in most of the years.

Another way to test the assumptions of the differences in differences model is to include firm specific time trends (Angrist and Pischke, 2008). The result of the addition of the interaction between years after acquisition and citing firm dummies is presented in Table 2.8. Except for the most stringent definition of upstream[18], all equations exhibit a positive triple interaction term.

Still, it is possible that newly acquired upstream firms do not change their patenting behavior, but rather their citations behavior for equivalent patents. If that is true, we should not observe changes in lags of citations added by the examiners or other third parties. Table 2.8 shows the regressions using only the citations not added by the patent applicant. The results show that citations to competitors after acquisition added by other agents also exhibit higher lag.

Finally, if the increase in lag is generated exclusively by citations added by agents other than the applicant, we might be suspicious that the mechanism that results in higher lags does depend on

[18] That is, the definition where upstream industries are the ones that contribute with at least 8 percent of the total output.

the patent applicant. Table 2.9 runs the same regression as Table 2.8, but now only includes citations added by the patent applicant. Again, the triple difference coefficient is positive in all specifications, except for the most restrictive definition of upstream industry.

Although some alternative explanations are not supported by the data, the results are not causal. The targets were not acquired randomly, so it is possible that some unobserved characteristic that evolves over time can be causing the results. While these results do not show causality, they are a robust indication that acquired upstream firms cite older competitors' patents, as predicted by the theory. This fact offers support to the theoretical model presented, increasing the confidence in its applicability to the real world.

2.6 Conclusion

The model proposed here identifies how vertical integration affects the information flow between producers and suppliers. In doing that, it also presents a step towards a general equilibrium theory of the firm by taking into account the response of other agents to vertical integration. Disintegrated industries allow firms to tap into the information from competitors more easily. Integration, on the other hand, enables the integrated firm to reap the full benefits of the target's knowledge, but often at the expense of competitors' knowledge. The optimal market structure depends on characteristics of the demand and technology, in addition to the balance between the importance of a particular producer's information and the information from competitors.

Viewing integration decisions under the light of knowledge flows allows us to better advise managers about the benefits and pitfalls of integrated and disintegrated structures. The main benefit of using markets to acquire inputs is the access to a wide variety of informational sources. The benefit of the wisdom of the crowds is well known (Clemen and Winkler, 1986; Page, 2008; Palm and Zellner, 1992) and can be an important source of uncertainty mitigation, especially when managers face a highly uncertain environment. This can help explain why integrated firms can be perceived as slow movers in dynamic markets. As Michael Dell said (Magretta, 1998), "With vertical integration, you can be an efficient producer – as long as the world is not changing very much."

In summary, managers need to weigh how important various sources of information are when deciding to integrate vertically. If input quality depends crucially on the knowledge generated by all market participants, the market solution is optimal. Conversely, if one producer has better information, the market solution might be providing too little incentive for that producer to share it.

The results help to explain a series of empirical facts that the literature has identified so far. First, it sheds light on why some firms remain integrated while others move toward disintegration (Argyres and Bigelow, 2010; Helfat and Campo-Rembado, 2016; Kapoor, 2013; Kapoor and Adner, 2012). The integrated firm suffers when it cannot take advantage of the informational synergies caused by the integration. Firms then disintegrate to take advantage of competitors' information, since using markets induces competitors to share information. Firms that keep the integrated structure need to have an informational advantage from the interface input/final product to balance the loss of information from competitors. This is consistent with the empirical finding that integrated firms have a bigger emphasis on systemic innovations (Kapoor, 2013).

The problem of why incumbent firms sometimes fail to adopt new technologies (Dosi, 1982; Henderson and Clark, 1990; Utterback and Acee, 2005) can also benefit from the intuition built from this model. Here, the wrong choice of organizational structure can decrease the quantity or quality of the information received by incumbents. Integrated incumbents will be isolated from the flow of information and, unless they are able to anticipate market movements by themselves, they will have an informational disadvantage. Managers can overestimate the integrated firm's capacity for producing knowledge, or underestimate the usefulness of competitors' information, leading to an unwelcomed integration. Also, managers can pay too little attention to how to correctly incentivize individuals to share information within the firm, leading to inefficient hoarding of information (Pierce, 2012) even if the firm as a whole would benefit from it. Alternatively, managers could fail to integrate when this structure would produce more information. In any case, the wrong vertical structure of the cluster of producers and suppliers could result in a knowledge deficiency, leading to failures in discerning useful innovations.

Finally, the model shows how internal organization can influence firm capabilities, responding to a call to integrate those two concepts (Argyres et al., 2012). The way a firm organizes changes the flow of information from market participants to the firm. This difference in information flow can cause a difference in the firm's knowledge set and end up shaping future firm capabilities.

2.7 References

Alcácer J, Chung W (2007) Location Strategies and Knowledge Spillovers. *Manag. Sci.* 53(5):760–776.

Alcácer J, Oxley J (2014) Learning by supplying. *Strateg. Manag. J.* 35(2):204–223.

Anand KS, Goyal M (2009) Strategic Information Management Under Leakage in a Supply Chain. *Manag. Sci.* 55(3):438–452.

Angrist JD, Pischke JS (2008) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press).

Argyres N, Bigelow LS (2010) Innovation, Modularity, and Vertical Deintegration: Evidence from the Early U.S. Auto Industry. *Organ. Sci.* 21(4):842–853.

Argyres NS, Felin T, Foss N, Zenger T (2012) Organizational Economics of Capability and Heterogeneity. *Organ. Sci.* 23(5):1213–1226.

Arora A, Cohen WM, Walsh JP (2016) The acquisition and commercialization of invention in American manufacturing: Incidence and impact. *Res. Policy* 45(6):1113–1128.

Atalay E, Hortaçsu A, Syverson C (2014) Vertical integration and input flows. *Am. Econ. Rev.* 104(4):1120–1148.

Baccara M (2007) Outsourcing, information leakage, and consulting firms. *Rand J. Econ.* 38(1):269–289.

Barney J (1991) Firm Resources and Sustained Competitive Advantage. *J. Manag.* 17(1):99–120.

Bertrand M, Duflo E, Mullainathan S (2004) How Much Should We Trust Differences-in-Differences Estimates? *Q. J. Econ.* 119(1):249–275.

Brandenburger A, Stuart H (2007) Biform Games. *Manag. Sci.* 53(4):537–549.

Brusoni S, Prencipe A, Pavitt K (2001) Knowledge Specialization, Organizational Coupling, and the Boundaries of the Firm: Why Do Firms Know More than They Make? *Adm. Sci. Q.* 46(4):597–621.

Clemen RT, Winkler RL (1986) Combining Economic Forecasts. *J. Bus. Econ. Stat.* 4(1):39–46.

Cohen WM, Levinthal DA (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Adm. Sci. Q.* 35(1):128–152.

Cohen WM, Nelson RR, Walsh JP (2000) *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)* (National Bureau of Economic Research).

Conner KR, Prahalad CK (1996) A Resource-Based Theory of the Firm: Knowledge versus Opportunism. *Organ. Sci.* 7(5):477–501.

Demsetz H (1988) The Theory of the Firm Revisited. *J. Law Econ. Organ.* 4(1):141–161.

Dosi G (1982) Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Res. Policy* 11(3):147–162.

Dushnitsky G, Shaver JM (2009) Limitations to interorganizational knowledge acquisition: the paradox of corporate venture capital. *Strateg. Manag. J.* 30(10):1045–1064.

Dyer JH, Hatch NW (2006) Relation-specific capabilities and barriers to knowledge transfers: creating advantage through network relationships. *Strateg. Manag. J.* 27(8):701–719.

Dyer JH, Singh H (1998) The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage. *Acad. Manage. Rev.* 23(4):660–679.

Grant RM (1996) Toward a knowledge-based theory of the firm. *Strateg. Manag. J.* 17(S2):109–122.

Grossman SJ, Hart OD (1986) The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *J. Polit. Econ.* 94(4):691–719.

Hakansson H, Havila V, Pedersen AC (1999) Learning in networks. *Ind. Mark. Manag.* 28(5):443–452.

Hall BH, Mairesse J, Mohnen P (2010) Measuring the Returns to R&D. *Handb. Econ. Innov.* 2:1033–1082.

Hart O, Moore J (1990) Property Rights and the Nature of the Firm. *J. Polit. Econ.* 98(6):1119–1158.

Hayek FA (1945) The Use of Knowledge in Society. *Am. Econ. Rev.* 35(4):519–530.

Helfat CE, Campo-Rembado MA (2016) Integrative Capabilities, Vertical Integration, and Innovation Over Successive Technology Lifecycles. *Organ. Sci.* 27(2):249–264.

Henderson RM, Clark KB (1990) Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Adm. Sci. Q.* 35(1):9–30.

Huckman RS, Pisano GP (2010) Flextronics International, Ltd.

Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Q. J. Econ.* 108(3):577–598.

Kapoor R (2013) Persistence of Integration in the Face of Specialization: How Firms Navigated the Winds of Disintegration and Shaped the Architecture of the Semiconductor Industry. *Organ. Sci.* 24(4):1195–1213.

Kapoor R, Adner R (2012) What Firms Make vs. What They Know: How Firms' Production and Knowledge Boundaries Affect Competitive Advantage in the Face of Technological Change. *Organ. Sci.* 23(5):1227–1248.

Katila R, Rosenberger JD, Eisenhardt KM (2008) Swimming with Sharks: Technology Ventures, Defense Mechanisms and Corporate Relationships. *Adm. Sci. Q.* 53(2):295–332.

Klepper S (1997) Industry Life Cycles. *Ind. Corp. Change* 6(1):145–182.

Knott AM, Posen HE, Wu B (2009) Spillover Asymmetry and Why It Matters. *Manag. Sci.* 55(3):373–388.

Kogut B, Zander U (1992) Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organ. Sci.* 3(3):383–397.

Kogut B, Zander U (1996) What Firms Do? Coordination, Identity, and Learning. *Organ. Sci.* 7(5):502–518.

Kotabe M, Martin X, Domoto H (2003) Gaining from vertical partnerships: knowledge transfer, relationship duration, and supplier performance improvement in the U.S. and Japanese automotive industries. *Strateg. Manag. J.* 24(4):293–316.

Langlois RN (2003) The vanishing hand: the changing dynamics of industrial capitalism. *Ind. Corp. Change* 12(2):351–385.

Magretta J (1998) The power of virtual integration: An interview with Dell Computer's Michael Dell. *Harv. Bus. Rev.*

Mesquita LF, Anand J, Brush TH (2008) Comparing the resource-based and relational views: knowledge transfer and spillover in vertical alliances. *Strateg. Manag. J.* 29(9):913–941.

Modi SB, Mabert VA (2007) Supplier development: Improving supplier performance through knowledge transfer. *J. Oper. Manag.* 25(1):42–64.

Nickerson J, Yen CJ, Mahoney JT (2012) Exploring the Problem-Finding and Problem-Solving Approach for Designing Organizations. *Acad. Manag. Perspect.* 26(1):52–72.

Nickerson JA, Zenger TR (2004) A Knowledge-Based Theory of the Firm: The Problem-Solving Perspective. *Organ. Sci.* 15(6):617–632.

Nonaka I (1994) A Dynamic Theory of Organizational Knowledge Creation. *Organ. Sci.* 5(1):14–37.

Oxley J, Wada T (2009) Alliance Structure and the Scope of Knowledge Transfer: Evidence from U.S.-Japan Agreements. *Manag. Sci.* 55(4):635–649.

Page SE (2008) *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* (Princeton University Press).

Pahnke EC, McDonald R, Wang D, Hallen B (2015) Exposed: Venture Capital, Competitor Ties, and Entrepreneurial Innovation. *Acad. Manage. J.* 58(5):1334–1360.

Palm FC, Zellner A (1992) To combine or not to combine? issues of combining forecasts. *J. Forecast.* 11(8):687–701.

Pierce L (2012) Organizational Structure and the Limits of Knowledge Sharing: Incentive Conflict and Agency in Car Leasing. *Manag. Sci.* 58(6):1106–1121.

Schildt H, Keil T, Maula M (2012) The temporal effects of relative and firm-level absorptive capacity on interorganizational learning. *Strateg. Manag. J.* 33(10):1154–1173.

Takeishi A (2002) Knowledge Partitioning in the Interfirm Division of Labor: The Case of Automotive Product Development. *Organ. Sci.* 13(3):321–338.

Tan KH, Wong WP, Chung L (2015) Information and Knowledge Leakage in Supply Chain. *Inf. Syst. Front.*

Tsui E, Waters R (2009) Asustek to spin off motherboard arm. *Financial Times* (December 11) https://www.ft.com/content/699416d2-e67a-11de-98b1-00144feab49a.

Utterback JM, Acee HJ (2005) Disruptive technologies: an expanded view. *Int. J. Innov. Manag.* 9(01):1–17.

Vasudeva G, Anand J (2011) Unpacking absorptive capacity: A study of knowledge utilization from alliance portfolios. *Acad. Manage. J.* 54(3):611–623.

Wernerfelt B (1984) A resource-based view of the firm. *Strateg. Manag. J.* 5(2):171–180.

Williamson OE (1979) Transaction-Cost Economics: The Governance of Contractual Relations. *J. Law Econ.* 22(2):233–261.

Williamson OE (1985) The economic institutions of capitalism. *Simon Schuster*.

Zenger TR, Felin T, Bigelow LS (2011) *Theories of the Firm-Market Boundary* (Social Science Research Network, Rochester, NY).

Figure 2.1: Timing of the game

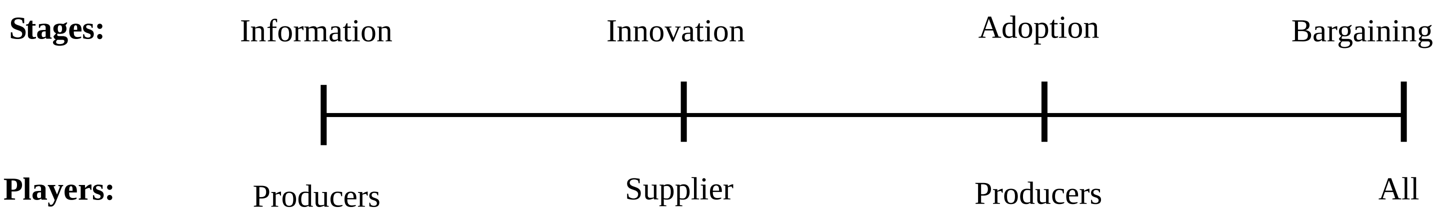


Figure 2.2: Cooperative Game

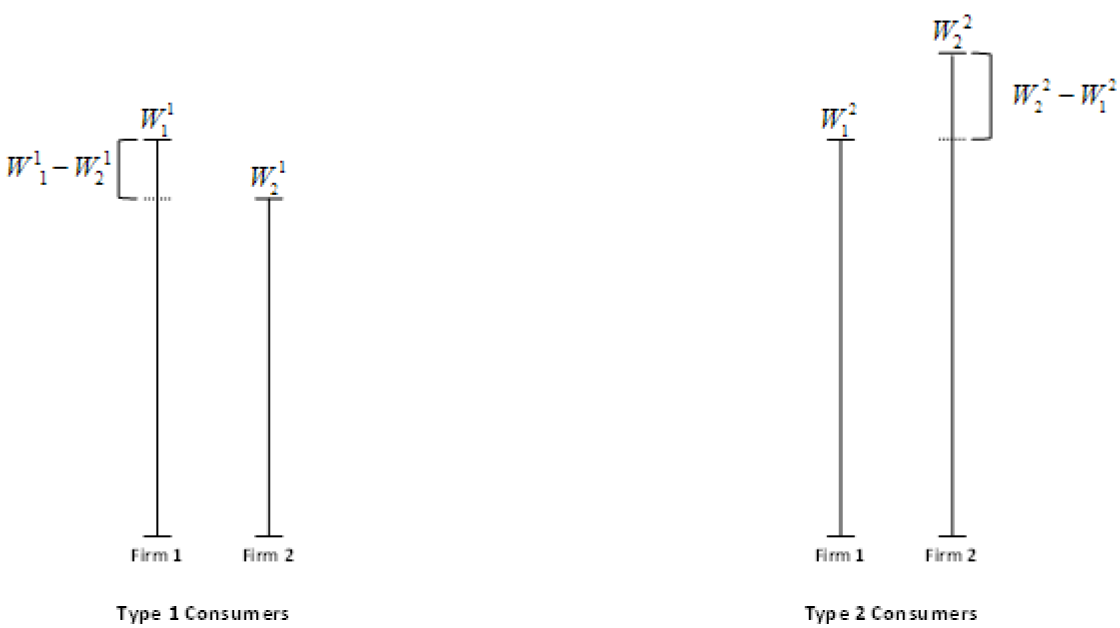


Figure 2.3: Producer 2's information loss on market 2 with integration

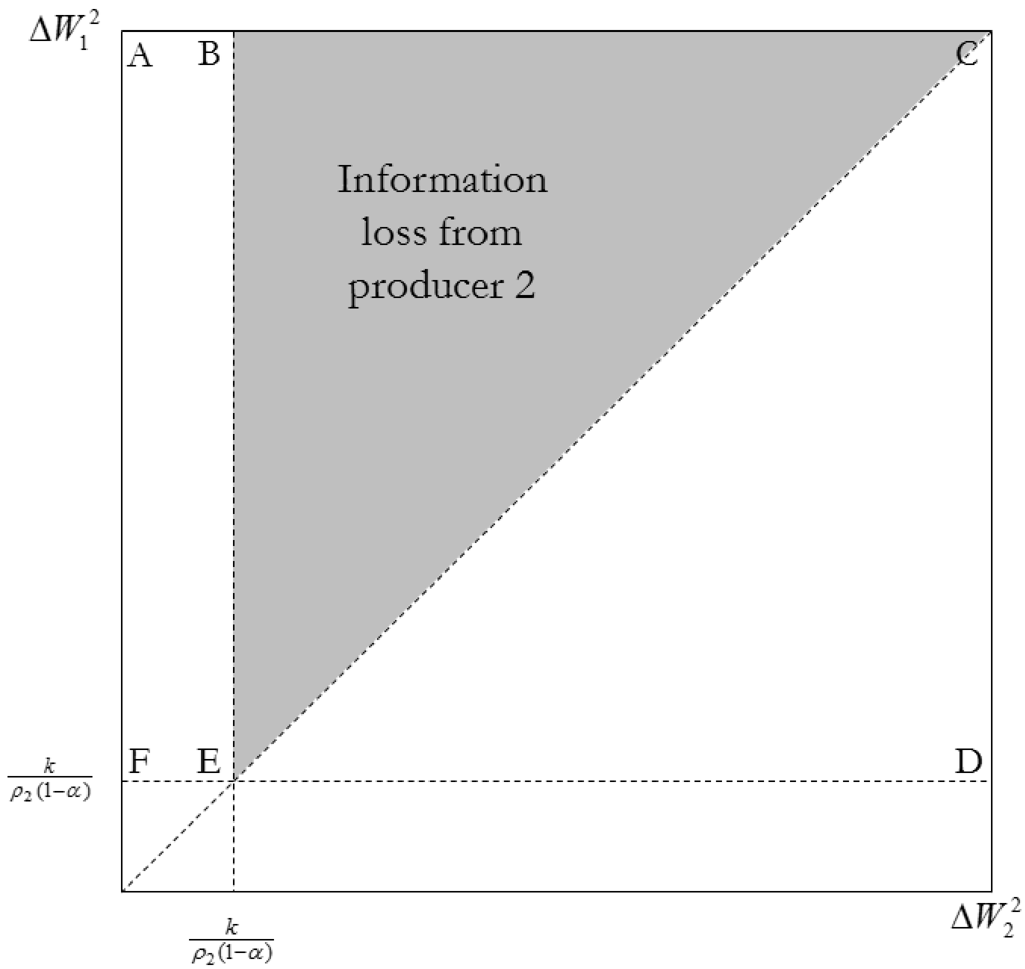
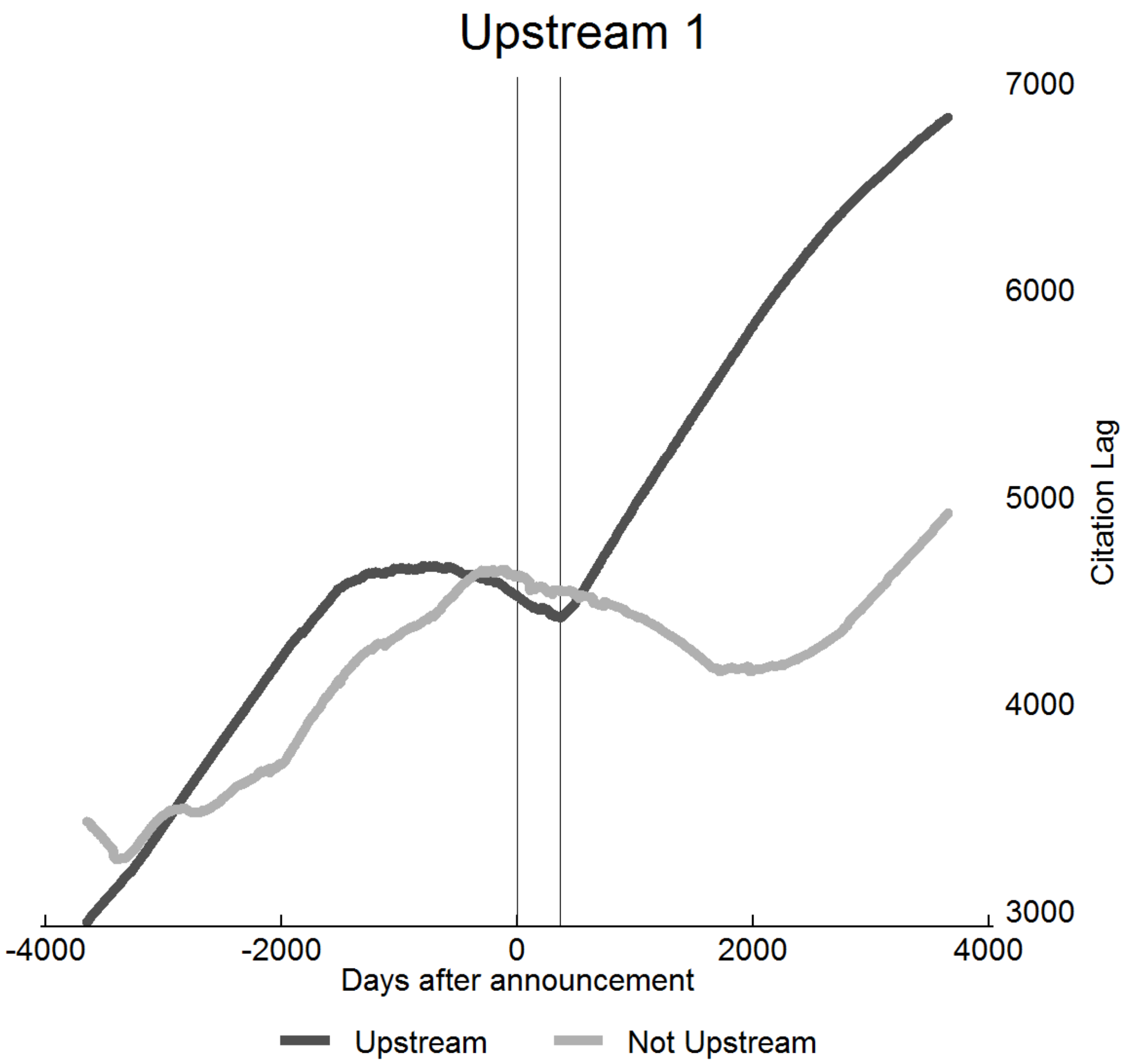
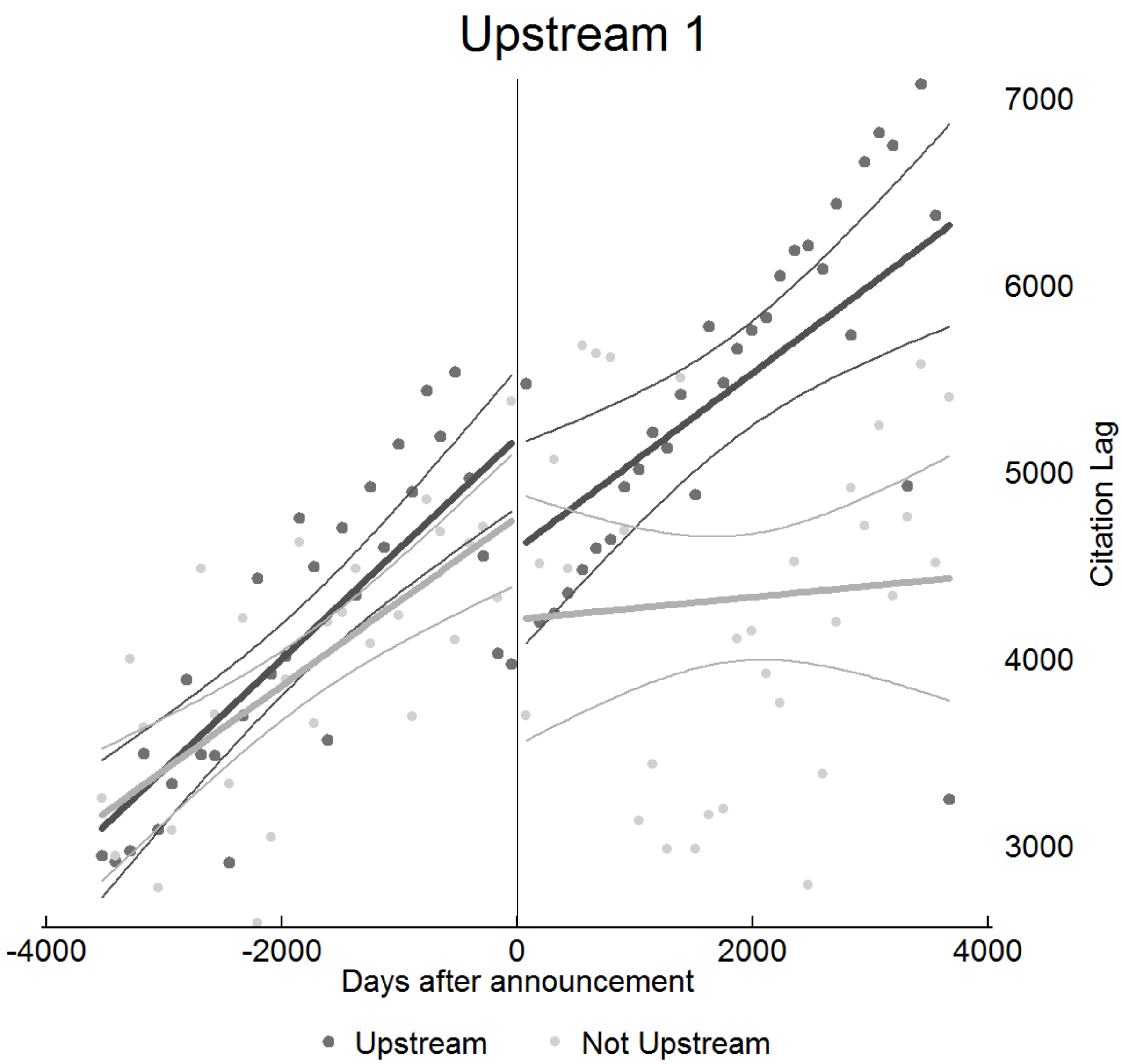


Figure 2.4: Competitor's patents citation lag



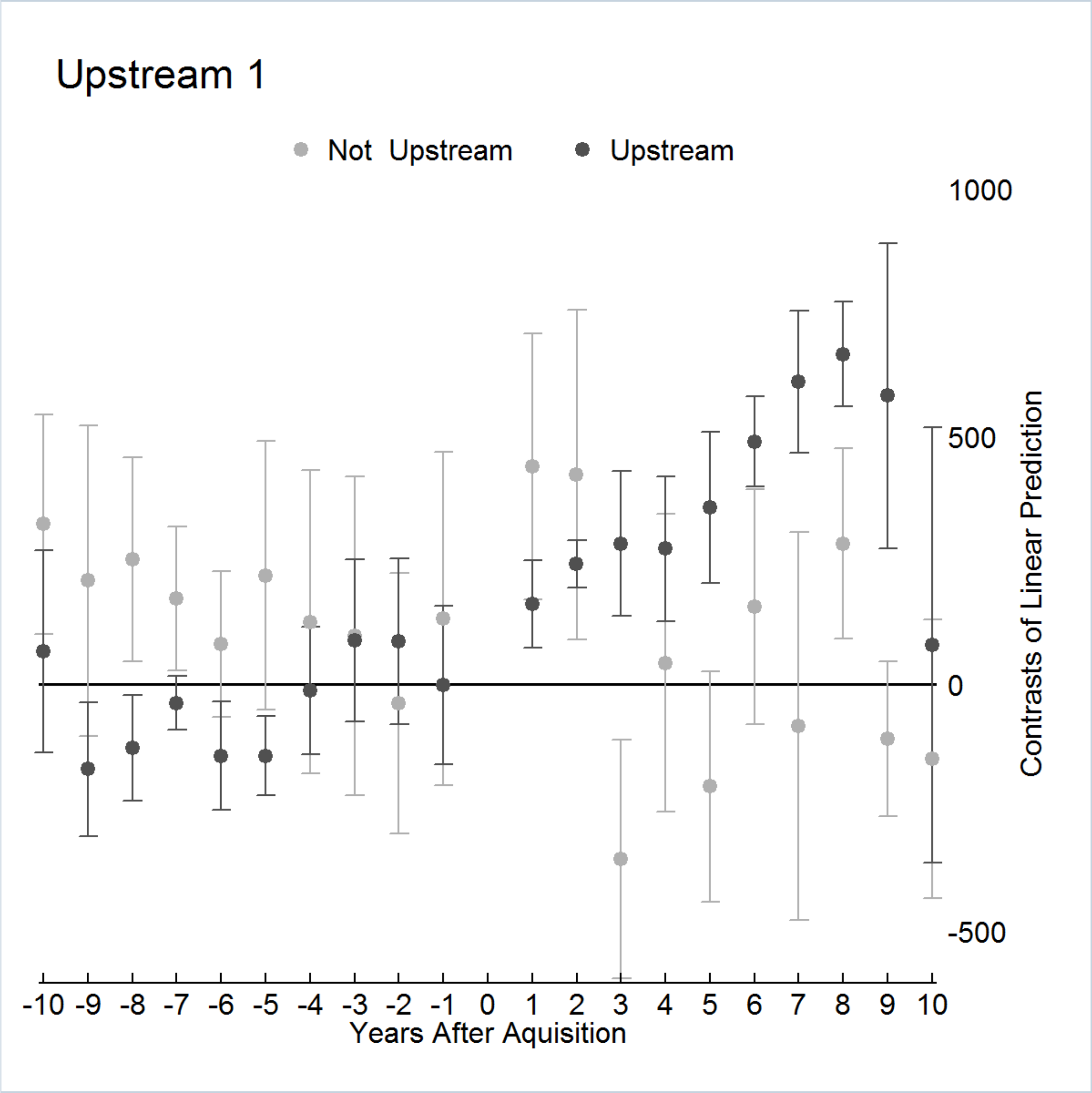
Note. This graph represents the raw data using the lowess smoother with bandwidth 0.4. The vertical line to the right of zero marks one year after the M&A announcement. Firms are classified as being in the upstream if target's industry supplies at least one percent of the total output for the buyer's industry according to the IO tables provided by the Bureau of Economic Analysis. Results with other thresholds for upstream classification and different bandwidth values are in the Appendix.

Figure 2.5: Piecewise linear regression of the citation lag of competitors' patents



Note. Every point is constructed by discretizing the days after announcement into intervals of 120 days and then averaging out all the citation lags within the interval. Firms are classified as being in the upstream if target's industry supplies at least one percent of the total output for the buyer's industry according to the IO tables provided by the Bureau of Economic Analysis. Results with other thresholds for upstream classification and different interval lengths and are shown in the Appendix.

Figure 2.6: Leads and lags, triple difference.



Note Firms are classified as being in the upstream if target's industry supplies at least one percent of the total output for the buyer's industry according to the IO tables provided by the Bureau of Economic Analysis. Results with other thresholds for upstream classification and different interval lengths and are shown in the Appendix.

Table 2.1: Summary statistics

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Time Lag	649,593	4,109	2,647	0	15,721
Cite Competitor	649,593	0.0538	0.226	0	1
Num Claims (citation)	649,592	20.26	17.40	1	868
Mean Lag (citation)	649,593	3,654	1,872	0	14,534
NumCit (citation)	649,593	96.90	147.0	1	3,603
Num Claims (citing)	649,593	21.46	13.98	1	299
NumCit made (citing)	649,593	176.8	256.7	1	1,179
daysafterbuy	649,593	-645.1	1,991	-3,650	3,650
Acquired	649,593	0.370	0.483	0	1
SQRdaysafterbuy	649,593	4.380e+06	3.835e+06	0	1.332e+07

Table 2.2: Effect of acquisition on competitors' patents citation lag; citing firm is on buyer's upstream. Regression without controls.

VARIABLES	(1) Upstream 1	(2) Upstream 2	(3) Upstream 3	(4) Upstream 4	(5) Upstream 5	(6) Upstream 6	(7) Upstream 7	(8) Upstream 8
Acquired	-61.92	-90.66	-100.2	-76.06	-206.3	-228.9	-284.5	-205.1*
	(126.3)	(135.9)	(142.3)	(151.8)	(139.6)	(152.9)	(194.0)	(110.3)
Cite Competitor	74.80	73.69	75.49	118.8	109.5	104.1	203.2	695.3***
	(121.9)	(122.1)	(122.8)	(127.5)	(126.2)	(126.6)	(169.9)	(136.8)
Acquired * Cite Competitor	1,505***	1,519***	1,523***	1,466***	1,486***	1,496***	1,399***	892.8***
	(122.9)	(121.8)	(122.7)	(126.3)	(125.7)	(125.6)	(173.6)	(99.53)
Target FE	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	351,712	339,388	329,237	270,172	247,809	235,711	190,377	100,018
R-squared	0.153	0.149	0.145	0.167	0.162	0.158	0.149	0.125

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.3: Effect of acquisition on competitors' patents citation lag; citing firm is not on buyer's upstream. Regression without controls.

VARIABLES	(1) Upstream 1	(2) Upstream 2	(3) Upstream 3	(4) Upstream 4	(5) Upstream 5	(6) Upstream 6	(7) Upstream 7	(8) Upstream 8
Acquired	125.8	132.5	140.3	99.07	116.5	119.3	140.3	31.80
	(114.7)	(108.0)	(106.2)	(106.1)	(99.15)	(94.49)	(92.47)	(120.0)
Cite Competitor	51.01	48.79	39.44	-58.01	-52.99	-46.31	-124.0	33.45
	(287.7)	(285.5)	(282.2)	(295.5)	(278.6)	(273.8)	(229.0)	(118.4)
Acquired * Cite Competitor	138.7	107.8	136.2	235.6	381.4	380.9	445.7	259.0
	(385.6)	(374.7)	(364.9)	(364.4)	(337.9)	(331.8)	(295.3)	(257.2)
Target FE	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	297,875	310,199	320,350	379,415	401,778	413,876	459,210	549,568
R-squared	0.221	0.222	0.223	0.199	0.200	0.200	0.199	0.195

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.4: Effect of acquisition on competitors' patents citation lag; citing firm is on buyer's upstream.

VARIABLES	(1) Upstream 1	(2) Upstream 1	(3) Upstream 2	(4) Upstream 3	(5) Upstream 4	(6) Upstream 5	(7) Upstream 6	(8) Upstream 7	(9) Upstream 8
Acquired		-108.8*	-108.6	-112.7	-105.1	-87.24	-55.00	-58.11	-16.25
		(63.30)	(66.67)	(69.16)	(76.61)	(88.44)	(95.19)	(117)	(120.9)
Cite Competitor		-53.56**	-53.23**	-54.36**	-58.23**	-60.32**	-659**	-74.00***	-125.5
		(26.10)	(25.82)	(25.43)	(24.34)	(25.06)	(24.10)	(18.93)	(135.6)
Acquired * Cite Competitor		383.2***	392.2***	392.5***	384.5***	370.0***	367.8***	374.1***	430.8***
		(48.53)	(44.32)	(45.43)	(46.43)	(40.54)	(38.62)	(30.40)	(105.0)
Num Claims (citation)	-3.012***	-2.778***	-2.646***	-2.546***	-2.774***	-2.718***	-2.501**	-2.150**	-904
	(0.685)	(0.759)	(0.756)	(0.767)	(0.892)	(0.951)	(0.942)	(0.819)	(155)
Mean Lag (citation)	1.187***	1.185***	1.181***	1.181***	1.199***	1.197***	1.196***	1.191***	1.189***
	(0.0170)	(0.0172)	(0.0167)	(0.0171)	(0.0179)	(0.0191)	(0.0199)	(0.0194)	(0.0264)
NumCit (citation)	0.0796	-0.00742	-0.0164	-0.000500	-0.0693	0.193	0.246	0.356	0.237*
	(0.315)	(0.283)	(0.288)	(0.294)	(0.307)	(0.277)	(0.286)	(0.312)	(0.117)
Num Claims (citing)	-0.965	-0.886	-0.617	-0.696	-0.996	-0.662	-0.589	-2.310**	-260***
	(061)	(057)	(047)	(110)	(425)	(470)	(551)	(010)	(0.386)
NumCit made (citing)	0.00952	-0.0229	-0.0256	-0.0350	-0.117	-0.126	-0.139	-0.0656	0.230***
	(0.237)	(0.223)	(0.221)	(0.224)	(0.319)	(0.342)	(0.353)	(0.323)	(0.0185)
daysafterbuy	0.728***	0.752***	0.765***	0.760***	0.704***	0.743***	0.751***	0.746***	0.794***
	(0.0756)	(0.0802)	(0.0797)	(0.0810)	(0.0896)	(0.0861)	(0.0883)	(0.0968)	(0.0454)
SQRdaysafterbuy	6.55e-06	8.26e-06	8.62e-06	8.71e-06	1.42e-05*	1.74e-05**	1.79e-05**	2.18e-05***	41e-05
	(8.08e-06)	(7.94e-06)	(7.94e-06)	(8.12e-06)	(7.34e-06)	(7.25e-06)	(7.00e-06)	(5.23e-06)	(48e-05)
Target FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	351,712	351,712	339,388	329,237	270,172	247,809	235,711	190,377	100,018
R-squared	0.854	0.855	0.858	0.859	0.837	0.845	0.846	0.852	0.898

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.5: Effect of acquisition on competitors' patents citation lag; citing firm is not on buyer's upstream.

VARIABLES	(1) Upstream 1	(2) Upstream 1	(3) Upstream 2	(4) Upstream 3	(5) Upstream 4	(6) Upstream 5	(7) Upstream 6	(8) Upstream 7	(9) Upstream 8
Acquired		65.65	57.25	57.01	47.58	37.83	33.69	551	63.03
		(84.18)	(80.72)	(78.95)	(76.73)	(72.49)	(70.00)	(67.60)	(67.28)
Cite Competitor		145.6*	142.5*	140.8*	135.7*	133.2*	136.4**	127.3**	20.50
		(75.87)	(75.35)	(74.76)	(75.15)	(719)	(69.43)	(57.22)	(60.08)
Acquired * Cite Competitor		28.18	12.14	30.55	398	-4.993	-6.999	3.691	105.5
		(152.5)	(149.3)	(143.6)	(144.6)	(143.0)	(140.8)	(134.9)	(133.9)
Num Claims (citation)	-4.823***	-4.802***	-4.858***	-4.862***	-4.409***	-4.381***	-4.464***	-4.512***	-4.092***
	(0.866)	(0.878)	(0.854)	(0.820)	(0.807)	(0.796)	(0.787)	(0.778)	(0.665)
Mean Lag (citation)	1.234***	1.234***	1.236***	1.235***	1.211***	1.211***	1.211***	1.212***	1.210***
	(0.0195)	(0.0195)	(0.0188)	(0.0183)	(0.0233)	(0.0222)	(0.0215)	(0.0207)	(0.0176)
NumCit (citation)	-0.334	-0.335	-0.317	-0.328	-0.253	-0.370*	-0.385*	-0.398*	-0.250
	(0.259)	(0.257)	(0.251)	(0.248)	(0.242)	(0.222)	(0.218)	(0.208)	(0.223)
Num Claims (citing)	-2.958***	-2.930***	-3.060***	-2.927***	-2.482***	-2.613***	-2.641***	-2.062**	-2.264***
	(0.806)	(0.815)	(0.792)	(0.769)	(0.757)	(0.730)	(0.722)	(0.811)	(0.794)
NumCit made (citing)	0.299	0.288	0.302	0.317	0.169	0.166	0.175	0.0987	-0.106
	(0.244)	(0.245)	(0.245)	(0.243)	(0.151)	(0.147)	(0.149)	(0.167)	(0.233)
daysafterbuy	0.605***	0.594***	0.585***	0.593***	0.624***	0.611***	0.604***	0.623***	0.629***
	(0.0782)	(0.0771)	(0.0752)	(0.0738)	(0.0682)	(0.0664)	(0.0650)	(0.0610)	(0.0710)
SQRdaysafterbuy	-1.32e-05	-1.39e-05	-1.35e-05	-1.28e-05	-1.62e-05	-1.53e-05*	-1.51e-05*	-1.37e-05	-1.16e-05
	(11e-05)	(09e-05)	(05e-05)	(02e-05)	(9.98e-06)	(9.15e-06)	(9.06e-06)	(9.11e-06)	(8.75e-06)
Target FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	297,874	297,874	310,198	320,349	379,414	401,777	413,875	459,209	549,567
R-squared	0.827	0.827	0.826	0.826	0.845	0.840	0.840	0.838	0.833

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.6: Effect of acquisition on competitors' patents citation lag; triple difference.

VARIABLES	(1) Upstream 1	(2) Upstream 2	(3) Upstream 3	(4) Upstream 4	(5) Upstream 5	(6) Upstream 6	(7) Upstream 7	(8) Upstream 8
Acquired	31.97	27.20	23.81	24.05	5.496	0.883	-8.503	-14.66
	(102.6)	(98.14)	(95.88)	(94.76)	(91.72)	(89.28)	(86.24)	(83.77)
Cite Competitor	142.3*	139.3*	137.4*	132.5*	128.6*	131.9*	122.8**	19.46
	(78.64)	(78.40)	(77.70)	(75.37)	(71.47)	(69.81)	(57.65)	(60.41)
Acquired * Cite Competitor	21.37	6.273	25.10	29.92	-7.131	-8.695	6.748	110.0
	(161.1)	(157.3)	(150.7)	(149.9)	(146.0)	(143.8)	(137.9)	(135.6)
Acquired * Upstream	-113.5	-110.9	-107.5	-112.8	-47.15	-27.18	35.00	75.09
	(142.7)	(146)	(149)	(142)	(142.8)	(149.4)	(146.5)	(155.5)
Cite Competitor * Upstream	-194.2**	-191**	-189.5**	-186.9**	-182.3**	-187.2**	-194.8***	-169.0
	(83.56)	(83.33)	(82.54)	(80.08)	(76.19)	(74.22)	(59.95)	(113.4)
Acquired * Cite Competitor * Upstream	334.0*	352.1**	332.3**	329.9**	373.9**	377.0**	376.1**	350.1**
	(175.5)	(173)	(165.1)	(163.6)	(160.1)	(157.4)	(150.1)	(153.3)
Target FE	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	649,586	649,586	649,586	649,586	649,586	649,586	649,586	649,586
R-squared	0.841	0.841	0.841	0.841	0.841	0.841	0.841	0.841

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. All controls in tables 4 and 5 are also taken into consideration, but not shown. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.7: Effect of acquisition on competitors' patents citation lag; triple difference. Patent classification fixed effect.

VARIABLES	(1) Upstream 1	(2) Upstream 2	(3) Upstream 3	(4) Upstream 4	(5) Upstream 5	(6) Upstream 6	(7) Upstream 7	(8) Upstream 8
Acquired	2.586	-2.249	-5.654	-9.738	-26.69	-27.58	-35.61	-35.66
	(102.8)	(97.03)	(94.76)	(94.07)	(88.10)	(85.89)	(865)	(79.96)
Cite Competitor	110.9	107.7	107.2	102.2	95.96	102.4	92.40*	577
	(73.31)	(73.18)	(72.25)	(70.17)	(66.07)	(64.24)	(519)	(38.79)
Acquired * Cite Competitor	-58.90	-69.82	-59.72	-53.78	-80.24	-968	-68.17	-24.28
	(93.27)	(89.85)	(88.72)	(87.72)	(926)	(90.84)	(85.34)	(83.06)
	(76.20)	(75.96)	(75.07)	(73.10)	(69.18)	(66.99)	(55.06)	(129.1)
Acquired * Cite Competitor * Upstream	365.9***	379.1***	368.7***	361***	390.4***	405.9***	385.8***	465.0***
	(105.2)	(100.8)	(99.53)	(97.50)	(109)	(101)	(96.25)	(137.8)
Target FE	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES
Main class	YES	YES	YES	YES	YES	YES	YES	YES
Observations	498,167	498,167	498,167	498,167	498,167	498,167	498,167	498,167
R-squared	0.841	0.841	0.841	0.841	0.841	0.841	0.841	0.841

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. All controls in tables 4 and 5 are also taken into consideration, but not shown. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.8: Effect of acquisition on competitors' patents citation lag; triple difference. Firm specific time trends.

VARIABLES	(1) Upstream 1	(2) Upstream 2	(3) Upstream 3	(4) Upstream 4	(5) Upstream 5	(6) Upstream 6	(7) Upstream 7	(8) Upstream 8
Acquired	-21.34	-28.28	-24.05	-27.80	-22.56	-26.00	-18.59	-23.02
	(36.92)	(36.29)	(36.55)	(36.97)	(34.75)	(34.51)	(33.93)	(33.19)
Cite Competitor	148.3*	144.6*	144.1*	138.6*	130.5*	133.0*	117.1*	17.13
	(81.05)	(81.22)	(80.27)	(78.09)	(75.21)	(73.58)	(62.63)	(59.29)
Acquired * Cite Competitor	-29.41	-40.99	-20.80	-14.94	-38.61	-37.84	-7.606	93.53
	(106.6)	(103.7)	(99.32)	(99.09)	(100.8)	(99.00)	(96.40)	(98.39)
Acquired * Upstream	-8.660	10.43	-1.739	9.389	-7.326	6.593	-31.58	24.74
	(61.22)	(63.17)	(62.82)	(63.41)	(68.52)	(72.94)	(82.41)	(99.32)
Cite Competitor * Upstream	-177.5**	-172.0*	-172.9*	-168.1*	-161.5*	-165.6**	-165.4**	84.16
	(88.46)	(89.14)	(88.04)	(86.15)	(83.60)	(81.92)	(68.90)	(112.0)
Acquired * Cite Competitor * Upstream	400.9***	413.4***	393.4***	387.8***	423.2***	423.8***	409.4***	160.9
	(115.5)	(113.1)	(108.6)	(107.9)	(113.5)	(111.7)	(107.2)	(140.3)
Firm-Specific Time Trend	YES	YES	YES	YES	YES	YES	YES	YES
Target FE	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	649,586	649,586	649,586	649,586	649,586	649,586	649,586	649,586
R-squared	0.845	0.845	0.845	0.845	0.845	0.845	0.845	0.845

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. All controls in tables 4 and 5 are also taken into consideration, but not shown. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.9: Time Lag, triple difference, patent classification fixed effect, citations not added by applicants

VARIABLES	(1) Upstream 1	(2) Upstream 2	(3) Upstream 3	(4) Upstream 4	(5) Upstream 5	(6) Upstream 6	(7) Upstream 7	(8) Upstream 8
Acquired	-38.79	-39.40	-35.91	-36.61	-43.54	-43.65	-48.22	-50.96
	(103.0)	(96.46)	(94.92)	(94.28)	(88.14)	(86.08)	(81.48)	(79.88)
Cite Competitor	95.99	92.54	93.17	88.79	82.89	88.62	75.67	41.40
	(69.39)	(69.56)	(68.92)	(66.58)	(62.92)	(60.94)	(47.37)	(37.31)
Acquired * Cite Competitor	-198.5**	-198.0**	-189.5**	-185.1**	-201.9**	-208.0**	-192.7**	-154.6**
	(85.02)	(83.51)	(80.57)	(78.74)	(88.37)	(86.96)	(77.38)	(71.33)
Acquired * Upstream	15.72	21.15	9.129	11.80	46.20	52.09	91.76	108.8
	(135.2)	(133.3)	(135.8)	(135.4)	(137.7)	(141.9)	(140.3)	(163.1)
Cite Competitor * Upstream	-104.2	-98.78	-101.3	-97.17	-87.99	-97.62	-95.10*	-101.7
	(71.45)	(71.59)	(70.98)	(69.07)	(66.04)	(63.95)	(54.17)	(144.2)
Acquired * Cite Competitor * Upstream	449.0***	449.3***	442.8***	438.4***	461.6***	471.6***	466.2***	468.3***
	(93.92)	(91.03)	(88.15)	(85.00)	(94.63)	(92.89)	(84.46)	(142.8)
Target FE	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES
Mainclass	YES	YES	YES	YES	YES	YES	YES	YES
Observations	358,514	358,514	358,514	358,514	358,514	358,514	358,514	358,514
R-squared	0.840	0.840	0.840	0.840	0.840	0.840	0.840	0.840

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. All controls in tables 4 and 5 are also taken into consideration, but not shown. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 2.10: Time Lag, triple difference, patent classification fixed effect, citations added by applicants

VARIABLES	(1) Upstream 1	(2) Upstream 2	(3) Upstream 3	(4) Upstream 4	(5) Upstream 5	(6) Upstream 6	(7) Upstream 7	(8) Upstream 8
Acquired	18.73	14.18	13.47	16.33	-8.413	-8.444	-7.176	-0.507
	(76.40)	(76.67)	(76.72)	(76.18)	(80.61)	(79.31)	(73.79)	(70.48)
Cite Competitor	248.3**	251.4**	246.4**	241.4**	245.4**	249.6**	261.1***	50.03
	(106.6)	(105.5)	(102.6)	(101.6)	(99.14)	(98.25)	(93.92)	(70.74)
Acquired * Cite Competitor	67.16	50.81	69.74	73.88	66.05	48.57	98.16	309.8***
	(125.1)	(124.5)	(121.6)	(121.1)	(112.7)	(113.0)	(114.2)	(115.2)
Acquired * Upstream	-79.42	-64.68	-64.07	-75.00	12.86	16.29	10.78	-79.29
	(108.4)	(109.5)	(109.5)	(109.9)	(98.95)	(101.2)	(127.1)	(160.5)
Cite Competitor * Upstream	-232.5**	-236.6**	-232.3**	-227.1**	-233.1**	-238.3**	-253.8**	607.6**
	(117.8)	(116.5)	(113.6)	(112.7)	(109.9)	(108.7)	(103.2)	(303.6)
Acquired * Cite Competitor * Upstream	286.0**	305.2**	285.4**	281.1**	292.9**	313.4**	259.4**	-603.7*
	(136.5)	(135.8)	(132.6)	(132.1)	(123.4)	(123.5)	(121.6)	(324.2)
Target FE	YES	YES	YES	YES	YES	YES	YES	YES
Citing year FE	YES	YES	YES	YES	YES	YES	YES	YES
Buyer FE	YES	YES	YES	YES	YES	YES	YES	YES
Main class	YES	YES	YES	YES	YES	YES	YES	YES
Observations	109,369	109,369	109,369	109,369	109,369	109,369	109,369	109,369
R-squared	0.879	0.879	0.879	0.879	0.879	0.879	0.879	0.879

Notes. The dependent variable is the difference in application dates between the citing and the cited patents. All controls in tables 4 and 5 are also taken into consideration, but not shown. Each column represents a different threshold to determine which targets are in the upstream. Clustered standard errors at the citing firm level in parentheses.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Chapter 3: A new direction or lack of direction? The story of innovation with outsider CEOs

LEONARDO M. KLÜPPEL
Washington University in St. Louis

TREY CUMMINGS
Washington University in St. Louis

3.1 Introduction

External hires bring new information and perspectives to the firm, impacting its innovation output (Aime *et al.*, 2010; Görg and Strobl, 2005; Kaiser, Kongsted, and Rønde, 2015; Lewis and Yao, 2006). Although all external hires are possible sources of changes in innovation practices, the CEO is a special case. Besides having a large impact on firm performance (Mackey, 2008), the CEO sits at the top of the organizational chart, a privileged position to spot resource complementarities across all business units and to take care of long term planning (Chandler, 1962; Hambrick, 1994). For those reasons, the CEO position is uniquely suited for the delineation of the firm's innovation direction. This chapter is a first effort at understanding how and when outside CEOs affect the direction of innovation in relation to inside CEOs. The effect of outside CEOs on innovation is an important phenomenon to study because outside CEOs have been on the rise over the last half century (Cummings and Knott 2018, Murphy and Zabojnik

2007) and innovation is a primary driver of growth (Romer 1990, Solow 1957). Although prior literature has determined that outside CEOs are associated with lower patent impact (Balsmeier and Buchwald 2014) and decreased R&D productivity (Cummings and Knott 2018) which is hypothesized to be due to the lack of expertise in the firm's technology, little is known about the direction outside CEOs take the firm's innovation. We use the multi-dimensional vector and norm of each firm's patents to explore differences in innovation direction between firms with outside and inside CEOs. Extant literature indicates that outside CEOs are hired to make changes which likely manifest in a shift in innovation direction. However, our empirical analysis indicates that outside CEOs explore less (lower relative change in the firm's patent vector across time) compared to their inside counterparts and this relationship holds when using a propensity score matching model. We also find that the outside CEO's impact is muted when the firm's organizational structure decreases the influence of the CEO which provides circumstantial evidence that the CEO is, at least in part, driving the effect. Taken together, our results surprisingly indicate that outside CEOs do not necessarily take the firm in a new innovative direction but rather may generate a lack of innovative direction for the firm. These results have direct implications for firms that hire outside CEOs to generate technological change.

Recent management literature has begun to explore how CEO heterogeneity affects the firm's innovation (Balsmeier and Buchwald 2014, Cummings and Knott 2018, Galasso and Simcoe 2011, Hirshleifer et al. 2012). A firm's innovation path depends heavily on the CEO's experience, knowledge, and beliefs about the world (Hambrick and Mason, 1984). Those characteristics can be impacted by a multitude of experiences in the life of the CEO, as varied as living through economic depressions (Malmendier and Nagel, 2011) to his/her functional experiences (Crossland *et al.*, 2014). Past employment can also change the way CEOs process and respond to new information. It is known that experiences beyond the focal firm and industry can change the CEO's commitment to the firm's current strategy (Hambrick, Geletkanycz, and Fredrickson, 1993) and strategic distinctiveness from the other industry competitors (Crossland *et al.*, 2014).

We classify the literature on the relationship of CEOs and innovation into two broad categories. The first is governance of the CEO (from both internal and external sources) to generate optimal innovation. This literature includes the firm's choices such as CEO compensation incentives (Gormley et al. 2013, Lerner and Wulf 2007) and takeover provisions (Becker-Blease 2011) which affect innovation by altering the CEO's time horizon. Ownership and outside factors also affect the CEO's incentive to innovate including the presence of institutional ownership, firm liquidity, and analyst coverage (Aghion et al. 2013, Fang et al. 2014, Yu 2008). The second literature category is the impact of CEO heterogeneity on firm innovation which includes personal attributes such as age as well as personality differences such as overconfidence that alter the CEO's incentive to innovate (Dechow and Sloan 1991, Galasso and Simcoe 2011, Hirshleifer et al. 2012). We leverage both categories in our investigation of the innovation direction of outside versus inside CEOs. Firms must make a governance choice in the CEO selection process to hire a CEO with firm specific knowledge and skills versus a CEO with more general skills. The results likely have profound effects on innovation.

We generate four hypotheses that predict firms with outside CEOs will usher in more innovation change and the magnitude of this change will be negatively impacted by the size of the R&D organization, the decentralization of the R&D organization, and the distance of the CEO's prior experience. To explore our hypotheses we utilize detailed executive employment records across time from Execucomp to identify CEOs hired from another firm (outside CEOs) and CEOs hired from within the firm (inside CEOs). We then use USPTO patent data to construct a combined patent vector for each firm from the entire yearly stock of firm patents. The dynamics of that patent vector are compared on a yearly basis (the comparative angle of the two vectors) for firms with outside and inside CEOs and reveal a surprising basic finding that firms with outside CEOs have less dynamic patent vectors compared to firms with inside CEOs. This rejection of our main hypothesis holds when utilizing a propensity score matching model to account for observable factors that affect the choice of outside CEO. We strengthen this interesting result with circumstantial evidence analyzing organizational attributes that will lessen the ability of the CEO to affect innovation. In particular, we look at proxies for R&D structure (centralization of patent inventors in metropolitan areas) and R&D size (R&D expenditures) which act to alter innovation control. We find that the decreased dynamics associated with outside CEOs are magnified when the R&D structure is more centralized (more CEO control) and with smaller R&D size (less inertia). Importantly, this result that the relationship we find between outside CEOs and a lack of direction is decreased in decentralized and larger organizations lends support that the result is potentially CEO driven. We also find that more remote prior CEO experience leads to decreased innovation dynamics.

These surprising results have interesting implications for firms. In particular, firms that hire outside CEOs should fully take into account the potential consequences which may include not only a decrease in innovation but also the inability or lack of desire to take the firm's innovation in a new direction. However, we caveat these findings in three ways. First, although we utilize propensity score matching and firm structure to strengthen the circumstantial case, the relationship should not be interpreted as causal because the assignment of outside CEOs to firms is not random. Second, this chapter utilizes firm's patents to proxy innovation yet many firms do not patent and patenting likely does not account for the entirety of the firm's innovation. Third, we also caution firms from interpreting the results as a prescription not to hire outside CEOs. In fact, outside CEOs likely have valuable capabilities other than innovation for which the firm should consider.

This chapter makes a contribution to multiple literature streams including innovation and CEO heterogeneity by finding the surprising result that outside CEOs may decrease innovation change in firms. The chapter also makes a contribution to firm innovation governance literature by finding a new potential impact of CEO succession choice (innovation direction) and identifying organizational characteristics that act to intensify or mute that potential impact. The chapter proceeds as follows. First, the background of the relationship between CEO heterogeneity and innovation, outside CEOs, and organizational structure and innovation are developed and utilized to generate four hypotheses. Second, the empirical section explains methods and tests the hypotheses. Lastly, the implications and conclusions are developed.

3.2 Background and hypotheses development

There has been a renewed focus in the literature on the how managerial differences translate into firm performance differences (Bandiera et al. 2017, Bloom et al. 2013, Malmendier et al. 2011, Malmendier and Tate 2005). A key managerial task is leading the long-term firm strategy which depends on his/her mental map about where the best opportunities for profit are located (Nickerson and Zenger, 2004; Siggelkow and Rivkin, 2005) and on the information gathered and interpreted by the CEO about the world. One of the firm's fundamental tools available to achieve a desired long term objective in a competitive market is management of the R&D function. With R&D the firm can create a capability bundle that delivers an expected future market position. Therefore, it is natural to think that the CEO's characteristics and the direction of the innovation of a firm are intimately connected, even if the CEO is not directly dictating what research projects should be tackled next.

The broader connection between the CEO and firm innovation has received significant attention in the literature. These studies provide evidence that CEOs do affect the firm's innovation and that understanding differences in these CEOs and how they are governed is key to any firm's innovation strategy. First, the literature shows that changes in the CEO's incentive to innovate leads to changes in innovative output. The most direct evidence in this spirit is research that shows that CEO's with greater long term incentive compensation increase firm innovative output (Gormley, Matsa, and Milbourn, 2013; Lerner and Wulf, 2007; Makri, Lane, and Gomez-Mejia, 2006). Also in this same direction, research looking at CEO (Bange and De Bondt 1998) and institutional ownership (Aghion et al. 2013) also links ownership incentives with greater innovative performance. Other governance factors have been shown to change the long-term versus short-term outlook horizon for CEOs which has implications for innovation. Research finds that higher firm liquidity (Fang et al. 2014), weaker takeover provisions (Becker-Blease 2011), and higher analyst coverage (Yu 2008) all increase CEO short-termism which leads to a decrease in innovation due to its long-term nature. Taken together, the evidence indicates that governance of the CEO affects the firm's innovative performance which implicitly assumes that CEOs have significant influence over the firm's innovation.

The literature also has studied how heterogeneous personal CEO traits impact innovation. A simple example of differences across CEOs is the effect of age on firm innovation. As CEOs approach retirement age, their ability to gain from the firm's long-term innovation time horizon is decreased which should lead to a focus on more short-term activities. Research shows that this is the case in the absence of long-term incentives to mute the impact (Dechow and Sloan 1991). Heterogeneity in the CEO's perception of risk has also been shown to affect firm innovation. The lack of diversification of CEOs compared to shareholders (due to the stickiness of firm specific career investment) likely leads to a departure in the risk appetite between the two. However, overconfident CEOs that underestimate innovation risk are likely to increase firm innovation which provides more alignment with shareholders. Research using the CEO's stock option exercising behavior when the options are "in the money" (Malmendier and Tate 2005) indicates that this is the case (Galasso and Simcoe 2011, Hirshleifer et al. 2012). These research examples indicate that not only do CEOs have influence over firm innovation but that differences in these CEOs may also result in heterogeneous firm innovation outcomes.

The firm's CEO succession decision intersects both the firm's governance of innovation as well as the potential effect of CEO heterogeneity. Firms must make tradeoffs in the succession choice that spans across all aspects of the enterprise. One such tradeoff is the decision to hire a CEO from inside the organization versus broadening the pool of applicants to hire outside the organization. Inside CEO candidates provide firm and industry knowledge, contacts, and skills as well as increase motivation of future potential candidates inside the firm (Harris and Helfat 1997, Howard 2001). Inside candidates also decrease information asymmetry between the potential CEO and firm (Zajac 1990) which can lead to less turbulence (Zhang and

Rajagopalan 2004). Conversely, firms look outside the organization to increase the applicant pool and acquire knowledge from outside the firm and industry (Virany et al. 1992, Zhang 2008, Zhang and Rajagopalan 2010). These tradeoffs indicate that the choice of inside versus outside CEO is likely not universally beneficial for firm performance which is corroborated by the empirical record (Lubatkin et al. 1986, Shen and Cannella 2002, Worrell and Davidson III 1987, Zajac 1990, Zhang 2008).

The association of inside versus outside CEOs to innovation, a key dimension of long-term performance, has been further investigated in the literature to find that outside CEOs are associated with decreased patenting activity (Balsmeier and Buchwald 2014) as well as a decrease in R&D productivity (Cummings and Knott 2018) due to the lack of firm and industry specific technological expertise. Balsmeier and Buchwald (2014) utilize a small sample of German firms in the early 2000s to analyze the patenting implications associated with this decision and find that firms with outside CEOs patent less and with less impact. Cummings and Knott (2018) utilize a broader sample of US firms and find that outside CEOs are associated with degraded RQ (R&D productivity) and this degradation increases as the CEO's prior experience is further away from the current firm and the more reliant the firm is on R&D. Both studies indicate that outside CEOs may be detrimental to firm innovation but we do not know the direction outside CEOs take innovation.

The extant literature reveals that hiring an inside versus an outside CEO may have a profound impact on innovation. Indeed, the literature documents the importance of hiring external CEOs during organizational change (Virany, Tushman, and Romanelli, 1992; Zhang and Rajagopalan, 2010). Outside CEOs have a different professional experience which means that external hires have different knowledge sets, professional networks, and life experiences that change the way those professionals interpret and process information. Additionally, CEOs with more career variety tend to exhibit higher strategic dynamism (Crossland *et al.*, 2014). Those differences should help an outside CEO spot complementarities not seen by internal members of a firm and act upon those new discoveries, inducing firms that hire outside CEOs to explore more. More exploration should manifest as bigger changes in innovation direction which leads to the first hypothesis.

Hypothesis 1: Firms with Outside CEOs should exhibit more changes in innovation direction than firms with inside CEOs.

We next explore aspects of organizations that may act to govern the CEOs influence over innovation. Given that outside CEOs seem to have a negative impact on firm innovation but may bring benefits to the firm in other areas, it may be in the interest of the firm to understand ways in which to maximize the potential positive aspects while minimizing the negative aspects. One potential impact of the CEOs influence on firm innovation is R&D structure. Research has investigated the potential impact of the firm's centralized versus decentralized R&D structure decision (Argyres and Silverman 2004, Arora et al. 2011, 2014, Knott 2017, Leiponen and Helfat 2011, Lerner and Wulf 2007). Centralized organizations have two key distinctions from decentralized organizations. First, centralized R&D organizations have a single group that handles the entire enterprise's R&D and reports directly to the CEO while decentralized R&D organizations utilize R&D silos such that each business unit has its own R&D and those organizations report to the head of that business unit. Second, centralized R&D budgets are allocated from the corporate office while decentralized R&D budgets are allocated at the business unit level (Argyres and Silverman 2004). The R&D structure not only has implications for the basic nature of technology produced, spillovers across business units, and sourcing of the technology (Knott 2017) but also for the control of the CEO over innovation. CEOs have much more control in a centralized R&D structure due to direct oversight of the organization and control of budgets. This is empirically demonstrated by Lerner and Wulf (2007) by showing CEO incentive compensation only has a significant effect on innovation if the firm's R&D structure is centralized. These insights lead to our second hypothesis:

Hypothesis 2: The impact of outside CEOs in directing innovation should decrease in magnitude with R&D decentralization.

Outside CEOs also come with heterogeneous prior experience which may impact their new firm's innovation direction (Harris and Helfat 1997). The applicability of prior experience to help spot innovation complementarities that lead to a shift in innovation direction may differ with the relation of the CEOs prior firm experience to the new firm. For example, CEOs that have outside experience within the same industry may have accumulated the technological expertise to recognize and capitalize on innovative opportunities. Conversely, CEOs that are hired with more remote technological experience may lack the expertise to lead the firm in a new technological direction but may make changes to other aspects of the firm (e.g. marketing, manufacturing, etc.) that conform more to the general skills that they bring. We expect this heterogeneous experience to translate into innovation direction in accordance with Hypothesis 3:

Hypothesis 3: The impact of outside CEOs in directing innovation should decrease as prior experience becomes more remote from the current firm.

Other aspects of the firm's innovation are also likely to affect the CEO's ability to control the firm's innovation direction and not all firms are suited for a change in innovation directed by outside CEOs. Firms that produce a lot of innovation usually have a stable organizational structure that manages that innovation. In those firms, CEOs have little leeway to change the innovation direction in an impactful way for four reasons. First, those tend to be large firms, likely with Chief Technology Officers and other employees in charge of the R&D. Those middlemen can lessen the impact of change in direction intended by the CEO. Second, firms with many patents usually have many projects in different stages of maturation in the pipeline. This increases the inertia of the firm, increasing the difficulty of shifting innovation towards a new path. Third, outside CEOs can have difficulties in gaining support in an environment in which they do not have enough power to make changes by themselves. One of the advantages of an internal hire is that inside CEOs generate hope and loyalty for some internal employees (Howard, 2001). This lack of loyalty makes it harder for external CEOs to implement changes. This problem is compounded in large firms, since CEO power tends to be less concentrated. In large firms, CEOs need to convince more people to implement changes. Finally, internal CEOs already have in place a dense knowledge network (Harris and Helfat, 1997) and have less information asymmetry than an outside hire (Zajac, 1990). This advantage is bigger in large firms, since there is less to learn when firms are smaller. Hypothesis 4 summarizes this discussion.

Hypothesis 4: The impact of outside CEOs in directing innovation direction should decrease in magnitude with the amount of R&D produced in the firm.

Next, we use data on CEOs and patents to test these hypotheses.

3.3 Data and methods

3.3.1 Empirical Model

The main objective of the chapter is to understand how outside CEOs influence the direction of the firm's innovation. To do that, we rely on the following specification:

$$(\Delta \textit{ in innovation direction})_{i,\, t+1} = \beta_1 \,(\textit{Outside CEO})_{i,\, t+1} + \beta_2 \,(\textit{Controls})_{i,t} + \eta_i + \lambda_t + \varepsilon_{i,\, t} \qquad (1)$$

The model in (1) uses firm fixed effects (η) and year fixed effects (λ) alongside a set of control variables to tease out the relationship between hiring an outside CEO in year t and the change in the direction of innovation in year t+1. To calculate the direction of innovation, we used the patent's main technology classification. Each patent has a main technology classification that indicates the technical content of the innovation claimed by the patent. Thus, by looking at how the technical mix of patents changes year by year, we hope to get insights about the firm's direction of innovation. To operationalize that change, we aggregated the number of patent applications in a year for each technology classification for every firm. This operation generates a vector for each firm-year that has dimensions equal to the number of possible technical classifications. Each entry in this vector corresponds to how many patent applications in that specific technology classification the firm has made in a particular year.

To calculate differences in innovation direction, we calculated one minus the cosine distance to indicate how far away two innovation portfolios are from one another. The further away are the vectors, the higher is the technological distance between them. The formula to calculate the distance between the vectors A and B in an N-dimensional technology classification is described by equation (2) below.

$$\text{Vector distance} = 1 - \frac{\sum_{i=1}^{N} A_i B_i}{\sqrt{\sum_{i=1}^{N} A_i^2}\sqrt{\sum_{i=1}^{N} B_i^2}} \quad (2)$$

Figure 3.1 illustrates the idea captured by equation (2). In that representation, there are only two technological classifications. In year T the firm applied for three patents with technology classification 1 and one patent with technology classification 2, resulting in the vector (3,1). In the next year, the firm applied for three patents with technology classification 1 and three patent with technology classification 2, that is, vector (3,3). The distance between the vectors is 1 minus the cosine of angle α, approximately 0.1. If the firm had a vector of (3,2) in year T+1 instead, the measure would have decreased to 0.03.

Since all vectors are non-negative, the distance as calculated above yields a number between 0 (identical vectors) and 1 (orthogonal vectors). The benefit of using a measure based on the cosine distance is that the scale of patenting does not impact the measure since the distance between a vector and n times that same vector is zero. We calculate not only the change in vector direction within the same firm over time, but also the distance between a firm's vector and the average vector composed of the other firms in the industry.

To measure the size of the innovation vector itself, we use the Euclidean norm. This is more akin to the amount of innovation taking place in the firm, instead of the overall direction as in the case of the cosine measure. With the cosine distance and the norm, we have a complete picture of the innovation change between the two time periods.

3.3.2 Data

We used three main sources of data: Execucomp, Compustat, and the USPTO database. Execucomp data allows us to identify up to 5 top executives in each firm. Compustat provides firm information, and the USPTO contains patent information. The USPTO database was matched with Compustat using fuzzy string matching of company name and patent application assignee. Execucomp is matched with Compustat using the Standard & Poor's firm identification. In total, the dataset spans from 1992 to 2017.

Variables

The main independent variable is a dummy *Outside CEO* that is equal to one if the CEO was in a different firm two years prior his/her CEO appointment[19]. To identify previous appointments, we classified CEOs as an outsider if they show up in the Execucomp dataset as working for a different firm two years before their appointment. To that, we added the outside CEOs manually compiled by Cummings and Knott (2018). In the end, we have 309 outside CEOs and 1616 CEOs hired from within the firm. This dataset has a proportion of outside CEOs around 16%, lower than similar calculations in the literature. There are two reasons for that. First, many firms were lost in the name matching process. Although there is selection in that sample, it is hard to imagine why it would bias the results. Second, we only look at firms that patent. This is a known limitation intrinsic to all studies using patenting.

As control variables, we first consider the R&D intensity of the firm. This allows us to control for the amount of resources available to the firm employed in innovation. To control for the total

[19] The reason for the two year lag is to take care of CEOs hired from outside to groom them for CEO, i.e. the "heir apparent" problem (Harris and Helfat, 1997).

amount of available resources, we added revenue growth of the prior two years. We control for CEO incentives using salary. Additionally, the regressions control for the distance between the firm current technological vector and the vector composed by the average firm in the firm's industry. This control the uniqueness of the firm innovation portfolio in relation to its competition. Finally, we control for the number of patent applications and the average number of claims each patent application makes. This controls for the firm's propensity to patent and how broad the patents tend to be.

As mentioned before, the independent variables are measures over the firm's technological vector. We used the Cooperative Patent Classification technology classification at the subclass level, yielding 638 possible technological classifications. Although there are finer classifications, using them would provide additional insight into innovation change. Since the measure used here gives equal weight for all differences in technology, we need the different classifications to have a meaningful impact on innovation direction.

We also utilize two measures of firm R&D structure based on the geographic location of the firm's patent inventors which proxy for the control that the CEO exerts over the firm's R&D organization (Cummings 2018). We use USPTO patent data to identify the home location of each individual inventor of patents assigned to a firm. The first R&D structure measure is the proportion of the firm's inventors that are located in the most frequent census statistical area (CSA). Thus, the proxy measures the geographic agglomeration of the firm's inventors. The second R&D structure measure takes into account the potential that the geographic agglomeration may be spread over multiple CSAs by calculating a Herfindahl index of the concentration of a firm's inventors in the five most frequent CSAs. The geographic agglomeration of R&D increases with each measure. Prior research has utilized firm inventor geographic agglomeration as a measure of informal R&D structure (Singh 2008) but this chapter also utilizes the proxy as a measure of formal R&D structure which is supported by its robust correlation with measures for patent type (Leiponen and Helfat 2011) and impact (Singh 2008) that are consistent with formal R&D structure (Argyres and Silverman 2004, Knott 2017).

The dependent variable, *Vector change (prior year),* is the vector distance between year t+1 and t, whereas *Norm change* measures the change in the norm. We also discuss the change in the firm's technology direction in relation to the industry average vector in year t, represented by *Vector dist., (mean ind.)*. Finally, we deleted observations that had R&D intensity and revenue growth over two years that had values higher than 3 times the 99^{th} percentile of each variable. This was done to get rid of extreme outliers in the data. The appendix contains the regressions using the full dataset. The main results remain similar, and any relevant discrepancy is addressed in the text. The summary statistics for the restricted sample is displayed in Table 3.1.

3.4 Results

The result of estimations of equation (1) are shown in Table 3.2. The table presents the impact of Outside CEOs on the change in innovation direction. The first column is the baseline, without the addition of the Outside CEO dummy. The distance between the innovation vector of the firm and the mean vector of the industry is positive and significant. This suggests that firms that have a research portfolio different than its peers in the industry tend to shift the direction of innovation more. If we interpret more change in the innovation vector as more exploration, this means that more unique firms tend to explore more than firms that are similar to the industry average.

Similar ideas were shown in the literature, linking technological diversity of alliance partners and exploratory innovation (Phelps 2010).

Columns (2) to (4) include the impact of Outside CEOs on change of innovation direction. The models show the influence of firm and year fixed effects on the Outside CEO coefficient. Column (4) is the main specification used to test hypothesis 1, since it includes both firm and year effects, as well as clustered standard errors at the firm level. This regression indicates that a firm headed by an outside CEO tend to shift innovation direction less that if it were managed by an inside CEO. Comparing to the mean change in year-to-year innovation direction, having an outside CEO decreases change in direction by a little more than 10%. Columns (5) to (8) show that this result is robust to inclusions of yearly revenue growth and measures of centralization to control for firm organizational structure. These regressions strongly reject hypothesis 1 which was the expectation that outside CEOs would generate more innovation change.

Table 3.3 analyzes the impact of outside CEOs on the innovation vector's norm (1) and on the distance between the firm and industry (2). Although outside CEOs tends to decrease the change in innovation direction, they seem to increase the vector norm. Outside CEOs tend to manage firms that have a more focused and deeper innovation portfolio. Column (2) shows that they are also not inclined to deviate from the average innovation direction of the industry. One possible explanation for mimicking other firms in the industry is the lack of expertise of outside CEOs (Cummings and Knott 2018).

Hypothesis 2 argues that if we believe hiring an outsider is creating the effects shown, we would expect that those effects are stronger the more power the CEO has to change the direction of innovation. To test that, we calculated the median of two centralization measures (Cummings

2018) and analyzed the impact of outside CEOs in the two subsamples: firms with a high degree of centralization (higher than the median) and firms with a low degree of centralization (lower than the median). The regressions in Table 3.4 use two measures of centralization to test hypothesis 2. Columns (1) and (2) split the sample using the yearly Herfindahl index of the firm's patent inventors in the firm's five most frequent metropolitan areas. Columns (3) and (4) use the concentration of the firm's patent inventors in the single most frequent metropolitan area. For both of those measures, centralization increases with the value of the measure. The coefficient for outside CEO is negative and significant for all high centralization subsamples. For low centralization, outside CEO is not precisely estimated and has higher coefficients. Thus, we find strong support for hypothesis 2.

The same analysis is performed for the innovation vector norm and distance to the industry. Tables 3.5 and 3.6 show that the impact of outside CEOs in those dependent variables is much less clear, since most of the coefficients for Outside CEO are not as precisely estimated. The increase in norm is more pronounced in firms with high centralization if we use the Herfindal index and concentration of inventors in metropolitan areas. However, the same regression using the full sample show conflicting results, as evidenced by Tables A2.5 and A2.6.

So far, we have presented evidence supporting the claim that outside CEOs decrease the change in direction of innovation. If that is indeed the case, we should expect that the more "outside" the CEO is, the higher is going to be its impact. In other words, the further away the firm is that the CEO comes from, the lower is the CEO expertise, likely magnifying the innovation direction impact. Table 3.7 splits the sample of CEOs into two: CEOs coming from a firm further away than the median distance between current and previous firm, and CEOs coming from firms nearer. CEOs from more distant firms tend to decrease the direction of innovation change more than CEOs coming from firms with innovation portfolios similar to the current firm. This finding corroborates hypothesis 3.

Again, the same analysis is performed for norm and distance to the industry. Tables 3.8 and 3.9 show the results. The analysis of the norm change does not show a significant coefficient for outsiders. Compatible with the idea of lack of expertise, CEOs from distant firms tend to work for firms that get closer to the industry's average innovation vector.

Finally, hypothesis 4 speculates about the impact of outside CEOs on firms with different R&D intensities. Table 3.10 splits the sample into firms with high and low R&D intensity in relation to the median R&D intensity. Regression (2) estimates that the reduction in the change of innovation direction comes from firms with lower R&D intensity, lending empirical support to hypothesis 4. The same happens with the effect of outside CEOs on the norm and distance to industry's innovation vector: only the coefficients for firms with low R&D intensity are significant. It is important to notice that the regressions using the full sample do not support the higher impact of outside CEOs in firms with low R&D intensity, so caution should be exerted.

3.5 Robustness Tests

The analysis of the impact of outside CEO can be plagued by endogeneity issues. More specifically, firms do not choose CEOs randomly. This opens the door for unobserved characteristics to influence both the direction of innovation and the choice of outside CEO. To help assuage the issue, we ran a propensity score matching model. Following the example of Cummings and Knott (2018), we used the log of revenues minus cost of goods sold to proxy for performance, firm growth in two years to indicate internal resources, revenues to control for lack of internal candidates, and R&D spending to control for the need of expertise. These results are reported in Table 3.11. Although the average treatment effect is negative and significant, we do not think that all the possible sources of endogeneity are dealt with. Unfortunately, the data does not offer a good instrument to help mitigate the problem.

Finally, we run a placebo test by calculating the percentage of outside CEOs in the data and generating 100 samples randomly assigning the outside CEO indication. For each sample, we run the same regression as shown in Table 3.2, column (4). Figure 3.2 shows the coefficients for Outside CEO and the 95% confidence intervals for the placebo tests and the data. This figure shows that the association between external CEO and the vector distance is not occurring because of some mechanical property of the way we measure innovation change.

3.6 Conclusions and implications

In this chapter we sought to further understand the potential impact on the dynamics of innovation for firms with outside CEOs. We utilized Execucomp to determine each CEO's prior experience (outside or inside the firm) and USPTO patent data to construct a multi-dimensional vector based on the firm's patent characteristics. Contrary to expectations, our results indicate that firms with outside CEOs have less dynamic innovation in comparison to firms with inside CEOs. More specifically, we find that the angle of the firm's patent vector changes less when the firm's CEO is from outside. These results are strengthened with the use of a propensity score matching model as well as with supplementary results finding a larger relationship magnitude with subsamples that proxy for centralized R&D, more remote prior experience, and larger R&D organizations. In particular, the results associated with R&D structure and R&D size provide circumstantial support that the effect may be CEO driven. When the weight of the results are taken in aggregate, there are implications for managerial dynamic capabilities with respect to innovation (Helfat and Martin 2015). However, we caution interpreting the results as causal for multiple reasons including the lack of random assignment of CEOs to firms and the imperfect nature of patents as a proxy for innovation.

Although the main findings come as a surprise, the results may conform to prior literature that indicates outside CEOs lack the technological expertise to make optimal strategic innovation decisions (Balsmeier and Buchwald 2014, Cummings and Knott 2018). More specifically, Cummings and Knott (2018) find that firms with outside CEOs respond less to CEO incentive compensation which could be interpreted to signal that they do not know what project to fund. The results of the current chapter may also indicate that outside CEOs do not know where to invest R&D which leads to less dynamic firm innovation. We also caution against interpreting the results to indicate prior literature is incorrect in stating that outside CEOs are brought in to usher strategic organizational change. However, this change may occur in expertise areas of the new CEO other than innovation. Outside CEOs bring with them more general skills (Murphy and Zabojnik 2007) which may translate into strategic change in marketing, sales, manufacturing, etc., rather than innovation.

This chapter has multiple implication layers for firms. First, the results indicate that innovation strategy may be affected by the choice of outside CEOs. However, overall performance takes into account many other factors for which outside CEOs may excel; therefore, firms should generate CEO succession decisions based on the potential effect on innovation as well as these other factors. Second, the results indicate that firms may be able to minimize the innovation impact of outside CEOs with complementary R&D structure decisions while taking advantage of other positive aspects outside CEOs bring to the table. We speculate that an R&D structure which lessens the CEO's control over innovation may be advantageous for firms with outside CEOs.

3.7 References

Aghion P, Van Reenen J, Zingales L (2013) Innovation and Institutional Ownership. *Am. Econ. Rev.* 103(1):277–304.

Argyres NS, Silverman BS (2004) R&D, organization structure, and the development of corporate technological knowledge. *Strateg. Manag. J.* 25(8–9):929–958.

Arora A, Belenzon S, Rios LA (2011) *The organization of R&D in American corporations: the determinants and consequences of decentralization* (National Bureau of Economic Research).

Arora A, Belenzon S, Rios LA (2014) Make, buy, organize: The interplay between research, external knowledge, and firm structure. *Strateg. Manag. J.* 35(3):317–337.

Balsmeier B, Buchwald A (2014) Who promotes more innovations? Inside versus outside hired CEOs. *Ind. Corp. Change* 24(5):1013–1045.

Bandiera O, Hansen S, Prat A, Sadun R (2017) *CEO Behavior and Firm Performance* (National Bureau of Economic Research).

Bange MM, De Bondt WF (1998) R&D budgets and corporate earnings targets. *J. Corp. Finance* 4(2):153–184.

Becker-Blease JR (2011) Governance and innovation. *J. Corp. Finance* 17(4):947–958.

Bloom N, Eifert B, Mahajan A, McKenzie D, Roberts J (2013) Does Management Matter? Evidence from India. *Q. J. Econ.* 128(1):1–51.

Crossland C, Jinyong Zyung, Hiller NJ, Hambrick DC (2014) Ceo Career Variety: Effects on Firm-Level Strategic and Social Novelty. *Acad. Manage. J.* 57(3):652–674.

Cummings T (2018) Innovation Strategy and Structure: R&D Organization, CEO Characteristics, and Institutional Ownership.

Cummings T, Knott AM (2018) Outside CEOs and Innovation. *Forthcom. Strateg. Manag. J.*

Dechow PM, Sloan RG (1991) Executive incentives and the horizon problem: An empirical investigation. *J. Account. Econ.* 14(1):51–89.

Fang VW, Tian X, Tice S (2014) Does stock liquidity enhance or impede firm innovation? *J. Finance* 69(5):2085–2125.

Galasso A, Simcoe TS (2011) CEO overconfidence and innovation. *Manag. Sci.* 57(8):1469–1484.

Gormley TA, Matsa DA, Milbourn T (2013) CEO compensation and corporate risk: Evidence from a natural experiment. *J. Account. Econ.* 56(2):79–101.

Harris D, Helfat C (1997) Specificity of CEO human capital and compensation. *Strateg. Manag. J.* 18(11):895–920.

Helfat CE, Martin JA (2015) Dynamic managerial capabilities: Review and assessment of managerial impact on strategic change. *J. Manag.* 41(5):1281–1312.

Hirshleifer D, Low A, Teoh SH (2012) Are overconfident CEOs better innovators? *J. Finance* 67(4):1457–1498.

Howard A (2001) Identifying, assessing, and selecting senior leaders. *Nat. Organ. Leadersh. Underst. Perform. Imperatives Confronting Todays Lead.*:305–346.

Knott AM (2017) *How Innovation Really Works: Using the Trillion-Dollar R&D Fix to Drive Growth: Using the Trillion-Dollar R&D Fix to Drive Growth* (McGraw Hill Professional).

Leiponen A, Helfat CE (2011) Location, decentralization, and knowledge sources for innovation. *Organ. Sci.* 22(3):641–658.

Lerner J, Wulf J (2007) Innovation and incentives: Evidence from corporate R&D. *Rev. Econ. Stat.* 89(4):634–644.

Lubatkin M, Chung K, Rogers RC, Owers JE (1986) The Effect of Executive Succession on Stockholders Wealth of Large Corporations. *Acad. Manag. Proc.* (Academy of Management), 36–40.

Makri M, Lane PJ, Gomez-Mejia LR (2006) CEO incentives, innovation, and performance in technology-intensive firms: a reconciliation of outcome and behavior-based incentive schemes. *Strateg. Manag. J.* 27(11):1057–1080.

Malmendier U, Tate G (2005) CEO overconfidence and corporate investment. *J. Finance* 60(6):2661–2700.

Malmendier U, Tate G, Yan J (2011) Overconfidence and early-life experiences: the effect of managerial traits on corporate financial policies. *J. Finance* 66(5):1687–1733.

Murphy KJ, Zabojnik J (2007) Managerial capital and the market for CEOs. *Available SSRN 984376*.

Nickerson JA, Zenger TR (2004) A Knowledge-Based Theory of the Firm: The Problem-Solving Perspective. *Organ. Sci.* 15(6):617–632.

Phelps CC (2010) A Longitudinal Study of the Influence of Alliance Network Structure and Composition on Firm Exploratory Innovation. *Acad. Manage. J.* 53(4):890–913.

Romer PM (1990) Endogenous Technological Change. *J. Polit. Econ.* 98.

Shen W, Cannella AA (2002) Revisiting the performance consequences of CEO succession: The impacts of successor type, postsuccession senior executive turnover, and departing CEO tenure. *Acad. Manage. J.* 45(4):717–733.

Siggelkow N, Rivkin JW (2005) Speed and Search: Designing Organizations for Turbulence and Complexity. *Organ. Sci.* 16(2):101–122.

Singh J (2008) Distributed R&D, cross-regional knowledge integration and quality of innovative output. *Res. Policy* 37(1):77–96.

Solow RM (1957) Technical change and the aggregate production function. *Rev. Econ. Stat.*:312–320.

Virany B, Tushman ML, Romanelli E (1992) Executive succession and organization outcomes in turbulent environments: An organization learning approach. *Organ. Sci.* 3(1):72–91.

Worrell DL, Davidson III WN (1987) The effect of CEO succession on stockholder wealth in large firms following the death of the predecessor. *J. Manag.* 13(3):509–515.

Yu FF (2008) Analyst coverage and earnings management. *J. Financ. Econ.* 88(2):245–271.

Zajac EJ (1990) CEO selection, succession, compensation and firm performance: A theoretical integration and empirical analysis. *Strateg. Manag. J.* 11(3):217–230.

Zhang Y (2008) Information asymmetry and the dismissal of newly appointed CEOs: An empirical investigation. *Strateg. Manag. J.* 29(8):859–872.

Zhang Y, Rajagopalan N (2004) When the known devil is better than an unknown god: An empirical study of the antecedents and consequences of relay CEO successions. *Acad. Manage. J.* 47(4):483–500.

Zhang Y, Rajagopalan N (2010) Once an outsider, always an outsider? CEO origin, strategic change, and firm performance. *Strateg. Manag. J.* 31(3):334–346.

3.8 Figures and Tables

Figure 3.1: Change in innovation direction

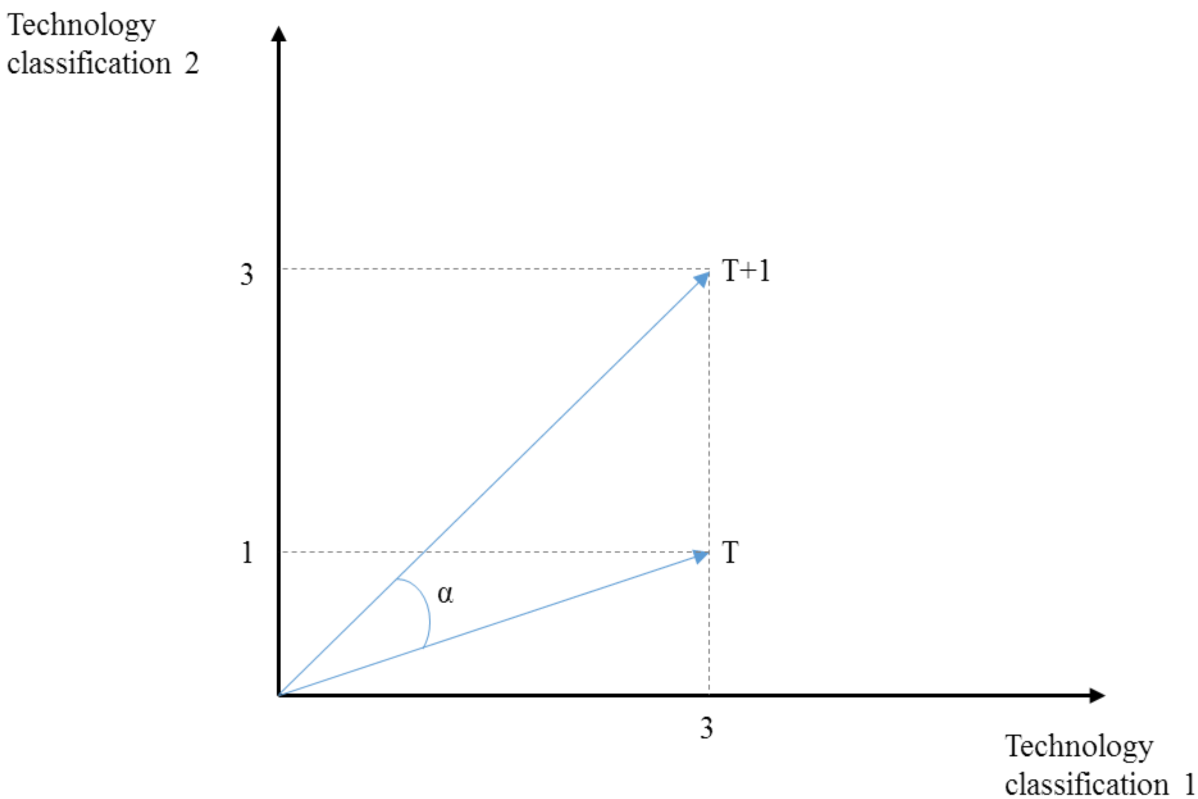


Figure 3.2: Placebo test

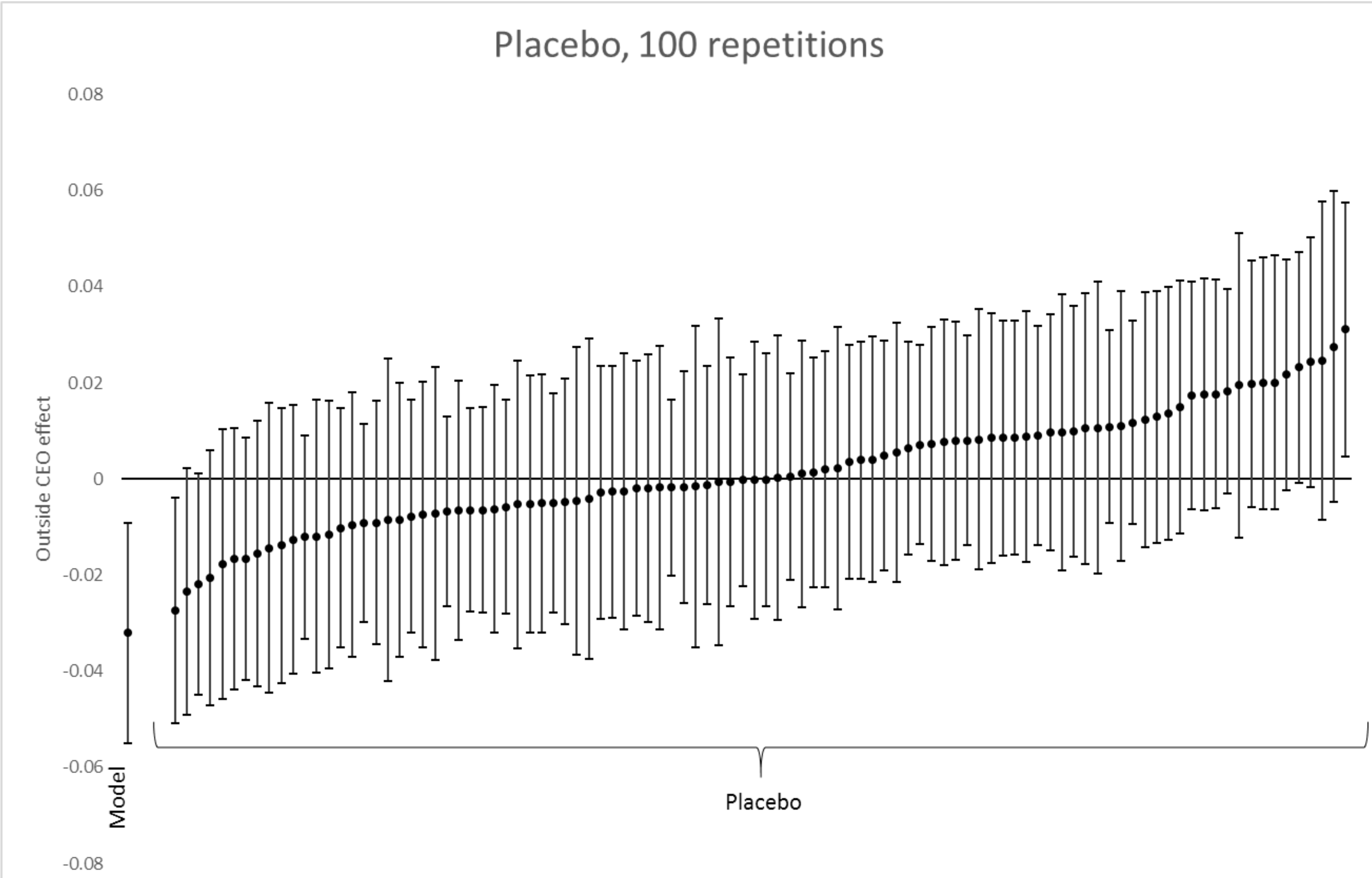


Table 3.1: Summary statistics

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Outside CEO	7,029	0.142	0.349	0	1
Vector change, (prior year)	6,883	0.290	0.302	0	1
Norm change, (prior year)	7,029	0.0633	59.17	-2,710	2,112
Vector change, (mean ind.)	7,029	0.403	0.289	0	1.000
Growth, (2 yrs)	7,029	0.371	1.014	-0.941	18.82
salary	7,029	680.0	409.8	0	5,500
Patent applications	7,029	119.6	486.8	1	12,271
Avg. num. claims	7,029	19.69	9.101	0	198
R&D intensity	7,029	0.155	0.403	0	7.158

Table 3.2: Outside CEO decrease change in innovation direction

VARIABLES	(1) Vector change (prior year)	(2) Vector change (prior year)	(3) Vector change (prior year)	(4) Vector change (prior year)	(5) Vector change (prior year)	(6) Vector change (prior year)	(7) Vector change (prior year)
R&D intensity	-0.0166*	-0.0164	-0.0485***	-0.0167*	-0.0177*	-0.0113	-0.0116
	(0.00997)	(0.01000)	(0.0115)	(0.0101)	(0.00993)	(0.00991)	(0.00990)
salary	5.75e-06	1.23e-05	1.15e-05	4.94e-06	4.85e-06	1.75e-05	1.76e-05
	(1.40e-05)	(1.24e-05)	(1.71e-05)	(1.36e-05)	(1.36e-05)	(1.40e-05)	(1.40e-05)
Avg. num. claims	-0.000682	-0.000761*	-0.00104*	-0.000695	-0.000692	-0.000528	-0.000530
	(0.000469)	(0.000438)	(0.000592)	(0.000469)	(0.000470)	(0.000477)	(0.000477)
Vector change, (mean ind.)	0.489***	0.490***	0.480***	0.487***	0.488***	0.482***	0.485***
	(0.0289)	(0.0282)	(0.0250)	(0.0286)	(0.0286)	(0.0319)	(0.0318)
Patent applications	-1.76e-05	-2.35e-05*	-5.19e-05*	-2.00e-05*	-1.99e-05*	-1.80e-05	-1.81e-05
	(1.12e-05)	(1.26e-05)	(2.65e-05)	(1.10e-05)	(1.10e-05)	(1.11e-05)	(1.11e-05)
Growth, (2 yrs)	-0.00348	-0.00448	-0.0148***	-0.00362	-0.00176	-0.00118	-0.00108
	(0.00292)	(0.00296)	(0.00383)	(0.00293)	(0.00308)	(0.00324)	(0.00325)
Growth, (prior yr)					-0.00742	-0.00404	-0.00405
					(0.00527)	(0.00590)	(0.00593)
Outside CEO		-0.0329***	-0.0406**	-0.0321***	-0.0320***	-0.0305**	-0.0304**
		(0.0114)	(0.0163)	(0.0116)	(0.0116)	(0.0119)	(0.0120)
Herfindal Centralization						0.0336	
						(0.0206)	
CSA Concentration							0.0307
							(0.0207)
Observations	5,767	5,767	5,833	5,767	5,767	5,096	5,096
R-squared	0.623	0.621	0.281	0.623	0.623	0.629	0.629
Firm FE	YES	YES	NO	YES	YES	YES	YES
Year FE	YES	NO	YES	YES	YES	YES	YES
Num. firms	570	570	636	570	570	481	481

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Outside CEO impact on norm and firm uniqueness

VARIABLES	(1) Norm change (prior year)	(2) Vector change (mean ind.)
R&D intensity	-1.488**	0.00288
	(0.669)	(0.00941)
salary	-0.00291	7.08e-07
	(0.0101)	(1.04e-05)
Avg. num. claims	0.0882**	-0.000360
	(0.0374)	(0.000379)
Vector change, (mean ind.)	4.981**	0.305***
	(2.180)	(0.0328)
Patent applications		-3.69e-05**
		(1.69e-05)
Growth, (2 yrs)	-0.0276	-0.00133
	(0.502)	(0.00217)
Outside CEO	9.845**	-0.0185*
	(4.566)	(0.0108)
Observations	5,767	5,767
R-squared	0.026	0.758
Firm FE	YES	YES
Year FE	YES	YES
Num. firms	570	570

Note: Standard errors clustered by firms in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3.4: CEO impact on innovation direction is higher in centralized firms

VARIABLES	(1) Vector change (prior year)	(2) Vector change (prior year)	(3) Vector change (prior year)	(4) Vector change (prior year)
R&D intensity	-0.00780	-0.0276**	-0.00780	-0.0259**
	(0.0145)	(0.0119)	(0.0142)	(0.0118)
salary	1.08e-06	1.65e-05	-2.78e-06	1.87e-05
	(2.13e-05)	(1.91e-05)	(1.96e-05)	(1.97e-05)
Avg. num. claims	-0.000666	-0.000787	-0.000856	-0.000596
	(0.000659)	(0.000661)	(0.000682)	(0.000633)
Vector change, (mean ind.)	0.521***	0.444***	0.495***	0.476***
	(0.0398)	(0.0391)	(0.0429)	(0.0375)
Patent applications	-1.57e-05	-1.69e-05	-1.22e-05	-2.21e-05
	(1.98e-05)	(1.30e-05)	(1.47e-05)	(1.67e-05)
Growth, (2 yrs)	-0.00774*	0.00150	-0.00629*	0.00106
	(0.00408)	(0.00450)	(0.00329)	(0.00557)
Outside CEO	-0.0600***	-0.00659	-0.0462***	-0.0175
	(0.0182)	(0.0140)	(0.0171)	(0.0153)
Observations	2,588	3,179	2,695	3,072
R-squared	0.619	0.621	0.640	0.611
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Centralization	High Herfindal Centralization	Low Herfindal Centralization	High CSA Concentration	Low CSA Concentration
Num. firms	303	267	302	268

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Outside CEO and norm change, centralized firms

VARIABLES	(1) Norm change (prior year)	(2) Norm change (prior year)	(3) Norm change (prior year)	(4) Norm change (prior year)
R&D intensity	-1.710*	-1.226	-2.508**	-0.647
	(0.888)	(0.960)	(1.100)	(0.753)
salary	-0.0179	0.0104	-0.0151	0.00948
	(0.0157)	(0.0127)	(0.0142)	(0.0133)
Avg. num. claims	0.0744	0.101	0.0848	0.0840
	(0.0483)	(0.0646)	(0.0527)	(0.0548)
Vector change, (mean ind.)	6.544**	3.326	6.125	3.795
	(3.245)	(2.891)	(3.783)	(2.339)
Growth, (2 yrs)	0.356	-0.0711	0.572	-0.647
	(0.345)	(0.716)	(0.441)	(0.722)
Outside CEO	12.48*	8.454	13.32**	7.914
	(6.719)	(6.448)	(6.189)	(6.889)
Observations	2,588	3,179	2,695	3,072
R-squared	0.090	0.024	0.092	0.021
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Centralization	High Herfindal Centralization	Low Herfindal Centralization	High CSA Concentration	Low CSA Concentration
Num. firms	303	267	302	268

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.6: Outside CEO and distance from industry, centralized firms

VARIABLES	(1) Vector change (mean ind.)	(2) Vector change (mean ind.)	(3) Vector change (mean ind.)	(4) Vector change (mean ind.)
R&D intensity	0.0136	-0.00493	0.0137	-0.00535
	(0.0121)	(0.0162)	(0.0121)	(0.0163)
salary	2.11e-06	3.17e-06	-9.48e-07	5.76e-06
	(1.46e-05)	(1.50e-05)	(1.30e-05)	(1.61e-05)
Avg. num. claims	-0.000193	-0.000609	-0.000354	-0.000407
	(0.000534)	(0.000549)	(0.000550)	(0.000516)
Vector change, (mean ind.)	0.232***	0.378***	0.226***	0.372***
	(0.0457)	(0.0454)	(0.0482)	(0.0434)
Patent applications	-6.70e-05***	-2.49e-05*	-5.14e-05***	-2.85e-05
	(1.64e-05)	(1.48e-05)	(1.74e-05)	(1.86e-05)
Growth, (2 yrs)	-0.000419	-0.00187	0.000136	-0.00327
	(0.00244)	(0.00357)	(0.00232)	(0.00400)
Outside CEO	-0.0302	-0.00764	-0.0311	-0.00590
	(0.0208)	(0.0107)	(0.0192)	(0.0120)
Observations	2,588	3,179	2,695	3,072
R-squared	0.735	0.776	0.755	0.765
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Centralization	High Herfindal Centralization	Low Herfindal Centralization	High CSA Concentration	Low CSA Concentration
Num. firms	303	267	302	268

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.7: CEO impact on innovation direction is higher when they come from more distant firms

VARIABLES	(1) Vector change (prior year)	(2) Vector change (prior year)
R&D intensity	-0.0178*	-0.0387
	(0.0103)	(0.0369)
salary	1.40e-05	7.61e-07
	(1.82e-05)	(1.61e-05)
Avg. num. claims	-0.000776	-0.000599
	(0.000531)	(0.000888)
Vector change, (mean ind.)	0.488***	0.414***
	(0.0343)	(0.0618)
Patent applications	-5.00e-05*	-8.55e-06
	(2.90e-05)	(6.00e-06)
Growth, (2 yrs)	-0.00455	-0.00630
	(0.00305)	(0.0131)
Outside CEO	-0.0385**	0.0232
	(0.0162)	(0.0155)
Observations	4,057	1,677
R-squared	0.619	0.702
Firm FE	YES	YES
Year FE	YES	YES
Previous firm distance	High Distance previous firm	Low Distance previous firm
Num. firms	531	210

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.8: Outside CEO and norm change, previous firm distance,

VARIABLES	(1) Norm change (prior year)	(2) Norm change (prior year)
R&D intensity	-0.518	-7.000
	(0.535)	(11.39)
salary	0.00319	0.00847
	(0.00343)	(0.0257)
Avg. num. claims	0.00387	0.473*
	(0.0251)	(0.246)
Vector change, (mean ind.)	0.630	23.10*
	(2.700)	(11.86)
Growth, (2 yrs)	0.594*	3.290
	(0.357)	(6.180)
Outside CEO	2.807	15.17
	(2.028)	(9.917)
Observations	4,057	1,677
R-squared	0.159	0.042
Firm FE	YES	YES
Year FE	YES	YES
Previous firm distance	High Distance previous firm	Low Distance previous firm
Num. firms	531	210

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.9: Outside CEO and distance from industry, previous firm distance

VARIABLES	(1) Vector change (mean ind.)	(2) Vector change (mean ind.)
R&D intensity	-0.00150	0.0519*
	(0.0100)	(0.0292)
salary	1.07e-05	-2.26e-05
	(1.76e-05)	(1.48e-05)
Avg. num. claims	-0.000264	-0.000149
	(0.000383)	(0.000931)
Vector change, (mean ind.)	0.238***	0.255***
	(0.0419)	(0.0630)
Patent applications	-7.93e-05**	-1.85e-05*
	(3.98e-05)	(9.53e-06)
Growth, (2 yrs)	-0.00364	-0.00689
	(0.00237)	(0.0105)
Outside CEO	-0.0331**	-0.0131
	(0.0137)	(0.0304)
Observations	4,057	1,677
R-squared	0.771	0.772
Firm FE	YES	YES
Year FE	YES	YES
Previous firm distance	High Distance previous firm	Low Distance previous firm
Num. firms	531	210

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.10: Impact of outside CEO for high and low R&D intensity firms

VARIABLES	(1) Vector change (prior year)	(2) Vector change (prior year)	(3) Norm change (prior year)	(4) Norm change (prior year)	(5) Vector change (mean ind.)	(6) Vector change (mean ind.)
R&D intensity	-0.0198**	1.033	-0.992	-175.9	-4.47e-05	0.342
	(0.00987)	(0.628)	(0.663)	(164.6)	(0.00989)	(0.540)
salary	2.41e-05	-1.20e-05	0.00126	-0.00581	-1.03e-05	1.69e-05
	(2.04e-05)	(1.81e-05)	(0.00484)	(0.0174)	(1.71e-05)	(1.10e-05)
Avg. num. claims	-0.000339	-0.00164**	0.0983*	0.0563	-0.000266	-0.000621
	(0.000603)	(0.000740)	(0.0515)	(0.0561)	(0.000433)	(0.000717)
Vector change, (mean ind.)	0.407***	0.566***	1.501	9.001***	0.295***	0.292***
	(0.0385)	(0.0430)	(3.534)	(2.772)	(0.0435)	(0.0479)
Patent applications	-2.08e-05	-1.92e-05*			-4.94e-05***	-3.55e-05*
	(2.20e-05)	(1.01e-05)			(1.77e-05)	(1.88e-05)
Growth, (2 yrs)	-0.00509*	0.00570	0.00888	-1.621	-0.00185	-0.0143
	(0.00291)	(0.0153)	(0.295)	(4.648)	(0.00220)	(0.0109)
Outside CEO	-0.0192	-0.0439**	2.815	20.59**	-0.00205	-0.0370*
	(0.0131)	(0.0215)	(1.880)	(9.972)	(0.0126)	(0.0203)
Observations	3,322	2,445	3,322	2,445	3,322	2,445
R-squared	0.557	0.646	0.078	0.025	0.775	0.744
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
R&D intensity	High	Low	High	Low	High	Low
Num. firms	319	251	319	251	319	251

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3.11: Propensity score matching

	Coefficient	Robust SD	p-value	95% Confidence Interval	
ATE Outsider (1 vs 0)	-0.03778	0.01275	0.003	-0.06277	-0.0128

Note: Logit as treatment model. Outcome is vector change in year t+1, as before.

Preface references

Ahuja G (2000) Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study. *Adm. Sci. Q.* 45(3):425.

Ahuja G, Katila R (2001) Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. *Strateg. Manag. J.* 22(3):197–220.

Alcácer J, Chung W (2007) Location Strategies and Knowledge Spillovers. *Manag. Sci.* 53(5):760–776.

Alcácer J, Chung W (2014) Location strategies for agglomeration economies. *Strateg. Manag. J.* 35(12):1749–1761.

Alcácer J, Oxley J (2014) Learning by supplying. *Strateg. Manag. J.* 35(2):204–223.

Arrow KJ (1974) *The Limits of Organization* (W. W. Norton & Company).

Bloom N, Schankerman M, Van Reenen J (2013) Identifying Technology Spillovers and Product Market Rivalry. *Econometrica* 81(4):1347–1393.

Brandenburger A, Stuart H (2007) Biform Games. *Manag. Sci.* 53(4):537–549.

Brandenburger AM, Stuart HW (1996) Value-based Business Strategy. *J. Econ. Manag. Strategy* 5(1):5–24.

Chandler AD (1962) *Strategy and Structure: Chapters in the History of the American Industrial Enterprise* (Beard Books).

Coase RH (1937) The Nature of the Firm. *Economica* 4(16):386–405.

Crossland C, Jinyong Zyung, Hiller NJ, Hambrick DC (2014) Ceo Career Variety: Effects on Firm-Level Strategic and Social Novelty. *Acad. Manage. J.* 57(3):652–674.

Cummings T, Knott AM (2018) Outside CEOs and Innovation. *Strateg. Manag. J.* Accepted Author Manuscript.

Dushnitsky G, Lenox MJ (2005) When do incumbents learn from entrepreneurial ventures?: Corporate venture capital and investing firm innovation rates. *Res. Policy* 34(5):615–639.

Dushnitsky G, Shaver JM (2009) Limitations to interorganizational knowledge acquisition: the paradox of corporate venture capital. *Strateg. Manag. J.* 30(10):1045–1064.

Dyer JH, Hatch NW (2006) Relation-specific capabilities and barriers to knowledge transfers: creating advantage through network relationships. *Strateg. Manag. J.* 27(8):701–719.

Gans J, Ryall MD (2017) Value capture theory: A strategic management review. *Strateg. Manag. J.* 38(1):17–41.

Hall BH, Khan B (2003) Adoption of New Technology. Jones DC, ed. *New Econ. Handb.* (Academic Press).

Hall BH, Mairesse J, Mohnen P (2010) Measuring the Returns to R&D. *Handb. Econ. Innov.* 2:1033–1082.

Hambrick D (1994) Top management groups: A conceptual integration and reconsideration of the team label. *Res. Organ. Behav.* (JAI Press, Greenwich, CT), 171–214.

Hambrick DC, Geletkanycz MA, Fredrickson JW (1993) Top Executive Commitment to the Status Quo: Some Tests of Its Determinants. *Strateg. Manag. J.* 14(6):401–418.

Hambrick DC, Mason PA (1984) Upper Echelons: The Organization as a Reflection of Its Top Managers. *Acad. Manage. Rev.* 9(2):193–206.

von Hippel E (2001) User toolkits for innovation. *J. Prod. Innov. Manag.* 18(4):247–257.

Hitt MA, Hoskisson RE, Ireland RD, Harrison JS (1991) Effects Of Acquisitions on R&D Inputs and Outputs. *Acad. Manage. J.* 34(3):693–706.

Hitt MA, Hoskisson RE, Johnson RA, Moesel DD (1996) The Market for Corporate Control and Firm Innovation. *Acad. Manage. J.* 39(5):1084–1119.

Howells J, Gagliardi D, Malik K (2008) The growth and management of R&D outsourcing: evidence from UK pharmaceuticals. *RD Manag.* 38(2):205–219.

Jeppesen LB, Molin MJ (2003) Consumers as Co-developers: Learning and Innovation Outside the Firm. *Technol. Anal. Strateg. Manag.* 15(3):363–383.

Kapoor R (2013) Persistence of Integration in the Face of Specialization: How Firms Navigated the Winds of Disintegration and Shaped the Architecture of the Semiconductor Industry. *Organ. Sci.* 24(4):1195–1213.

Katila R, Rosenberger JD, Eisenhardt KM (2008) Swimming with Sharks: Technology Ventures, Defense Mechanisms and Corporate Relationships. *Adm. Sci. Q.* 53(2):295–332.

Knott AM (2017) *The Record on R&D Outsourcing* (Social Science Research Network, Rochester, NY).

Knott AM, Posen HE, Wu B (2009) Spillover Asymmetry and Why It Matters. *Manag. Sci.* 55(3):373–388.

Kogut B, Zander U (1992) Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organ. Sci.* 3(3):383–397.

Kogut B, Zander U (1996) What Firms Do? Coordination, Identity, and Learning. *Organ. Sci.* 7(5):502–518.

Lee J, Kim M (2014) Market-Driven Technological Innovation Through Acquisitions The Moderating Effect of Firm Size. *J. Manag.*:0149206314535439.

Leiblein MJ, Miller DJ (2003) An empirical examination of transaction- and firm-level influences on the vertical boundaries of the firm. *Strateg. Manag. J.* 24(9):839–859.

Lippman SA, Rumelt RP (2003) A Bargaining Perspective on Resource Advantage. *Strateg. Manag. J.* 24(11):1069–1086.

Mackey A (2008) The effect of CEOs on firm performance. *Strateg. Manag. J.* 29(12):1357–1367.

Maggie ZY, Wan X (2017) Product variety and vertical integration. *Strateg. Manag. J.* 38(5):1134–1150.

Makri M, Hitt MA, Lane PJ (2010) Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. *Strateg. Manag. J.* 31(6):602–628.

Malmendier U, Nagel S (2011) Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?*. *Q. J. Econ.* 126(1).

Monteverde K (1995) Technical Dialog as an Incentive for Vertical Integration in the Semiconductor Industry. *Manag. Sci.* 41(10):1624–1638.

Mowery DC, Oxley JE, Silverman BS (1998) Technological overlap and interfirm cooperation: implications for the resource-based view of the firm. *Res. Policy* 27(5):507–523.

Mowery DC, Oxley JE, Silverman BS (2014) Strategic alliances and interfirm knowledge transfer. *Strateg. Manag. J.* 17(S2):77–91.

Nickerson JA, Zenger TR (2004) A Knowledge-Based Theory of the Firm: The Problem-Solving Perspective. *Organ. Sci.* 15(6):617–632.

Nonaka I (1994) A Dynamic Theory of Organizational Knowledge Creation. *Organ. Sci.* 5(1):14–37.

Pahnke EC, McDonald R, Wang D, Hallen B (2015) Exposed: Venture Capital, Competitor Ties, and Entrepreneurial Innovation. *Acad. Manage. J.* 58(5):1334–1360.

Reagans R, Singh PV, Krishnan R (2015) Forgotten Third Parties: Analyzing the Contingent Association Between Unshared Third Parties, Knowledge Overlap, and Knowledge Transfer Relationships with Outsiders. *Organ. Sci.* 26(5):1400–1414.

Ryall MD, MacDonald G (2004) How Do Value Creation and Competition Determine Whether a Firm Appropriates Value? *Manag. Sci.* 50(10):1319–1333.

Sampson RC (2007) R&D Alliances and Firm Performance: The Impact of Technological Diversity and Alliance Organization on Innovation. *Acad. Manage. J.* 50(2):364–386.

Williamson OE (1979) Transaction-Cost Economics: The Governance of Contractual Relations. *J. Law Econ.* 22(2):233–261.

Williamson OE (1991) Strategizing, economizing, and economic organization. *Strateg. Manag. J.* 12(S2):75–94.

Appendix 1: Appendix to Chapter 1

A1. Proof that feasibility and stability implies that players cannot appropriate more than their value added

Player $i's$ value added is equal to the value when all players are in the coalition minus the value created when i leaves the coalition. Define $G/\{i\}$ to be the groups of all players except player i, and $v_{G\setminus\{i\}}$ to be the value produced by the coalition $G/\{i\}$. We can show that no player can appropriate more than their added value. Suppose, by contradiction, that i appropriates in equilibrium more than their added value, that is $\pi_i > v_{G\setminus\{i\}}$. Then the distribution could not have satisfied (1.2), since:

$$\pi_i + \sum_{j\in G\setminus\{i\}} \pi_j = v_G$$
$$\sum_{j\in G\setminus\{i\}} \pi_j = v_G - \pi_i < v_G - (\overbrace{v_G - v_{G\setminus\{i\}}}^{\text{i's value added}}) = v_{G\setminus\{i\}} \tag{A.3}$$

A2. Generalized bargain solution, no vertical integration

Let Table 1A represents the valuation of each final product by the consumer, such that $A > B > C > D$.

Table A1.1: Product values

Product	Value
$FP_{1,1}$	A
$FP_{1,2}$	B
$FP_{2,1}$	C
$FP_{2,2}$	D

Table 2A shows the value created by all possible coalitions between the players.

Table A1.2: Value table, no integration

Group	Value	Group	Value
S1, S2, P1, P2, Co	A	S1, P1, Co	A
S1, S2, P1, Co	A	S1, P2, Co	B
S1, S2, P2, Co	B	S2, P1, Co	C
S1, P1, P2, Co	A	S2, P2, Co	D
S2, P1, P2, Co	C	Other Groups	0

Since no firm can appropriate more than its value added, the equilibrium has to satisfy the following inequalities:

$$\begin{aligned} &\pi_{S1} \le A - C \\ &\pi_{P1} \le A - B \\ &\pi_{S2} \le 0 \\ &\pi_{P2} \le 0 \\ &\pi_{Co} \le A \end{aligned} \tag{A.4}$$

To show that the inequalities in (A.4) are binding, consider the distributions of value shown in (A.5). All those distributions are in the core, showing that the added value of each player is tenable.

$$\begin{aligned} &\pi_{Co} = A, \pi_{S1} = \pi_{P1} = \pi_{S2} = \pi_{P2} = 0 \ (\pi_{Co} = Co's \text{ added value}) \\ &\pi_{Co} = C, \pi_{S1} = A - C, \pi_{P1} = \pi_{S2} = \pi_{P2} = 0 \ (\pi_{S1} = S1's \text{ added value}) \\ &\pi_{Co} = B, \pi_{P1} = A - B, \pi_{S1} = \pi_{S2} = \pi_{P2} = 0 \ (\pi_{P1} = P1's \text{ added value}) \end{aligned} \tag{A.5}$$

By stability, the one player coalitions require that $\pi_i \ge 0$ for all i. Two players coalitions impose the same conditions. From this, we can deduce that $\pi_{S2}, \pi_{P2} = 0$. Three players coalitions require:

$$\begin{aligned} &\pi_{S1} + \pi_{P1} + \pi_{Co} \ge A \text{ (value generated by S1,P1,Co)} \\ &\pi_{S1} + \pi_{P2} + \pi_{Co} \ge B \text{ (value generated by S1,P2,Co)} \\ &\pi_{S2} + \pi_{P1} + \pi_{Co} \ge C \text{ (value generated by S2,P1,Co)} \\ &\pi_{S2} + \pi_{P2} + \pi_{Co} \ge D \text{ (value generated by S2,P2,Co)} \end{aligned} \tag{A.6}$$

Substituting $\pi_{S2}, \pi_{P2} = 0$ in (A.6), we get

$$
\begin{aligned}
\pi_{S1} + \pi_{P1} + \pi_{Co} &\geq A \\
\pi_{S1} + \pi_{Co} &\geq B \\
\pi_{P1} + \pi_{Co} &\geq C \\
\pi_{Co} &\geq D
\end{aligned}
\tag{1.7}
$$

Neither four nor five player coalitions generate no new restriction over what is shown in (1.7) after substituting $\pi_{S2}, \pi_{P2} = 0$.

Supposing that $D \geq B + C - A$[20], the equilibrium is characterized by

$$
\begin{aligned}
D \leq \pi_{Co} &\leq A \\
\max\{0, B - \pi_{Co}\} \leq \pi_{S1} &\leq A - C \\
\max\{0, C - \pi_{Co}\} \leq \pi_{P1} &\leq A - B \\
\pi_{S2} &= 0 \\
\pi_{P2} &= 0
\end{aligned}
\tag{A.8}
$$

Introducing the confidence indices for the consumer (α_{Co}) , S1 (α_{S1}) and P1 $(1-\alpha_{S1})$[21], we obtain the following payoffs:

$$
\begin{aligned}
\pi_{Co} &= \alpha_{Co} A + (1-\alpha_{Co}) D \\
\pi_{S1} &= \alpha_{S1}(A-C) + (1-\alpha_{S1}).\max\{0,(B-\pi_{Co})\} \\
\pi_{P1} &= (1-\alpha_{S1})(A-B) + \alpha_{S1}.\max\{0,(C-\pi_{Co})\} \\
\pi_{S2} &= 0 \\
\pi_{P2} &= 0
\end{aligned}
\tag{A.9}
$$

To cover all possible scenarios, we need to divide the analysis into three parts; $\pi_{Co} \leq C$, $C < \pi_{Co} \leq B$, and $B < \pi_{Co}$. For each part, the equilibrium will be defined, and there will be a discussion of the impacts of innovation on the firms' payoffs.

A2.1 Low consumer's confidence index $\pi_{Co} \leq C$

If $\pi_{Co} \leq C$, the equilibrium values shown in (A.9) will be:

---

[20] This restriction comes from the fact that it is possible for S1, P1, S2, and P2 to appropriate their value added and the left over is higher than D. The assumption $D \geq B + C - A$ means that the product made by S2 and P2 is not too behind the other possible products.

[21] Refer back to the main text as to why P1 does not have an independent confidence index.

$$\begin{aligned}
\pi_{Co} &= \alpha_{Co} A + (1-\alpha_{Co}) D \\
\pi_{S1} &= \alpha_{S1}(A-C) + (1-\alpha_{S1}) B - \pi_{Co} \\
\pi_{P1} &= (1-\alpha_{S1})(A-B) + \alpha_{S1} C - \pi_{Co} \\
\pi_{S2} &= 0 \\
\pi_{P2} &= 0
\end{aligned} \tag{A.10}$$

Substituting $\pi_{Co} = \alpha_{Co} A + (1-\alpha_{Co}) D$ into the payoffs in (A.10), we get:

$$\begin{aligned}
\pi_{Co} &= \alpha_{Co} A + (1-\alpha_{Co}) D \\
\pi_{S1} &= (\alpha_{S1}(1+\alpha_{Co}) - \alpha_{Co}) A + (1-\alpha_{S1})(B-(1-\alpha_{Co})D) - \alpha_{S1} C \\
\pi_{P1} &= (1-\alpha_{S1}(1+\alpha_{Co})) A + \alpha_{S1}(C-(1-\alpha_{Co})D) - (1-\alpha_{S1}) B \\
\pi_{S2} &= 0 \\
\pi_{P2} &= 0
\end{aligned} \tag{A.11}$$

Having $\pi_{Co} \leq C$ is equivalent to requiring that the consumer bargaining ability satisfy condition (A.12).

$$\alpha_{Co} \leq \frac{C-D}{A-D} \tag{A.12}$$

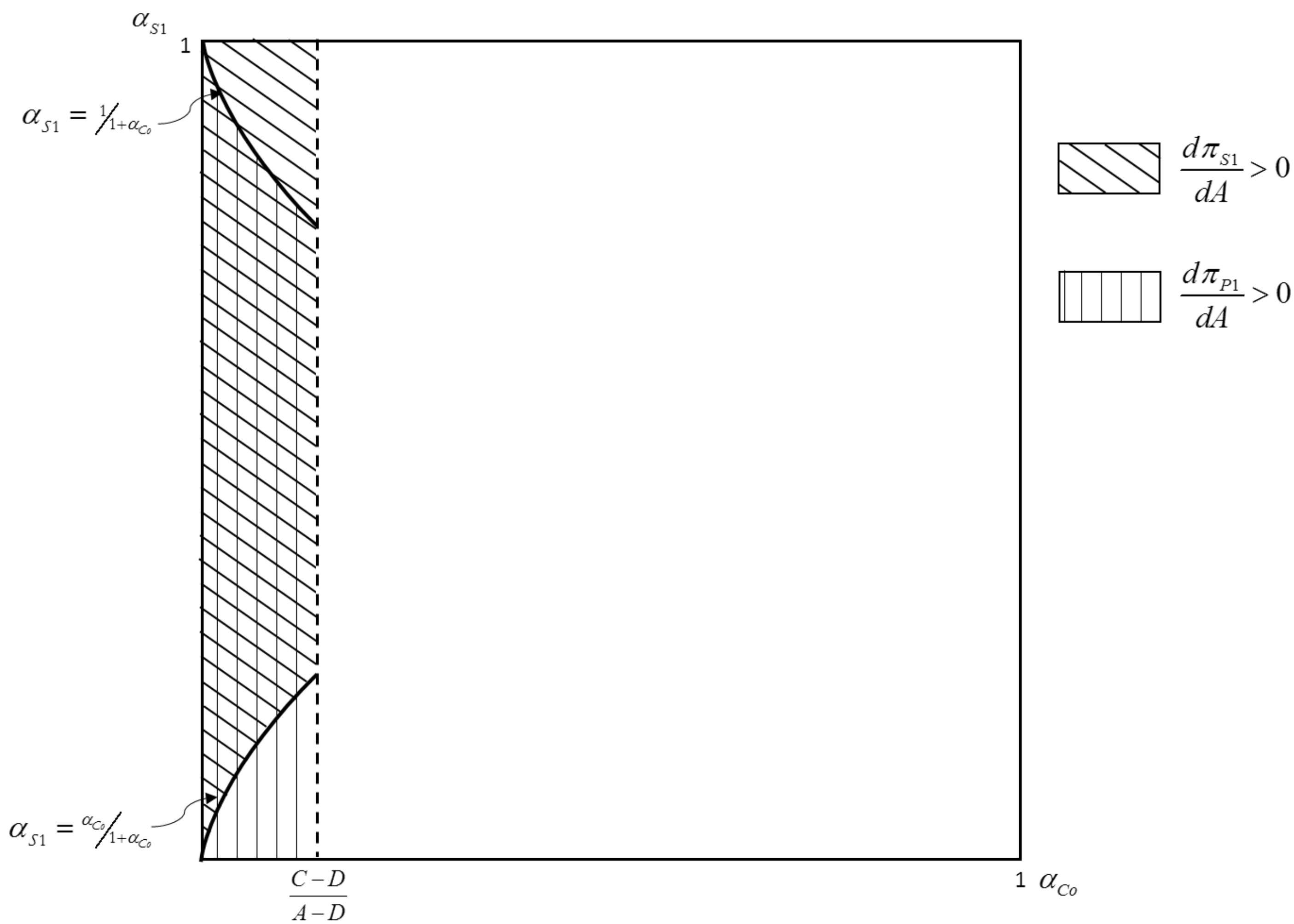
Given the consumer's confidence index, condition (A.12) is more likely to be satisfied the closer C is to A, that is, the more competitive product $FP_{2,1}$ is in relation to $FP_{1,1}$. In other words, if P1 views S2 input as a close substitute for S1, it is likely that (A.11) represents the equilibrium.

The equilibrium in (A.11) shows that firm S1 is better off by an increase in the value of $FP_{1,1}$ (that is, the derivative of π_{S1} with respect to A is positive) when $\alpha_{S1} \geq \alpha_{Co} / (1+\alpha_{Co})$. Likewise, increasing A is beneficial to P1 when $\alpha_{S1} \leq 1 / (1+\alpha_{Co})$. In summary,

$$\begin{aligned}
\frac{d\pi_{S1}}{dA} > 0 \text{ if } \alpha_{S1} > \frac{\alpha_{Co}}{1+\alpha_{Co}} \\
\frac{d\pi_{P1}}{dA} > 0 \text{ if } \alpha_{S1} < \frac{1}{1+\alpha_{Co}}
\end{aligned} \tag{A.13}$$

Figure A1 shows when an increase in A is beneficial for S1 (represented by the area in diagonal hachure) and P1 (represented by the area in vertical hachure).

Figure A1.1: Increase in A, low consumer bargaining ability, no integration



A2.2 Medium consumer's confidence index $C < \pi_{Co} \leq B$

If $C < \pi_{Co} \leq B$, the equilibrium values shown in (A.9) will be:

$$
\begin{aligned}
\pi_{Co} &= \alpha_{Co}A + (1-\alpha_{Co})D \\
\pi_{S1} &= \alpha_{S1}(A-C) + (1-\alpha_{S1})(B-\pi_{Co}) \\
\pi_{P1} &= \alpha_{P1}(A-B) \\
\pi_{S2} &= 0 \\
\pi_{P2} &= 0
\end{aligned}
\qquad (A.14)
$$

Requiring $C < \pi_{Co} \le B$ is equivalent to looking at a consumer with bargaining ability satisfying condition (A.15).

$$\frac{C-D}{A-D} < \alpha_{Co} \le \frac{B-D}{A-D} \tag{A.15}$$

If we require that $\pi_{Co} + \pi_{S1} + \pi_{P1} = A$, we need to set up α_{P1} as shown in (A.16).

$$\begin{gathered}
\pi_{Co} + \overbrace{\alpha_{S1}(A-C) + (1-\alpha_{S1})(B-\pi_{Co})}^{\pi_{S1}} + \overbrace{\alpha_{P1}(A-B)}^{\pi_{P1}} = A \\
\alpha_{S1}\pi_{Co} + \alpha_{S1}(A-C) + (1-\alpha_{S1})B + \alpha_{P1}(A-B) = A \\
\alpha_{S1}\pi_{Co} + \alpha_{S1}(A-B) - \alpha_{S1}C + B + \alpha_{P1}(A-B) = A \\
\alpha_{P1} = \frac{-\alpha_{S1}\pi_{Co} + (1-\alpha_{S1})(A-B) + \alpha_{S1}C}{A-B} \\
\alpha_{P1} = \frac{-\alpha_{S1}\alpha_{Co}A - (\alpha_{S1} - \alpha_{S1}\alpha_{Co})D + (1-\alpha_{S1})(A-B) + \alpha_{S1}C}{A-B} \\
\alpha_{P1} = \frac{\alpha_{S1}(C-D) - \alpha_{S1}\alpha_{Co}(A-D) + (1-\alpha_{S1})(A-B)}{A-B}
\end{gathered} \tag{A.16}$$

To ensure that α_{P1} defined in expression (A.16) lies in the interval $[0,1]$ and therefore is a valid confidence index, we need the extra condition[22] (A.17).

$$\begin{aligned}
&\frac{\alpha_{S1}(C-D) - \alpha_{S1}\alpha_{Co}(A-D) + (1-\alpha_{S1})(A-B)}{A-B} \ge 0 \\
&\alpha_{S1}(C-D) - \alpha_{S1}\alpha_{Co}(A-D) + (1-\alpha_{S1})(A-B) \ge 0 \\
&\alpha_{S1}(C-D-\alpha_{Co}(A-D) - A + B) \ge B - A \\
&\alpha_{S1} \le \frac{A-B}{\alpha_{Co}(A-D) + A - B + D - C}
\end{aligned} \tag{A.17}$$

Thus, if $\frac{C-D}{A-D} < \alpha_{Co} \le \frac{B-D}{A-D}$ and condition (A.17) is satisfied, we can substitute (A.16) in (A.14) and obtain the equilibrium payoffs

---

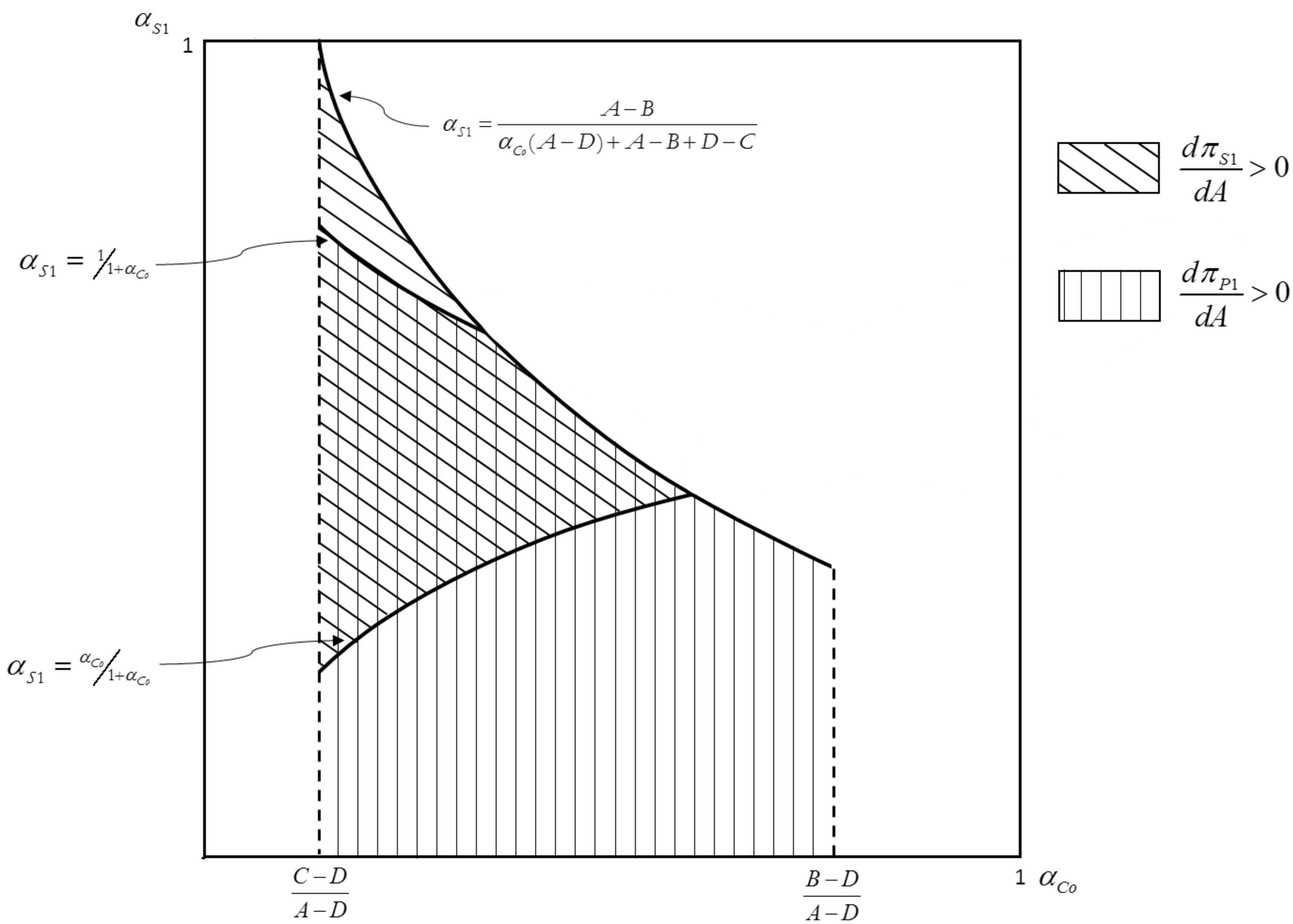
[22] The condition (A.17) implies that $\alpha_{P1} \ge 0$ because the assumption that $D \ge B + C - A$ already guarantees $\alpha_{P1} \le 1$.

$$\begin{aligned}
\pi_{Co} &= \alpha_{Co}A + (1-\alpha_{Co})D \\
\pi_{S1} &= \alpha_{S1}(A-C) + (1-\alpha_{S1})(B-[\alpha_{Co}A + (1-\alpha_{Co})D]) \\
\pi_{P1} &= \alpha_{S1}(C-D) - \alpha_{S1}\alpha_{Co}(A-D) + (1-\alpha_{S1})(A-B) \\
\pi_{S2} &= 0 \\
\pi_{P2} &= 0
\end{aligned} \tag{A.18}$$

Taking derivatives with respect to A on (A.18) shows that the benefit for firms S1 and P1 of increases in A remain the same as shown in (A.13): firm S1 is better off when $\alpha_{S1} \geq \alpha_{Co}/(1+\alpha_{Co})$ and P1 prefers a higher A when $\alpha_{S1} \leq 1/(1+\alpha_{Co})$.

Figure A2 shows when an increase in A is beneficial for S1 (represented by the area in diagonal hachure) and P1 (represented by the area in vertical hachure) for medium values of the consumer's confidence index.

Figure A1.2: Beneficiaries of increase in A, medium consumer bargaining ability, no integration



A2.3 High consumer's confidence index $B < \pi_{Co}$

The final case is equivalent to $\alpha_{Co} > \frac{B-D}{A-D}$. When the confidence index for the consumer is that high, the expected value for each player is:

$$\begin{aligned}
\pi_{Co} &= \alpha_{Co}A + (1-\alpha_{Co})D \\
\pi_{S1} &= \alpha_{S1}(A-C) \\
\pi_{P1} &= \alpha_{P1}(A-B) \\
\pi_{S2} &= 0 \\
\pi_{P2} &= 0
\end{aligned} \tag{A.19}$$

The confidence indices consistent with feasibility are:

$$\begin{aligned}
\frac{B-D}{A-D} &< \alpha_{Co} \le 1 \\
0 \le \alpha_{S1} &\le \frac{(A-D)(1-\alpha_{Co})}{A-C} \\
\alpha_{P1} &= \frac{(A-D)(1-\alpha_{Co}) + \alpha_{S1}(C-A)}{A-B}
\end{aligned} \tag{A.20}$$

The range for α_{S1} come from the fact that the highest α_{S1} possible will be when $\alpha_{Co} = \frac{B-D}{A-D}$ and $\alpha_{S1} = 0$. If α_{S1} is bigger than zero, then the value for α_{P1} in (A.20) guarantees that $\pi_{Co} + \pi_{S1} + \pi_{P1} = A$. In this case, both α_{Co} and α_{S1} are exogenous, so that S1 find beneficial a higher A as shown in (A.19). For P1, an increase in A increase its value appropriated if $\alpha_{S1} \le 1-\alpha_{Co}$.

$$\begin{aligned}
&\frac{d\pi_{S1}}{dA} > 0 \\
&\frac{d\pi_{P1}}{dA} > 0 \text{ if } \alpha_{S1} \le 1-\alpha_{Co}
\end{aligned} \tag{A.21}$$

Figure A3 shows when an increase in A is beneficial for S1 (represented by the area in diagonal hachure) and P1 (represented by the area in vertical hachure) for high values of the consumer's confidence index.

Figure A1.3: Beneficiaries of increase in A, high consumer bargaining ability, no integration

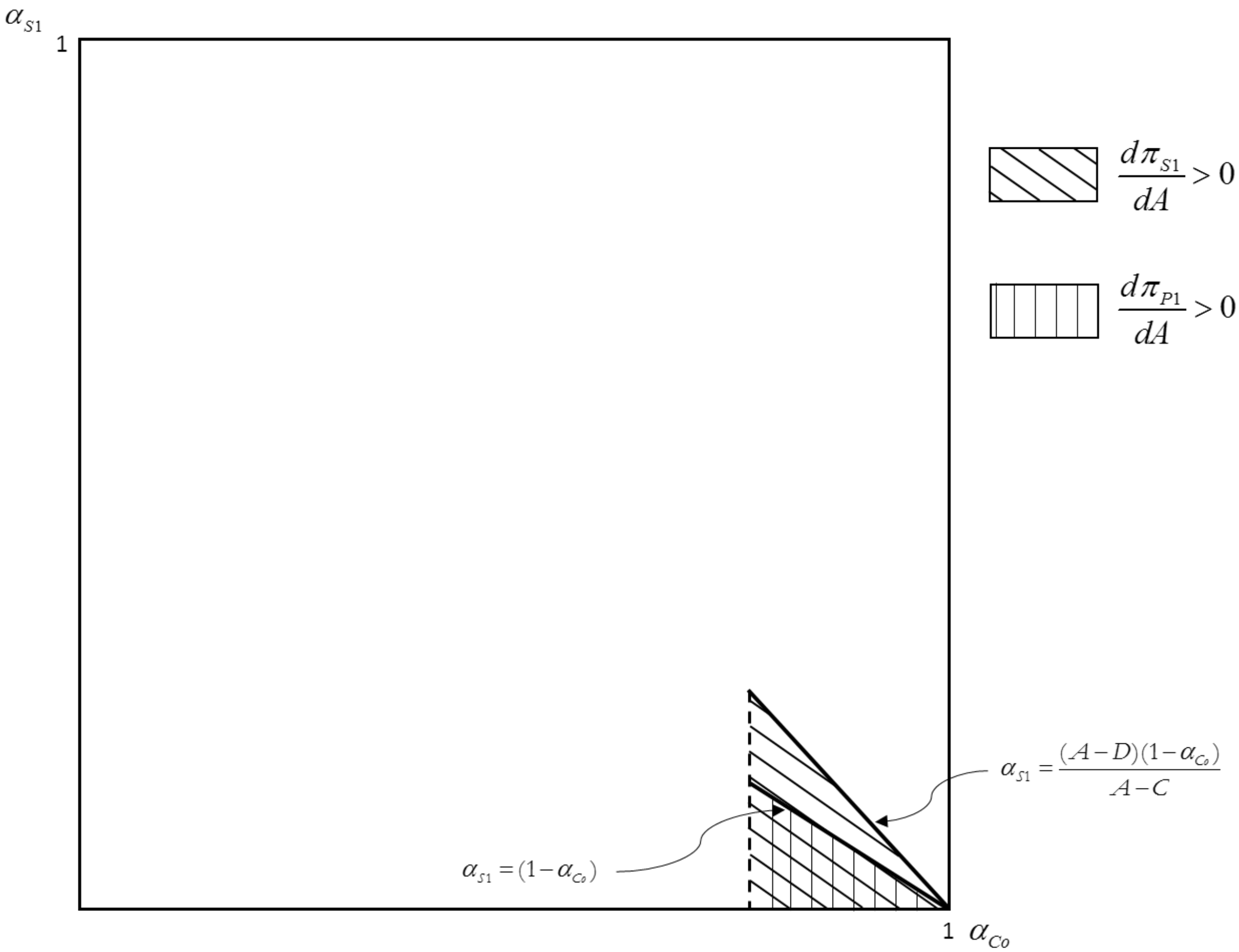


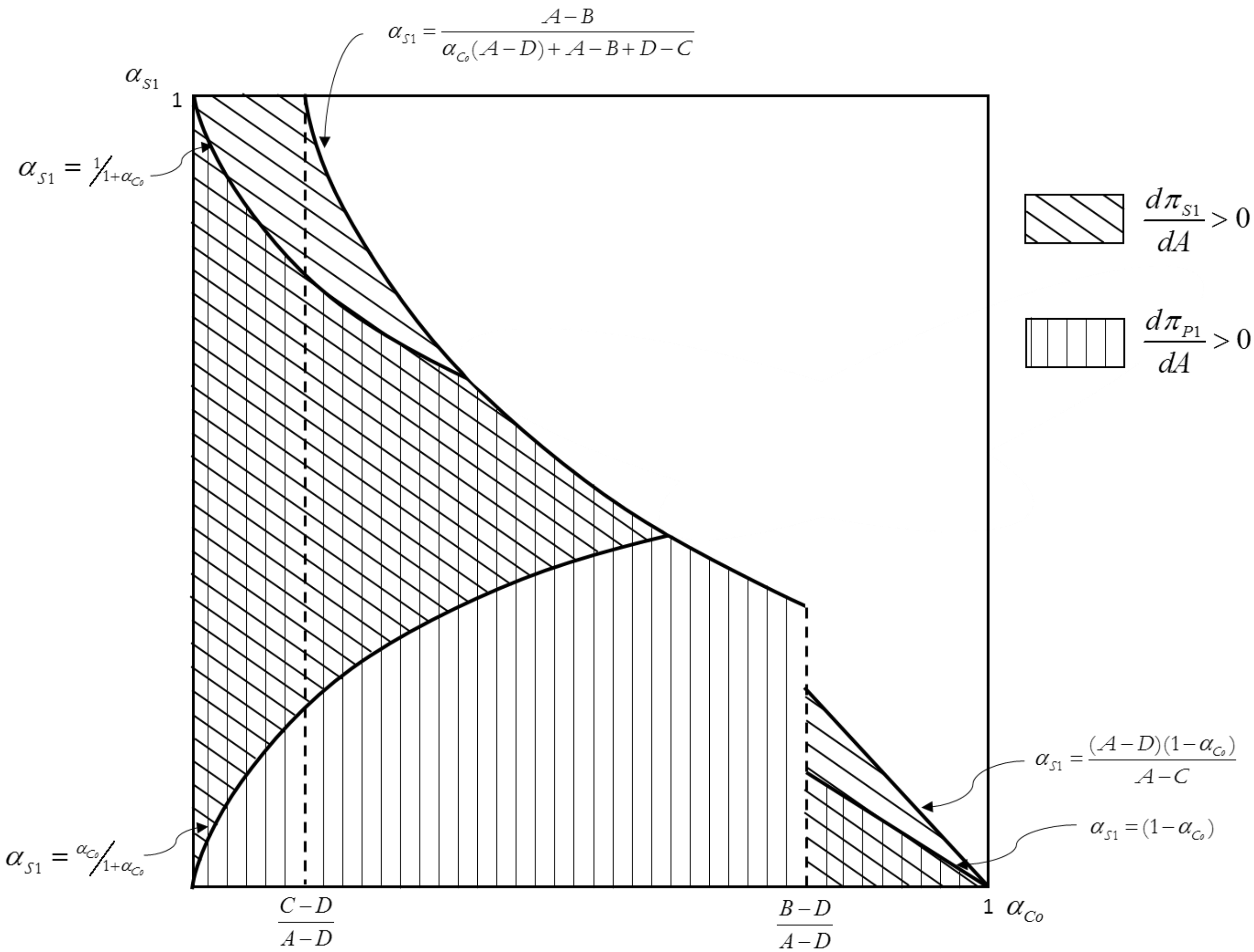
Table 3A summarizes when an increase in A is beneficial for firms S1 and P1.

Table A1.3: Willingness to adopt new input, no integration.

α_{Co}	$\alpha_{Co} \le \frac{C-D}{A-D}$	$\frac{C-D}{A-D} < \alpha_{Co} \le \frac{B-D}{A-D}$	$\alpha_{Co} > \frac{B-D}{A-D}$
α_{S1}	$0 \le \alpha_{S1} \le 1$	$0 \le \alpha_{S1} \le \frac{A-B}{\alpha_{Co}(A-D)+A-B+D-C}$	$0 \le \alpha_{S1} \le \frac{(A-D)(1-\alpha_{Co})}{A-C}$
$\frac{d\pi_{S1}}{dA} > 0$	$\alpha_{S1} \ge \frac{\alpha_{Co}}{1+\alpha_{Co}}$	$\alpha_{S1} \ge \frac{\alpha_{Co}}{1+\alpha_{Co}}$	Always
$\frac{d\pi_{P1}}{dA} > 0$	$\alpha_{S1} \le \frac{1}{1+\alpha_{Co}}$	$\alpha_{S1} \le \frac{1}{1+\alpha_{Co}}$	$\alpha_{S1} \le 1-\alpha_{Co}$

Figure A4 shows Table A3, joining Figures A1, A2, and A3

Figure A1.4: Willingness to adopt new input, no integration.



Adoption. Let the adoption cost be $k > 0$ arbitrarily small. Then, P1 is willing to adopt innovation from supplier S1 if $\frac{d\pi_{P1}}{dA} > 0$.

Innovation. S1 will innovate when its payoff is increasing in A, that is, when $\frac{d\pi_{S1}}{dA} > 0$ and when P1 would adopt, $\frac{d\pi_{P1}}{dA} > 0$ (because of the innovation cost).

Information Sharing. The intersection that determines if the increase in A is beneficial for both S1 and P1 is the parameters that satisfy $\frac{d\pi_{S1}}{dA} > 0$ and $\frac{d\pi_{P1}}{dA} > 0$. Given any information transmission cost, P1 would only release information if S1 would innovate the input and P1 would adopt the new input.

As before, P2 would not send information because its payoff is zero in all cases. Likewise, S2 would not get information from producers.

The determination of the derivative of the payoff functions with respect to A is enough to determine when information is going to get shared and the input will be improved and adopted. The same reasoning will apply to the cases where S1 is integrated with one of the producers.

A3. Generalized bargain solution, integration between S1 and P1

The value table for the case where S1 is integrated with P1 depicted in Table 4A.

Table A1.4: Value table, integrated S1and P1

Group	Value	Group	Value
S1P1, S2, P2, Co	A	S2, P2, Co	D
S1P1, S2, Co	A	S1P1, Co	A
S1P1, P2, Co	A	Other Groups	0

The added value of each player imposes the conditions (A.22).

$$\begin{gathered} \pi_{S1P1} \leq A - D \\ \pi_{Co} \leq A \\ \pi_{S2}, \pi_{P2} \leq 0 \end{gathered} \tag{A.22}$$

To show that the inequalities in (A.22) are binding, consider the distributions of value shown in. All those distributions are in the core, showing that the added value of each player is tenable.

$$\begin{gathered} \pi_{Co} = A, \pi_{S1P1} = \pi_{S2} = \pi_{P2} = 0 \ (\pi_{Co} = Co\text{'s added value}) \\ \pi_{Co} = D, \pi_{S1P1} = A - D, \pi_{S2} = \pi_{P2} = 0 \ (\pi_{S1} = S1P1\text{'s added value}) \end{gathered} \tag{A.23}$$

By stability, the one player coalitions require that $\pi_i \geq 0$ for all i. From this and (A.22) we can deduce that $\pi_{S2}, \pi_{P2} = 0$. The other coalitions require:

$$\begin{aligned}
\pi_{S1P1} + \pi_{Co} &\geq A \text{ (value generated by S1P1,Co)} \\
\pi_{S1P1} + \pi_{Co} &\geq A \text{ (value generated by S1P1,P2,Co)} \\
\pi_{S1P1} + \pi_{Co} &\geq A \text{ (value generated by S1P1,S2,Co)} \\
\pi_{Co} &\geq D \text{ (value generated by S2,P2,Co)} \\
\pi_{S1P1} + \pi_{Co} &\geq A \text{ (value generated by S1P1,S2,P2,Co)}
\end{aligned} \tag{A.24}$$

In summary, the following conditions are imposed in the core:

$$\begin{aligned}
D \leq \pi_{Co} \leq A \\
\pi_{S1P1} = A - \pi_{Co} \\
\pi_{S2}, \pi_{P2} = 0
\end{aligned} \tag{A.25}$$

The equilibrium outcomes in the bargaining phase are given by:

$$\begin{aligned}
\pi_{Co} &= A\alpha_{Co} + D(1-\alpha_{Co}) \\
\pi_{S1P1} &= (A-D)(1-\alpha_{Co}) \\
\pi_{S2} &= 0 \\
\pi_{P2} &= 0
\end{aligned} \tag{A.26}$$

It is clear that the payoff for firm S1P1 is always increasing in A. In this case, the information will always flow internally from P1 to S1, independently of the value of α_{Co}. The integrated firm S1P1 will also innovate and, adopt the input. P2 will not adopt or send information to any supplier.

A3. Generalized bargain solution, integration between S1 and P2

The value table for when S1 and P2 are integrated is represented in table 4.

Table A1.5: Value table, integrated S1and P2

Group	Value	Group	Value
S1P2, S2, P1, Co	A	S2, P1, Co	C
S1P2, S2, Co	B	S1P2, Co	B
S1P2, P1, Co	A	Other Groups	0

The added value of each player imposes the following conditions:

$$\begin{gathered}\pi_{S1P2} \le A - C \\ \pi_{P1} \le A - B \\ \pi_{Co} \le A \\ \pi_{S2} \le 0\end{gathered} \tag{A.27}$$

To show that the inequalities in (A.27) are binding, consider the distributions of value shown in. All those distributions are in the core, showing that the added value of each player is tenable.

$$\begin{gathered}\pi_{Co} = A, \pi_{S1P2} = \pi_{S2} = \pi_{P1} = 0 \ (\pi_{Co} = Co's \text{ added value}) \\ \pi_{Co} = C, \pi_{S1P2} = A - C, \pi_{S2} = \pi_{P1} = 0 \ (\pi_{S1} = S1P1's \text{ added value}) \\ \pi_{Co} = B, \pi_{P1} = A - B, \pi_{S2} = \pi_{S1P2} = 0 \ (\pi_{S1} = P1's \text{ added value})\end{gathered} \tag{A.28}$$

By stability, the one player coalitions require that $\pi_i \ge 0$ for all i. From this and (A.27) we can deduce that $\pi_{S2} = 0$. The other coalitions require:

$$\begin{gathered}\pi_{S1P2} + \pi_{Co} \ge B \text{ (value generated by S1P2,Co)} \\ \pi_{S1P2} + \pi_{P1} + \pi_{Co} \ge A \text{ (value generated by S1P2,P1,Co)} \\ \pi_{S1P2} + \pi_{Co} \ge B \text{ (value generated by S1P2,S2,Co)} \\ \pi_{P1} + \pi_{Co} \ge C \text{ (value generated by S2,P1,Co)} \\ \pi_{S1P2} + \pi_{P1} + \pi_{Co} \ge A \text{ (value generated by S1P2,S2,P1,Co)}\end{gathered} \tag{A.29}$$

In summary, the following conditions are imposed in the core:

$$D \le \pi_{Co} \le A$$
$$\pi_{S1P1} = A - \pi_{Co} \quad \text{(A.30)}$$
$$\pi_{S2}, \pi_{P2} = 0$$

Equilibrium outcomes follow the set of conditions:

$$\max\{0, B + C - A\} \le \pi_{Co} \le A$$
$$\max\{0, B - \pi_{Co}\} \le \pi_{S1P2} \le A - C$$
$$\max\{0, C - \pi_{Co}\} \le \pi_{P1} \le A - B \quad \text{(A.31)}$$
$$\pi_{S2} = 0$$

Notice that the lower bound for π_{Co} is the maximum between zero and $B + C - A$. This is because it is possible for both S1P2 and P1 to appropriate their added value and there is still value leftover for Co to appropriate. Indeed, if the market is competitive enough, it is likely that $B + C - A > 0$. This chapter will maintain this assumption.

A3.1 Low consumer's confidence index $\pi_{Co} \le C$

If $\pi_{Co} \le C$, then $\alpha_{Co} \le \frac{A - B}{2A - B - C}$. The equilibrium values shown in (A.31) will be:

$$\pi_{Co} = \alpha_{Co} A + (1 - \alpha_{Co})(B + C - A)$$
$$\pi_{S1P2} = \alpha_{S1P2}(A - C) + (1 - \alpha_{S1P2})(B - \pi_{Co})$$
$$\pi_{P1} = \alpha_{P1}(A - B) + (1 - \alpha_{P1})(C - \pi_{Co}) \quad \text{(A.32)}$$
$$\pi_{S2} = 0$$

If we to set up $\alpha_{P1} = 1 - \alpha_{S1P2}$, then $\pi_{Co} + \pi_{S1} + \pi_{P1} = A$. Substituting $\alpha_{P1} = 1 - \alpha_{S1P2}$ and the expanded form for π_{Co} into (A.32), we get (A.33).

$$\pi_{Co} = (2\alpha_{Co} - 1) A + (1 - \alpha_{Co})(B + C)$$
$$\pi_{S1P2} = [1 + 2\alpha_{Co}(\alpha_{S1P2} - 1)] A + [\alpha_{Co}(1 - \alpha_{S1P2})] B - [1 - \alpha_{Co}(1 - \alpha_{S1P2})] C$$
$$\pi_{P1} = (1 - 2\alpha_{Co}\alpha_{S1P2}) A - (1 - \alpha_{Co}\alpha_{S1P2}) B + \alpha_{Co}\alpha_{S1P2} C \quad \text{(A.33)}$$
$$\pi_{S2} = 0$$

When looking at the benefit of an increase in A for players, we get that:

$$\frac{d\pi_{S1P2}}{dA} > 0 \text{ if } \alpha_{S1P2} > \frac{2\alpha_{Co} - 1}{2\alpha_{Co}}$$
$$\frac{d\pi_{P1}}{dA} > 0 \text{ if } \alpha_{S1P2} < \frac{1}{2\alpha_{Co}} \tag{A.34}$$

Note that (A.34) displays the change in value for players when A ($FP_{1,1}$) increases. The first thing to notice here is that the derivatives do not depend on the values of B or C. Therefore, P1's choice to adopt A does not depend on adoption by P2 – it only depends on α_{S1P2} and α_{Co}. Second, the benefit for S1P2 do adopt an innovation in P2 (that is, increase the value of B) is $\alpha_{Co}(1-\alpha_{S1P2}) > 0$, a positive number. This means that S1P2 is better off by having P2 send information to S1, creating a better input and adopting the new input in P2 regardless of P1's decision to adopt. Therefore, P2 will always send information to the upstream unit of the integrated firm. It also means that P2 will adopt all input innovation, including the innovation that was born from P1's information sharing.

Firm's P1 decision to adopt a innovation follows (A.34), independently of where the information needed to innovate came from. However, P1's knowledge that the integrated P2 will always adopt impacts its decision to share information. To understand how S1P2's adoption affects the equilibrium, suppose that the value consumers give to a product made by firm i using the input from S1 ($FP_{1,i}$) is some function of the input quality I and producers' quality K_i so that $FP_{1,i} = f(I, K_i) \equiv f_i(I)$[23]. When both S1P2 and P1 adopts the new input, the total change in payoff for firm P1 is calculated in (A.35).

$$(1-2\alpha_{Co}\alpha_{S1P2})\frac{df_A(I)}{dI} - (1-\alpha_{Co}\alpha_{S1P2})\frac{df_B(I)}{dI} \tag{A.35}$$

Since (A.35) is smaller than $(1-2\alpha_{Co}\alpha_{S1P2})\frac{df_A(I)}{dI}$, producer P1 will choose to send information less frequently. Moreover, the bigger the impact of I on $f_B(I)$, the less information P1 will send to S1P2. If we assume that P1 is a higher quality producer ($K_{P1} > K_{P2}$) and consumers view

[23] The effect of increases in I is positive for both functions $f_i(I)$.

input quality and producers quality as substitutes ($\frac{\partial f(I,K_i)}{\partial I \partial K_i} < 0$), then $\frac{df_A(I)}{dI} < \frac{df_B(I)}{dI}$ and thus producer P1 will never send information to S1P2 because (A.35) will be negative for all values of α_{S1P2} and α_{Co}.

In general, P1 sends information if (A.35) is positive, that is, if (A.36) is satisfied.

$$\alpha_{S1P2} < \frac{\frac{df_A(I)}{dI} - \frac{df_B(I)}{dI}}{\alpha_{Co}(2\frac{df_A(I)}{dI} - \frac{df_B(I)}{dI})} \tag{A.36}$$

For notation simplicity, let $\gamma = \frac{\frac{df_B(I)}{dI}}{\frac{df_A(I)}{dI}}$. Then we can write the right hand side of (A.36) as (A.37).

$$\frac{1-\gamma}{\alpha_{Co}(2-\gamma)} \tag{A.37}$$

A3.2 Medium consumer's confidence index $C < \pi_{Co} \le B$

If $C < \pi_{Co} \le B$, then $\frac{A-B}{2A-B-C} < \alpha_{Co} \le \frac{A-C}{2A-B-C}$. The equilibrium values shown in (A.31) will be:

$$\begin{aligned}
\pi_{Co} &= \alpha_{Co}A + (1-\alpha_{Co})(B+C-A) \\
\pi_{S1P2} &= \alpha_{S1P2}(A-C) + (1-\alpha_{S1P2})(B-\pi_{Co}) \\
\pi_{P1} &= \alpha_{P1}(A-B) \\
\pi_{S2} &= 0
\end{aligned} \tag{A.38}$$

Since the chapter assumes that all confidence indices are consistent, we need to set up to α_{P1} be:

$$\alpha_{P1} = 1 + \frac{\alpha_{S1P2}\alpha_{Co}(B + C - 2A)}{A - B} \tag{A.39}$$

We need to insure $0 \le \alpha_{P1} \le 1$. Since $B + C - 2A < 0$, we can guarantee that $\alpha_{P1} \le 1$. The condition for α_{P1} to be bigger than zero implies

$$\alpha_{S1P2} \le \frac{A - B}{A - C} \tag{A.40}$$

Assuming that (A.40) is satisfied, we can substitute (A.39) into (A.38) to yield (A.41). Note that (A.41) is identical to (A.33). Therefore, the discussion about information sharing and adoption is the same as before.

$$\begin{aligned}
\pi_{Co} &= (2\alpha_{Co} - 1)A + (1 - \alpha_{Co})(B + C) \\
\pi_{S1P2} &= [1 + 2\alpha_{Co}(\alpha_{S1P2} - 1)]A + [\alpha_{Co}(1 - \alpha_{S1P2})]B - [1 - \alpha_{Co}(1 - \alpha_{S1P2})]C \\
\pi_{P1} &= (1 - 2\alpha_{Co}\alpha_{S1P2})A - (1 - \alpha_{Co}\alpha_{S1P2})B + \alpha_{Co}\alpha_{S1P2}C \\
\pi_{S2} &= 0
\end{aligned} \tag{A.41}$$

A3.3 High consumer's confidence index $B < \pi_{Co}$

Finally, we need to look at the case when $\alpha_{Co} > \frac{A - C}{2A - B - C}$. In this case, the payoffs will be:

$$\begin{aligned}
\pi_{Co} &= \alpha_{Co}A + (1 - \alpha_{Co})(B + C - A) \\
\pi_{S1P2} &= \alpha_{S1P2}(A - C) \\
\pi_{P1} &= \alpha_{P1}(A - B) \\
\pi_{S2} &= 0
\end{aligned} \tag{A.42}$$

In this case, the confidence indices compatible with the equilibrium are:

$$\begin{aligned}
\frac{A - C}{2A - B - C} &< \alpha_{Co} \le 1 \\
0 \le \alpha_{S1P2} &\le \frac{(2A - B - C)(1 - \alpha_{Co})}{A - C} \\
\alpha_{P1} &= \frac{(2A - B - C)(1 - \alpha_{Co}) + \alpha_{S1P2}(C - A)}{A - B}
\end{aligned} \tag{A.43}$$

Substituting (A.43) into (A.42) yields (A.44).

$$\begin{aligned} \pi_{Co} &= \alpha_{Co} A + (1-\alpha_{Co})(B+C-A) \\ \pi_{S1P2} &= \alpha_{S1P2}(A-C) \\ \pi_{P1} &= (1-\alpha_{Co})(2A-B-C) + \alpha_{S1P2}(C-A) \\ \pi_{S2} &= 0 \end{aligned} \tag{A.44}$$

The derivatives of payoffs with relation to A are calculated in (A.45).

$$\begin{aligned} &\frac{d\pi_{S1P2}}{dA} > 0 \\ &\frac{d\pi_{P1}}{dA} > 0 \text{ if } \alpha_{S1P2} < 2(1-\alpha_{Co}) \end{aligned} \tag{A.45}$$

Since $\frac{(2A-B-C)(1-\alpha_{Co})}{A-C} < 2(1-\alpha_{Co})$, $\frac{d\pi_{P1}}{dA} > 0$ for all equilibria when $\alpha_{Co} > \frac{A-C}{2A-B-C}$.

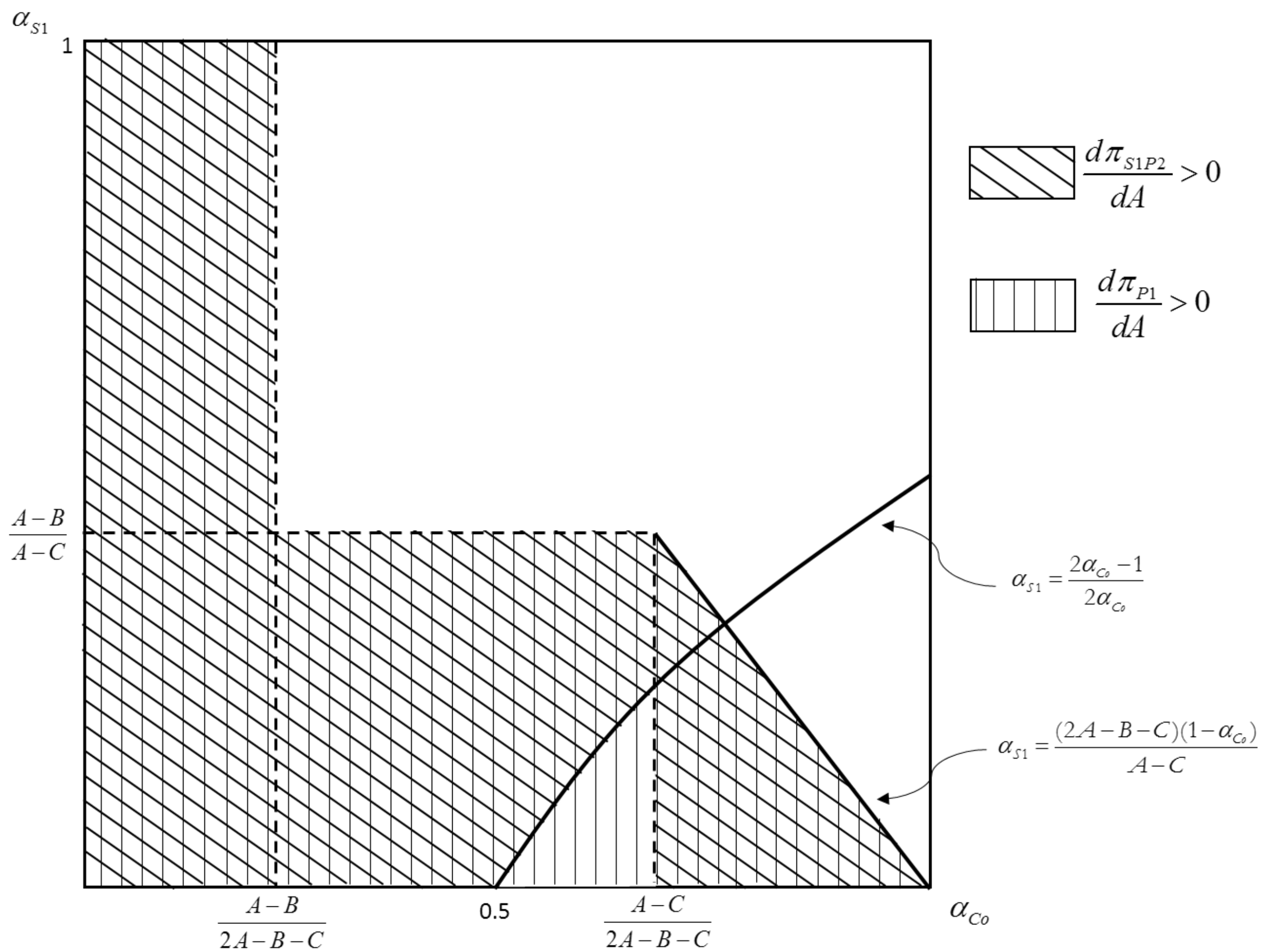
Firm S1P2 does not gain from increasing B, since its payoff function only depends on A and C, as shown in (A.42).

Table 6A and Figure 2A show when S1P2 and P1 are better off with an increase in A.

Table A1.6: Impact of larger A on firms' payoffs, integration between S1 and P2.

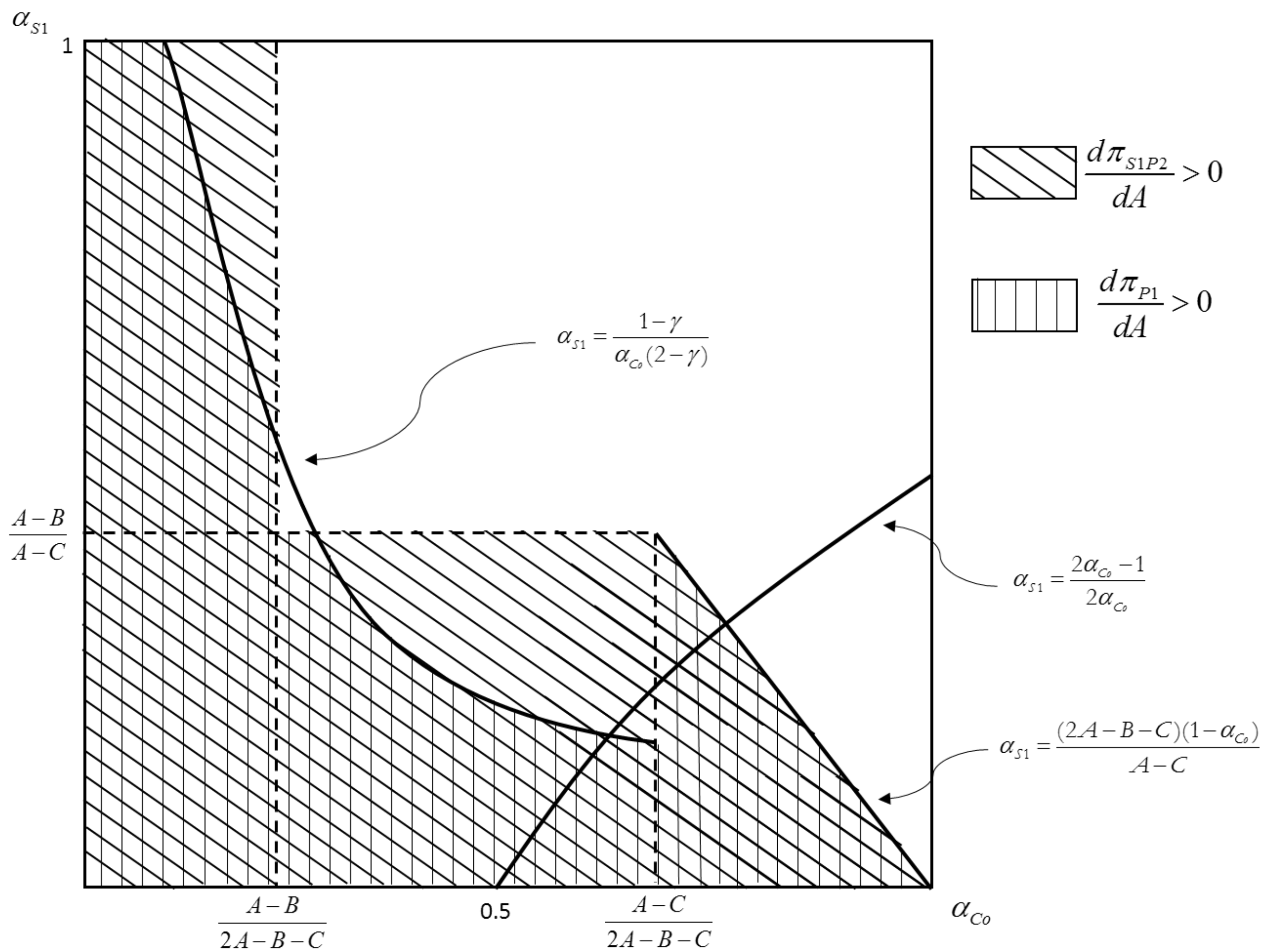
α_{Co}	$\alpha_{Co} \le \frac{A-B}{2A-B-C}$	$\frac{A-B}{2A-B-C} < \alpha_{Co} \le \frac{A-C}{2A-B-C}$	$\alpha_{Co} > \frac{A-C}{2A-B-C}$
α_{S1}	$0 \le \alpha_{S1} \le 1$	$0 \le \alpha_{S1P2} \le \frac{A-B}{A-C}$	$0 \le \alpha_{S1} \le \frac{(2A-B-C)(1-\alpha_{Co})}{A-C}$
S1P2	$\alpha_{S1P2} \ge \frac{2\alpha_{Co}-1}{2\alpha_{Co}}$	$\alpha_{S1P2} \ge \frac{2\alpha_{Co}-1}{2\alpha_{Co}}$	Always
P1	$\alpha_{S1P2} \le \frac{1}{2\alpha_{Co}}$	$\alpha_{S1P2} \le \frac{1}{2\alpha_{Co}}$	$\alpha_{S1P2} \le 2(1-\alpha_{Co})$

Figure A1.5: Impact of larger A on firms' payoffs, integration between S1 and P2.



However, as stated before, S1P2 would always want to adopt the innovation on product $FP_{1,2}$, thus increasing the value of B. Figure 3A takes into consideration this effect.

Figure A1.6: Impact of larger A and B on firms' payoffs, integration between S1 and P2.



Note that Figure 3A depends critically on the value of γ. Figure 3A depicts a value of $\gamma = 0.9$, that is, when P2 adopts the input, it gets only 90 percent of the increase that would be gained by P1 adoption. If both firms gained the same from input adoption $(\gamma = 1)$, then P1 would never benefit from sending information.

Appendix 2: Appendix to chapter 3

Table A2.1: Summary statistics, full sample

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Outside CEO	8,237	0.138	0.345	0	1
Vector change, (prior year)	7,934	0.308	0.315	0	1
Norm change, (prior year)	8,237	0.138	54.74	-2,710	2,112
Vector dist., (mean ind.)	8,237	0.404	0.292	0	1.000
Growth, (2 yrs)	8,105	0.870	17.40	-1	1,338
salary	8,237	682.8	411.6	0	5,500
Patent applications	8,237	104.6	451.8	1	12,271
Avg. num. claims	8,237	19.92	9.584	0	198
R&D intensity	7,177	0.347	8.520	0	509.4

Table A2.2: Outside CEO decrease change in innovation direction, full sample

VARIABLES	(1) Vector change (prior year)	(2) Vector change (prior year)	(3) Vector change (prior year)	(4) Vector change (prior year)	(5) Vector change (prior year)	(6) Vector change (prior year)	(7) Vector change (prior year)
R&D intensity	-0.000412***	-0.000441***	-7.58e-05	-0.000411***	-0.000520***	-0.000470**	-0.000472**
	(0.000139)	(0.000131)	(0.000221)	(0.000138)	(0.000178)	(0.000189)	(0.000187)
salary	7.98e-06	1.60e-05	2.51e-05	7.19e-06	7.04e-06	1.92e-05	1.93e-05
	(1.39e-05)	(1.23e-05)	(1.67e-05)	(1.35e-05)	(1.35e-05)	(1.39e-05)	(1.39e-05)
Avg. num. claims	-0.000747	-0.000827*	-0.00113*	-0.000760	-0.000743	-0.000575	-0.000578
	(0.000473)	(0.000440)	(0.000605)	(0.000473)	(0.000473)	(0.000480)	(0.000479)
Vector dist., (mean ind.)	0.485***	0.486***	0.482***	0.484***	0.484***	0.481***	0.483***
	(0.0287)	(0.0280)	(0.0249)	(0.0284)	(0.0284)	(0.0317)	(0.0316)
Patent applications	-1.83e-05	-2.39e-05*	-5.28e-05**	-2.06e-05*	-2.05e-05*	-1.84e-05*	-1.84e-05*
	(1.14e-05)	(1.28e-05)	(2.68e-05)	(1.12e-05)	(1.12e-05)	(1.12e-05)	(1.11e-05)
Growth, (2 yrs)	-0.000114*	-0.000126*	-0.000247***	-0.000133**	-0.000127*	-0.000121**	-0.000122**
	(6.70e-05)	(6.59e-05)	(8.73e-05)	(6.48e-05)	(6.48e-05)	(5.79e-05)	(5.69e-05)
Growth, (prior yr)					-0.00108***	-0.00102***	-0.00102***
					(0.000195)	(0.000180)	(0.000179)
Outside CEO		-0.0321***	-0.0388**	-0.0318***	-0.0320***	-0.0308***	-0.0307**
		(0.0114)	(0.0164)	(0.0116)	(0.0116)	(0.0119)	(0.0119)
Herfindal Central.						0.0316	
						(0.0202)	
CSA Concent.							0.0286
							(0.0203)
Observations	5,828	5,828	5,894	5,828	5,827	5,144	5,144
R-squared	0.621	0.619	0.271	0.622	0.622	0.628	0.628
Firm FE	YES	YES	NO	YES	YES	YES	YES
Year FE	YES	NO	YES	YES	YES	YES	YES
Num. firms	576	576	642	576	576	485	485

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2.3: Outside CEO impact on norm and firm uniqueness, full sample

VARIABLES	(1) Norm change (prior year)	(2) Vector distance (mean ind.)
R&D intensity	0.00271	0.000608***
	(0.0176)	(0.000142)
salary	-0.00284	-5.26e-07
	(0.0100)	(1.05e-05)
Avg. num. claims	0.0841**	-0.000333
	(0.0366)	(0.000374)
Vector change, (mean ind.)	4.854**	0.305***
	(2.154)	(0.0326)
Patent applications		-3.69e-05**
		(1.69e-05)
Growth, (2 yrs)	0.0109	-0.000128***
	(0.00791)	(4.38e-05)
Outside CEO	9.829**	-0.0178*
	(4.551)	(0.0107)
Observations	5,828	5,828
R-squared	0.026	0.759
Firm FE	YES	YES
Year FE	YES	YES
Num. firms	576	576

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2.4: CEO impact on innovation direction is higher in centralized firms, full sample

VARIABLES	(1) Vector change (prior year)	(2) Vector change (prior year)	(3) Vector change (prior year)	(4) Vector change (prior year)
R&D intensity	-0.000546***	-0.000201	-0.000554***	-0.000199
	(9.99e-05)	(0.000253)	(8.69e-05)	(0.000259)
salary	1.29e-05	1.39e-05	-2.88e-06	1.90e-05
	(2.40e-05)	(1.70e-05)	(2.36e-05)	(1.79e-05)
Avg. num. claims	-0.000635	-0.000875	-0.000864	-0.000690
	(0.000724)	(0.000619)	(0.000696)	(0.000635)
Vector change, (mean ind.)	0.489***	0.477***	0.488***	0.478***
	(0.0444)	(0.0370)	(0.0460)	(0.0353)
Patent applications	-9.59e-05*	-1.40e-05	-3.93e-06	-2.35e-05
	(5.17e-05)	(9.86e-06)	(1.60e-05)	(1.57e-05)
Growth, (2 yrs)	-0.000954**	-5.74e-05*	-0.000499**	-8.34e-05*
	(0.000462)	(2.99e-05)	(0.000253)	(4.63e-05)
Outside CEO	-0.0588***	-0.0123	-0.0520***	-0.0140
	(0.0213)	(0.0136)	(0.0180)	(0.0149)
Observations	2,222	3,606	2,462	3,366
R-squared	0.616	0.622	0.634	0.616
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Centralization	High Herfindal Central.	Low Herfindal Central.	High CSA Concent.	Low CSA Concent.
Num. firms	272	304	281	295

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2.5: Outside CEO and norm change, centralized firms, full sample

VARIABLES	(1) Norm change (prior year)	(2) Norm change (prior year)	(3) Norm change (prior year)	(4) Norm change (prior year)
R&D intensity	-0.000239	-0.0122	0.00387	-0.00862
	(0.00857)	(0.0624)	(0.0100)	(0.0653)
salary	0.000648	-0.00334	0.000458	-0.00472
	(0.00262)	(0.0155)	(0.00294)	(0.0167)
Avg. num. claims	0.0362	0.125*	0.0641	0.107*
	(0.0440)	(0.0639)	(0.0552)	(0.0557)
Vector change, (mean ind.)	3.821***	5.960	6.226	4.099*
	(1.391)	(3.670)	(4.078)	(2.114)
Growth, (2 yrs)	0.0124	0.0127	0.0708	0.00847
	(0.0119)	(0.00978)	(0.0613)	(0.00793)
Outside CEO	3.115	12.35*	6.563***	11.59
	(1.982)	(6.693)	(2.489)	(7.362)
Observations	2,222	3,606	2,462	3,366
R-squared	0.078	0.030	0.074	0.023
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Centralization	High Herfindal Central.	Low Herfindal Central.	High CSA Concent.	Low CSA Concent.
Num. firms	272	304	281	295

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2.6: Outside CEO and distance from industry, centralized firms, full sample

VARIABLES	(1) Vector distance (mean ind.)	(2) Vector distance (mean ind.)	(3) Vector distance (mean ind.)	(4) Vector distance (mean ind.)
R&D intensity	0.000550***	0.000932***	0.000515***	0.000918***
	(6.98e-05)	(8.58e-05)	(6.77e-05)	(8.25e-05)
salary	-7.62e-06	4.44e-06	-9.59e-06	7.09e-06
	(1.79e-05)	(1.29e-05)	(1.38e-05)	(1.42e-05)
Avg. num. claims	-0.000121	-0.000538	-0.000353	-0.000332
	(0.000601)	(0.000483)	(0.000589)	(0.000480)
Vector change, (mean ind.)	0.206***	0.381***	0.226***	0.363***
	(0.0501)	(0.0406)	(0.0495)	(0.0420)
Patent applications	-8.33e-05	-3.18e-05**	-3.96e-05**	-3.37e-05*
	(7.26e-05)	(1.52e-05)	(1.96e-05)	(2.01e-05)
Growth, (2 yrs)	-0.000196	-0.000121**	-0.000127	-0.000121**
	(0.000139)	(4.99e-05)	(0.000117)	(4.76e-05)
Outside CEO	-0.0192	-0.0129	-0.0285	-0.00833
	(0.0262)	(0.0102)	(0.0222)	(0.0114)
Observations	2,222	3,606	2,462	3,366
R-squared	0.725	0.779	0.757	0.764
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Centralization	High Herfindal Central.	Low Herfindal Central.	High CSA Concent.	Low CSA Concent.
Num. firms	272	304	281	295

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2.7: CEO impact on innovation direction is higher when they come from distant firms, full sample

VARIABLES	(1) Vector distance (mean ind.)	(2) Vector distance (mean ind.)
R&D intensity	-0.000386***	-0.0178***
	(0.000123)	(0.00235)
salary	1.65e-05	-9.21e-07
	(1.81e-05)	(1.61e-05)
Avg. num. claims	-0.000840	-0.000563
	(0.000535)	(0.000882)
Vector change, (mean ind.)	0.487***	0.409***
	(0.0341)	(0.0615)
Patent applications	-5.10e-05*	-8.64e-06
	(2.92e-05)	(6.03e-06)
Growth, (2 yrs)	-0.000140**	-0.0110
	(6.06e-05)	(0.0114)
Outside CEO	-0.0404**	0.0156
	(0.0160)	(0.0127)
Observations	4,100	1,696
R-squared	0.618	0.701
Firm FE	YES	YES
Year FE	YES	YES
Previous firm distance	High Distance previous firm	Low Distance previous firm
Num. firms	538	211

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2.8: Outside CEO and norm change, previous firm distance, full sample

VARIABLES	(1) Norm change (prior year)	(2) Norm change (prior year)
	(0.0146)	(0.756)
salary	0.00310	0.00875
	(0.00345)	(0.0253)
Avg. num. claims	0.00277	0.467*
	(0.0251)	(0.241)
Vector change, (mean ind.)	0.587	22.86*
	(2.700)	(11.79)
Growth, (2 yrs)	0.0127	2.417
	(0.00921)	(4.910)
Outside CEO	2.793	14.94**
	(2.002)	(7.548)
Observations	4,100	1,696
R-squared	0.159	0.041
Firm FE	YES	YES
Year FE	YES	YES
Previous firm distance	High Distance previous firm	Low Distance previous firm
Num. firms	538	211

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2.9: Outside CEO and distance from industry, previous firm distance, full sample

VARIABLES	(1) Vector distance (mean ind.)	(2) Vector distance (mean ind.)
R&D intensity	0.000594***	0.00869***
	(0.000102)	(0.00185)
salary	1.01e-05	-2.27e-05
	(1.78e-05)	(1.48e-05)
Avg. num. claims	-0.000226	-0.000175
	(0.000378)	(0.000924)
Vector change, (mean ind.)	0.239***	0.254***
	(0.0415)	(0.0624)
Patent applications	-7.99e-05**	-1.83e-05*
	(3.99e-05)	(9.54e-06)
Growth, (2 yrs)	-0.000141***	-0.00482
	(3.98e-05)	(0.00903)
Outside CEO	-0.0316**	-0.00878
	(0.0138)	(0.0228)
Observations	4,100	1,696
R-squared	0.772	0.774
Firm FE	YES	YES
Year FE	YES	YES
Previous firm distance	High Distance previous firm	Low Distance previous firm
Num. firms	538	211

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2.10: Impact of outside CEO for high and low R&D intensity firms, full sample

VARIABLES	(1) Vector change (prior year)	(2) Vector change (prior year)	(3) Norm change (prior year)	(4) Norm change (prior year)	(5) Vector change (mean ind.)	(6) Vector change (mean ind.)
R&D intensity	-0.000429***	1.283**	0.0143	-103.8	0.000561***	0.560
	(0.000128)	(0.616)	(0.0157)	(117.6)	(0.000144)	(0.475)
salary	2.83e-05	-1.09e-05	0.00132	-0.00585	-1.35e-05	1.62e-05
	(2.04e-05)	(1.84e-05)	(0.00493)	(0.0169)	(1.74e-05)	(1.08e-05)
Avg. num. claims	-0.000402	-0.00147**	0.0909*	0.0624	-0.000163	-0.000691
	(0.000623)	(0.000696)	(0.0536)	(0.0499)	(0.000445)	(0.000653)
Vector change, (mean ind.)	0.406***	0.554***	1.386	8.355***	0.308***	0.282***
	(0.0395)	(0.0416)	(3.641)	(2.534)	(0.0437)	(0.0469)
Patent applications	-1.87e-05	-1.79e-05*			-4.99e-05***	-3.59e-05*
	(2.14e-05)	(1.04e-05)			(1.78e-05)	(1.93e-05)
Growth, (2 yrs)	-0.000107*	-0.00521***	0.00493	-0.118	-0.000119***	-0.000778
	(5.84e-05)	(0.000601)	(0.00797)	(0.0862)	(4.19e-05)	(0.000596)
Outside CEO	-0.0267**	-0.0315	2.979	18.59**	0.00142	-0.0377**
	(0.0130)	(0.0207)	(2.002)	(9.100)	(0.0130)	(0.0187)
Observations	3,178	2,650	3,178	2,650	3,178	2,650
R-squared	0.560	0.639	0.080	0.023	0.771	0.753
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
R&D intensity	High	Low	High	Low	High	Low
Num. firms	305	271	305	271	305	271

Note: Standard errors clustered by firms in parentheses, *** p<0.01, ** p<0.05, * p<0.1