Broad Cross-license Negotiations [1]

Alberto Galasso
University of Toronto

December 7, 2011

[1] This is a revised version of a chapter of my PhD thesis defended at the London School of Economics. I thank Mark Schankerman for his guidance and for sharing data on patent litigation. I am particularly grateful to a Co-editor and two referees whose insightful comments led to a substantial improvement of the paper. I also thank Andrea Prat, Philippe Schmidt-Dengler, Saul Lach, Ronny Razin, Pierre Regibeau, John Van Reenen and seminar participants at Georgia Tech, Essex, Imperial College, Stony Brook, South Carolina, Wisconsin, LSE, Toronto, HEC Montreal, Alicante, Arizona and the European Patent Office for helpful suggestions. Brunna Parreiras provided excellent research assistance.

Abstract

In many industries, broad cross-license agreements are considered a useful method to obtain freedom to operate and to avoid patent litigation. In this paper I study firm incentives to sign a broad cross-license as well as the duration of broad cross-license negotiations. I develop a model of bargaining with learning, which predicts that two firms will enter a broad cross-license agreement only if their capital intensities are large enough. The model also predicts faster negotiations when firms have high capital intensities and when the frequency of future disputes is low. I confirm these predictions empirically using a novel dataset on cross-licensing and litigation in the U.S. semiconductor industry.

Keywords: cross-license, semiconductors, patent litigation, bargaining.

JEL Codes: 031, 032, C78, L63.

1 Introduction

During the past few years, various scholars and industry representatives have drawn attention to specific inefficiencies generated by the patent system in several industries. In particular, Shapiro (2001) has argued that a "patent thicket" has appeared that renders it difficult to commercialize a new technology. In some industries the number of intellectual property rights a firm requires to produce a new product is so large, and their ownership is so dispersed, that it is quite easy to infringe on a patent unintentionally. In this environment there is, therefore, a hold-up problem; when a manufacturer starts selling its product, a patentee might show up threatening to shut production down unless it is paid high royalties.

This issue's relevance is indicated by the endogenous reaction taken by firms operating in industries where the thicket is especially severe. In fact, various business arrangements enabling firms to cut through the thicket have appeared. The objective of this paper is to analyze one of these arrangements: broad cross-license agreements.[1]

A cross-license agreement is a contract between two companies that grants each the right to practice the other's patents. In other words, it is a bilateral agreement in which two firms choose not to enforce intellectual property rights between them. This paper focuses on broad cross-licensing, i.e., on agreements covering the entire patent portfolios or patents in some extensive technology class. Broad cross-licensing is a common practice in industries involved in "cumulative systems technologies" (e.g., computers and semiconductors), where products combine many patentable technologies and where it is easy to unintentionally infringe on a patent (Grindley and Teece, 1997; Shapiro, 2001; Arora et al., 2008). Press releases and companies' annual reports indicate that these contracts are especially widespread in the semiconductor industry, where major industry participants such as Intel, Texas Instruments and AMD are involved in a number of broad cross-licenses.[2] Despite their prevalence, broad cross-license agreements have so far attracted little empirical and theoretical attention from economists compared to other arrangements used to transfer intellectual property (e.g., patent licensing or patent pools). A study of these contracts that lie somewhere between arm's length contracting and firms full integration may help to improve our understanding of the costs

[1] Other possible business arrangements include patent pools. Their relevance and their efficiency properties are analyzed in Lerner and Tirole (2004).

[2] My data show that, of the ten companies with the largest R&D expenditure for the period 1990-1995 (carrying out more than sixty percent of the R&D expenditure of the industry), eight signed at least one agreement between 1990 and 2000.

associated with enforcing patent rights.

I develop a model in which two firms are involved in a series of patent infringement disputes. In the absence of a broad cross-license agreement, these disputes are litigated and the infringing firm stops producing if the court finds it liable. I assume that, when halting production, the infringer sustains a cost that increases with its capital intensity. If two firms sign a broad cross-license agreement, they commit to no litigation of present and future infringements and, thereafter, share duopoly profits. Firms reach an agreement for a cross-license contract through a bargaining procedure in which they learn the frequency of future intellectual property disputes. The model shows that both the decision to cross-license and the timing of the agreement depend crucially on firms' capital intensities. The main prediction is that two firms will sign a cross-license only if their capital intensities are large enough. Intuitively, a broad cross-license agreement is costly to a firm because it reduces each monopoly surplus from innovation to a duopoly surplus. On the other hand, a firm benefits from such a contract because it avoids the loss associated with discontinuing production. Because this benefit increases with firms' capital intensities, a cross-license agreement is profitable only for firm pairings with high capital intensities. In addition, the model predicts that broad cross-license negotiations will be shorter for firm pairings with high capital intensities and for firm pairings facing a low frequency of infringements. This occurs because firms have an incentive to delay the agreement in order to obtain additional information on the likelihood of future disputes. However, waiting is costly because it involves litigation that is particularly harmful for firms with high capital intensities. The model also suggests that broad cross-license agreements are more likely when the patent portfolios of the two firms include a lot of complementary technologies.

I test the predictions of the model using a new dataset that combines information on broad cross-license agreements and patent litigation in the semiconductor industry. The main empirical findings are consistent with the theoretical model and can be summarized as follows. First, high capital intensity increases the likelihood that firms will sign broad cross-license agreements and decreases the duration of licensing negotiations. Second, broad cross-license agreements will take longer to negotiate when firms patent in similar technology areas, which increases the likelihood of future infringement. I also find that firms are more likely to enter broad cross-license agreements when their patent portfolios are complementary, as indicated by a high frequency of cross-citations between the patents of the two firms.

The paper is organized as follows. Section 2 discusses links with previous literature.

Section 3 describes the main features of the semiconductor industry. Section 4 presents the model and the empirical predictions. Section 5 describes the data and presents the econometric results. Section 6 summarizes and concludes.

2 Related Literature

This paper is connected to various strands of the literature in economics of innovation. The first is theoretical research on cross-licensing. Fershtman and Kamien (1992) analyze a two-firm innovation race for two complementary technologies required to produce a product. They compare the case in which both firms need to develop the two technologies and the case of cross-license in which each firm needs to develop only one of the two technologies. Marjit et al. (2001) study how the probability of infringement, R&D costs and product market competition affect the likelihood of firms to avoid patent litigation in case of mutual infringements. Choi (2003) compares private and social incentives to cross-license in a setting where firms are uncertain about the validity of their patents. My model differs from these papers in two main aspects. First, I consider repeated interaction between firms (in the above papers firms face only one infringement problem). Second, I study broad cross-license negotiations and identify settings in which agreement is difficult to reach (in previous literature agreement is always immediate).

My paper is also closely connected to the empirical literature on patent licensing and patent litigation. Lerner (1995) shows that firms with high litigation costs take greater precautions to avoid litigation. Lanjouw and Schankerman (2001, 2004) and Somaya (2003) compare litigation and settlement rates across different industries. Galasso and Schankerman (2010) study how patent thickets and the strength of patent law affect the duration of settlement negotiations.

Hall and Ziedonis (2001) document how capital-intensive semiconductor firms (most vulnerable to hold-up) substantially increased their patenting activity in the '80s. One explanation for this finding is that capital-intensive firms enter patent portfolio races to increase their bargaining power in case of cross-license negotiations. My empirical findings support this conjecture and indicate that capital-intensive firms are indeed more likely to enter a broad cross-license agreement.

Hall and Ziedonis (2007) study the link between the 1982 pro-patent shift in the legal environment and the level of litigation in the semiconductor industry. They find little evidence that semiconductor firms adopted a more aggressive stance toward patent enforcement after the '80s. My paper documents a widespread use of broad cross-license agreements among firms that are most vulnerable to hold-up. This finding can explain why there was little change in litigation level among semiconductor firms.

Siebert and Von Gravenitz (2010a, 2010b) look at licensing contracts in the semiconductor industry and distinguish between ex-ante and ex-post contracts. They define ex-ante licensing as contracts showing a commitment to provide a certain technology before the technology has been explored and ex-post contracts as contracts granting rights for technologies already in place. Siebert and Von Gravenitz (2010a) show that ex-ante licensing is more likely to occur in the presence of blocking patents and with soft product market competition. Siebert and Von Gravenitz (2010b) show that patent thickets and experience also play an important role in the choice between licensing ex-ante and licensing ex-post. My paper complements the analysis of Siebert and Von Gravenitz (2010a, 2010b) by examining broad cross-license deals. These contracts have both an ex-ante and an ex-post nature and therefore are difficult to map in the framework of Siebert and Von Gravenitz (2010a, 2010b).

Finally, my analysis of broad cross-license negotiations is related to recent literature on negotiations and cooperative arrangements in industrial organization. (See Gans and Inderst, 2007 for a detailed survey.) Among these papers, Gans and Stern (2000, 2003) study the interplay of competition, negotiations and cooperation in the "market for ideas."

3 Industry Background

The semiconductor industry consists of companies that design and manufacture computer chips and other electronic devices made of silicon. Semiconductor manufacturing is a significant part of the U.S. economy, accounting to 1.7 percent of the U.S. GDP in 2008 (Bureau of Economic Analysis, 2010).

The semiconductor vertical chain involves a large number of production steps. While a number of key industry players (e.g., Intel) complete the entire process in-house, companies frequently outsource part of their production activities (Turley, 2003). At the extremes of the vertical chain spectrum, there are the "fab-less" companies that specialize in chip design and the "pure-play" foundries that manufacture chips without designing them. While most of the semiconductor firms position themselves at intermediate levels of vertical specialization, there is a trend toward organizational separation of semiconductor design from chip manufacturing (Monteverde, 1995).

The manufacture of semiconductor devices requires use of "clean rooms," where airborne particles are minimized and temperature, humidity and pressure are strictly controlled. Manufacturing equipment consists of two broad categories: *front-end* equipment - installed in clean rooms and used to make the silicon wafers and create the semiconductor chips - and *back-end* equipment - used to assemble, package and test the devices (U.S. International Trade Commission, 2006).

Front-end equipment is typically installed and configured to accommodate the production of a specific semiconductor device. Because minor changes in clean room operations may affect the sensitive equipment and result in damaged products (Turley, 2005), reconfiguring front-end equipment to accommodate changes in manufacturing is extremely expensive, to the extent that construction of a new fabrication facility (aka fab) is often a lower cost alternative (Theron et al., 1999).

Fabs are expensive and have a short life-span. A typical new manufacturing facility costs at least $3 billion and the economic life of manufacturing equipment is about three years (SIA, 2006). In addition, the production of semiconductor devices requires moving down a steep learning curve (Macher and Mowery, 2003). The vast majority of the chips that are produced in the first weeks of operation of a new fab are usually damaged. Over time the quality of the production improves, but it may take up to one entire year of non-stop operations to have production yields above 90 percent (Turley, 2005).

In short, clean rooms and front-end manufacturing equipment are expensive, difficult to redeploy and involve steep learning curves. These features of the semiconductor industry explain why the losses incurred when halting or altering production processes are more detrimental for firms that own and operate complex manufacturing facilities than for firms that do not own such facilities. In turn, this implies that patent litigation and preliminary injunctions are more costly for firms with large investments in manufacturing facilities. This idea is supported by the findings of Hall and Ziedonis (2001) and Ziedonis (2003) that document how capital intensive semiconductor firms tend to react to the risk of patent litigation by amassing large patent portfolios to improve their ex-post bargaining position in the event of settlement negotiation.

4 Theory

In this section, I develop a model of broad cross-license negotiations that borrows from the bargaining framework of Yildiz (2003, 2004). Reflecting the uncertain boundaries of intellectual property rights, firms in the model disagree about the chances of future patent infringements and learn the relative value of each other's patents during negotiations.[3]

4.1 Model

Consider a setting with two firms, $N = \{1, 2\}$, discrete time and infinite horizon. Firms discount the future with a common discount factor, $\delta < 1$. At any date, $t \geq 0$, one of the two firms (the patentee) experiences an infringing action by the other firm (the infringer). In the absence of a broad cross-license agreement, the parties resolve the dispute in court, each incurring a litigation cost of L. If the infringer is found not liable, the two firms share the market for that period, each deriving duopolistic profits, π^D. Conversely, if the infringer is found liable, then in that period the patentee obtains monopoly profits, π^M, and the infringer shuts its production down.

I indicate with kf the cost that the infringer sustains to stop producing for one period where $f > 0$ is a constant and the parameters $k \geq 0$ is a measure of the firm's capital intensity. The function kf captures the first key idea of this model: *the cost associated with stopping production increases with firms' capital intensities.* There are a number of empirical studies that provide evidence in support of this assumption. In particular, the findings of Hall and Ziedonis (2001) and Ziedonis (2003) indicate that the losses incurred when halting or altering production processes are more detrimental for firms investing intensively in product-specific manufacturing facilities.

For notational convenience, I assume that when a trial takes place the infringer is found liable with probability $1/2$.[4] Therefore, the patentee's expected payoff from the litigation period

[3] Two-sided asymmetric information models of non-cooperative bargaining typically have a very large number of equilibria and, thus, are not well suited to derive testable implications (among others see Chatterjee and Samuelson, 1983 and Crampton, 1984; 1992). Often these models have a continuum of possible equilibria, and it is not clear what refinement should be used (see the discussion in Crampton, 1984). An advantage of Yildiz's model is that it yields unique equilibria in the duration of the negotiations, thereby permitting testable implications.

[4] This is for simplicity. All the results can be shown to be valid as long as the infringer is found liable with a positive probability.

is

$$\overline{u} = \left[\frac{1}{2}\pi^M + \frac{1}{2}\pi^D\right] - L, \tag{1}$$

whereas the expected payoff for the infringer is

$$\underline{u}(k) = \left[\frac{1}{2}\pi^D - \frac{1}{2}kf\right] - L. \tag{2}$$

A micro-foundation of the cost kf can be provided assuming that in each period the firm spends f in equipment and that a fraction, k, of this expenditure consists of clean rooms and front-end tools that cannot be redeployed in the case of an unexpected stop in production. The remaining fraction, $1-k$, of the equipment can be resold if production is halted. Indicating with R^D the firm's duopolist net revenue gross of investment cost (i.e., $\pi^D = R^D - f$), the infringer's expected payoff from litigation is

$$\left[\frac{1}{2}R^D + \frac{1}{2}(1-k)f\right] - f - L,$$

which reduces to $\underline{u}(k)$. Notice that under this micro-foundation the cost kf captures the opportunity cost that firms face when halting production and that this cost is greater for capital-intensive firms.[5]

At each point in time a broad cross-license agreement can be signed. If this happens, the two firms do not litigate any intellectual property disputes in the following periods. In addition, the contract specifies the share of future duopolistic surplus ($2\pi^D$ per period) that each firm will obtain. I assume that in negotiating a cross-license agreement the patentee makes a fixed-fee take-it-or-leave-it offer, z.[6]

The identity of patentee and infringer are determined by the following stochastic process: In each period, the probability that a patent owned by firm 1 is infringed by firm 2 is ρ, and the probability that firm 1 infringes a firm 2 patent is $1-\rho$. I assume that in the first period the two firms do not know ρ and that they have different priors about it. This difference in prior beliefs captures the second key idea of the model: *It is difficult for firms to agree on the likelihood of future intellectual property disputes.*[7]

---

[5] I thank a co-editor for suggesting this micro-foundation. All the results in the paper hold for general cost functions $C(k) > 0$ with $C'(k) > 0$.

[6] I assume that firms do not adopt more sophisticated licensing mechanisms because in my sample of broad cross-license agreements, I do not observe any use of royalties or other forms of payments contingent on future usage. I conjecture that this is related to scrutiny from the Department of Justice.

[7] This is a natural assumption for complex technology areas, such as semiconductors, where a single product

As noted above, I model this learning process following Yildiz (2004) and assume that prior beliefs have beta distributions. Fixing any positive integers $\overline{m}_1$, $\overline{m}_2$ and n with $1 \leq \overline{m}_2 \leq \overline{m}_1 \leq n-2$, I assume that for any given dates t and s with $s \geq t$, at the beginning of date t if a firm i observes that firm 1 patents have been infringed m times (and firm 2 patents have been infringed $t-m$ times), then it assigns probability

$$\frac{\overline{m}_i + m}{t+n} \tag{3}$$

to the event that firm 1 patents will be infringed at date $s \geq t$. This belief structure arises when each player believes that the probability of firm 1 patents being infringed at any date t is identically and independently distributed with some unknown parameter ρ that is distributed with a beta distribution with parameters $\overline{m}_i$ and n. As in Yildiz (2004), I assume that everything about this belief structure is common knowledge.

Because $p_t^i(m)$ denotes the probability that firm i assigns at (m,t) to the event that it will be infringed at date $s \geq t$, firm i thinks at (m,t) that the probability that firm j will be infringed at date s is $1 - p_t^i(m)$ while firm j thinks that its patents will be infringed with probability $p_t^j(m)$.[8]

Define

$$y_t(m) = p_t^1(m) + p_t^2(m) - 1.$$

Because of the beta distribution:

$$y_t(m) = \frac{\overline{m}_1 - \overline{m}_2}{t+n} = \frac{\Delta}{t+n} \geq 0.$$

The parameters $\overline{m}_1$, $\overline{m}_2$ and n pin down the initial beliefs of the two players: firm 1 believes that in the future it will be infringed with probability $\overline{m}_1/n$, whereas firm 2 believes that firm 1 will be infringed with probability $\overline{m}_2/n$. When $\overline{m}_1 > \overline{m}_2$, the players have different priors: firm 1 believes that it will be next period patentee with a greater probability than the

may potentially infringe on a large number of patents. Among others, Grindley and Teece (1997) document how hard it is for semiconductor firms to predict future infringements when assessing the value of a potential cross-licensing candidate. To this end, firms perform extensive reverse engineering of the other's product (Texas Instruments engineers call this procedure "reading" the patents on the infringer's products). Because reverse engineering of a semiconductor product is a very demanding task (it takes 400-500 man-hours per device on average), it may take several months for firms to learn the relative value of each other's patent portfolios.

[8] Formula (3) implies that $p_t^1(m) = (\overline{m}_1 + m)/(t+n)$ and $p_t^2(m) = 1 - (\overline{m}_2 + m)/(t+n)$.

one firm 2 assigns to the same event. This means that firm 1 thinks that firm 2 is optimistic and vice versa. Notice that the greater the difference $\overline{m}_1 - \overline{m}_2$, the greater the level of perceived optimism. For this reason, following Yildiz (2008) I refer to $y_t(m)$ as the level of optimism at time t. This implies that the initial level of optimism is $y_o(m) = \Delta/n$. Because $y_t(m)$ is deterministic (it does not depend on m) I will refer to it as y_t. As t gets larger, firms' beliefs converge and firms learn the actual ρ. The speed of convergence decreases with the parameter n (i.e., n measures the firmness of the players' prior beliefs).

Figure 1 describes the timing of the game. At $t = 0$, nature identifies the first period patentee and firms update their beliefs. The chosen patentee offers the infringer a broad cross-license agreement. If the infringer accepts this offer the game ends and the firms enjoy the stream of future duopolistic profits according to the share proposed by the patentee. If the offer is rejected, both players receive the litigation payoffs for one period and nature chooses who makes the offer in the following period. Firms observe who is chosen, update their beliefs and the patentee makes an offer. This structure is maintained for an infinite horizon.

4.2 Testable Implications

In the Appendix, I characterize the Markov perfect equilibrium in which firms' strategies only depend on how many times firm 1 patents have been infringed. I show that, when there is no initial disagreement, the bargaining game is characterized by a discontinuity in agreement time: there is either immediate agreement (no delay) or perpetual disagreement (infinite delay). The presence of initial disagreement generates an intermediate case in which (finite) delay is observed. The equilibrium characterization yields four sets of testable implications. All proof are relegated to the Appendix.

The first prediction relates to the characteristics of firm pairings involved in broad cross-licensing. A broad cross-license agreement is costly because it reduces each innovation monopoly surplus to a duopoly surplus. On the other hand, such a contract is beneficial because firms do not incur the costs associated with a stop in production. Because this benefit increases with firms' capital intensities, a cross-license agreement is profitable only for firm pairings with high capital intensities.

Implication 1 (Capital Intensity and Cross-Licensing) *Broad-cross license agreements are signed by firm pairings with high capital intensities.*

Second, the model predicts the length of broad cross-license negotiations. Firms have

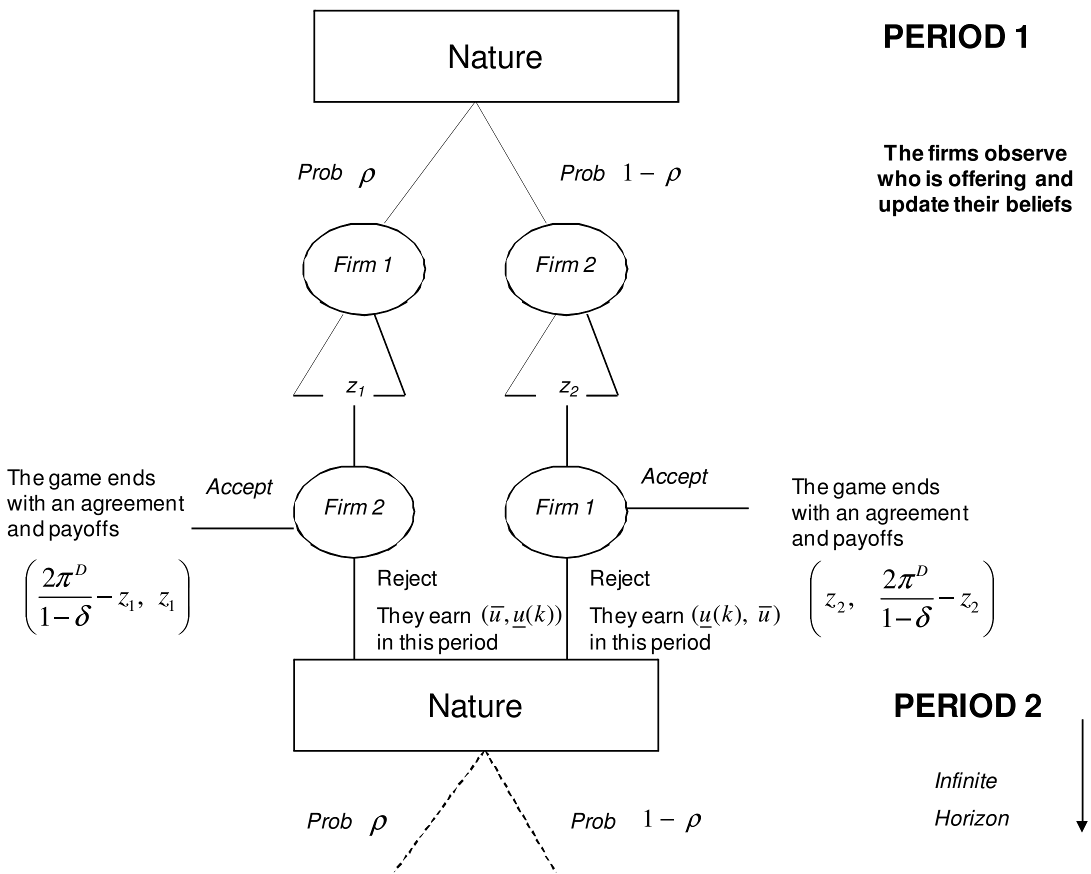

Figure 1: Game Tree

an incentive to delay the agreement in order to obtain additional information on the likelihood of future disputes. Specifically, each firm expects this information to reduce the opponent's optimism and therefore to decrease the opponent's requested surplus share. However, waiting is costly because it involves litigation that is particularly harmful for firms with high capital intensities. As a result, firms with high capital intensities have a lower incentive to delay the agreement.

Implication 2 (Capital Intensity and Negotiation Delay) *Broad cross-license negotiations will be shorter for firm pairings with high capital intensities.*

Third, the model shows that delay is related to the frequency of infringements. As standard in non-cooperative bargaining models, I consider the discount factor as a proxy of the time between successive interactions (Shaked and Sutton, 1984; Binmore et al., 1986); as $\delta \to 1$ the interval between successive infringement events becomes "negligibly small." In deciding whether to delay a broad cross-license agreement, firms face a trade off between a present cost and a future benefit. The present cost is related to the litigation that takes place if the agreement is delayed. The future benefit is related to the better deal that can be obtained in the future. If infringements are not very frequent (small δ), the present cost out-weighs the future benefit and firms have an incentive to agree as soon as possible. If infringements are very frequent (large δ), the present cost is small relative to the future gain and firms will find it profitable to delay the contract.

Implication 3 (Frequency of Infringements and Negotiation Delay) *Broad cross-license negotiations will be longer for firm pairings facing a high frequency of infringements.*

Finally, the model can be used to investigate the interaction between patent portfolio complementarities and broad cross-licensing negotiations. An increase in the degree of complementarity of the patenting activities of the two firms can be modeled as an increase in duopoly profits. If the innovations of the two firms are associated to complement products, the total duopoly profits may actually exceed the monopoly profits ($2\pi^D > \pi^M$). By increasing duopolistic surplus, greater complementarities have two main effects. First, complementarities increase the value of a broad cross-license agreement and therefore the incentives to sign an agreement. Second, when complementarities are strong, i.e. $2\pi^D > \pi^M$, the litigation payoffs will also be affected. If the court finds an infringer liable, the patentee will prefer to license the patent and not to ask for a permanent injunction. In this case, the total expected profits from litigation will be $2\pi^D - 2L$. Because with these litigation profits complementarities have similar effect both on the agreement surplus and on the outside option, they have an ambiguous impact on negotiation delay.

Implication 4 (Portfolio Complementarities and Cross-Licensing): *Broad cross-license agreements are signed by firms with high portfolio complementarities. Portfolio complementarities have an ambiguous impact on negotiation delay.*

4.3 Discussion of the Main Modeling Assumptions

The model builds on a number of assumptions that are worthy of additional discussion. First, I assumed that when the two firms engage in a broad cross-license negotiation, some of their investments are sunk. In the absence of market failures, firms could avoid litigation by signing "narrow" ex-ante deals that completely remove the hold-up problem. Specifically, in each period before sinking the investment, a firm could negotiate a deal obtaining the permission to use the specific patented input required. In the model, I implicitly assumed that those "narrow" ex-ante deals are not always possible and that firms may unintentionally infringe on a patent after their investments have been sunk. This assumption is consistent with the literature on thickets (e.g., Heller and Eisenberg, 1998; Shapiro, 2001) that emphasizes how in industries where firms require a large number of patents and patent ownership is dispersed, the likelihood of unintentionally infringing on a patent is substantial. A firm may sink investment and unintentionally infringe on a patent for a number of reasons. One is the high search costs associated with fragmented patent ownerships. Another is the uncertainty about patent applications and their validity (sub-marine patents). There is also a common practice in information technology firms of instructing their engineers not to read the patents coming out of the USPTO. Lemley and Tangri (2003) and Shapiro (2007) document how experienced lawyers often advise their clients to avoid reading new patents. This practice is followed because firms want to minimize the chances to be found to have willfully infringed and be subject to treble damages.[9]

Second, I assumed that in case of disagreement over a broad cross-license contract, firms resolve their period dispute by litigation. This assumption greatly simplifies the analysis but can be removed. If I introduce the possibility of private settlement, the testable predictions

[9] Moreover, because a broad cross-license mitigates the hold-up problem, firms may have an incentive to increase their asset specificity after having signed a deal. This would imply that firms anticipating a cross-licensing deal could be more inclined to sink their investments than those not anticipating such a cross-license. In the econometric analysis (section 5.3), I use data on lagged asset specificity to control for this possibility.

of the model remain valid as long as in the settlement subgame one of the parties has some private information about factual issues that is relevant to predict the expected outcome of the trial (as in Bebchuck, 1984). Because the presence of asymmetric information assures that there is a positive probability of incurring the cost kf, the main trade-offs described in the model are still present. In the litigation sub-game I also implicitly assumed that injunctions are automatically granted. This is technically convenient but the qualitative results are robust to including probabilistic injunctions. [10],[11]

Third, I assumed that there is no difference in the capital intensity of the two firms. Unfortunately, when $k_1 \neq k_2$ the surplus that the two parties perceive is no longer deterministic and I cannot characterize a unique agreement time. At the same time, I do not expect my model to deliver a clear prediction of the impact of asymmetries in asset specificity. To see this, suppose I take a symmetric firm pairing and I simultaneously increase k_1 and reduce k_2 of the same amount. In this case, if firms believe that firm 2 is more likely to infringe firm 1, the asymmetry is likely to reduce the incentive of cross-licensing and to increase negotiation delay. Conversely, if firms believe that firm 1 is more likely to infringe firm 2, the asymmetry is likely to increase the incentive of cross-licensing and to reduce the length of negotiations. For these reasons, in the empirical analysis I will expect an ambiguous impact of firm asymmetry.

Finally, I discuss the choice of Yildiz (2003; 2004) bargaining framework where delay arises because each firm is optimistic and overestimates the likelihood of being infringed by the other firm. While optimism is a plausible behavioral assumption in innovation settings,[12] the main reason for this modeling strategy is to avoid the multiplicity of equilibria associated with two-sided asymmetric information bargaining models. It is important to notice that the role that overconfidence plays for the testable implications is minimal; the first implication holds even in the absence of optimism and the other three hold (weakly) for any positive level of optimism.

---

[10] The 2006 eBay v. MercExchange Supreme Court decision significantly reduced the likelihood of an injunction and thus reduced the sensitivity of the infringer payoff to capital intensity. The model suggests that after this decision we should observe a lower impact of capital intensity on both the decision to cross-license and on negotiation delay. Unfortunately, I cannot examine the empirical validity of this conjecture because my dataset covers the period 1985-2005 and thus ends before the 2006 decision.

[11] Firms may settle a patent dispute by forming a patent pool (e.g., the Summit and VisX patent pool). The model can be generalized introducing this settlement option as long as one of the parties has some private information about the validity of the patent or the outcome of the trial in case the pool is not formed.

[12] Galasso and Simcoe (2011) study the interplay of managerial overconfidence and innovation.

5 Empirical Analysis

5.1 Data Description

The starting point for the empirical analysis is the set of 218 publicly traded U.S. firms whose principal line of business is semiconductors and related devices (SIC 3674), for which Compustat has data for at least three years between 1985 and 2005.

Examining the SEC annual filings (10-K) and press releases in the companies' web pages, I identified 38 broad cross-license agreements. More precisely, I focused on contracts that satisfied two requirements. First, both contracting parties are semiconductor firms; I exclude agreements between semiconductor firms and firms in other industries. Second, I focused on "broad" cross-license contracts, i.e., agreements in which firms agree not to litigate any of their patents for a period of time or any patent in some very extensive technology areas.[13] An example is Micrel Inc.'s 10-K filing for 2003, which states that "*on May 23, 2002, the Company entered into a Patent Cross License and Settlement Agreement with National Semiconductor which settled all outstanding patent disputes between the companies and cross licensed the entire patent portfolio of each company. Some of the National patents within certain field of use areas are licensed for the life of the patents, all other patents of both companies are licensed through May 22, 2009. Under the terms of the agreement Micrel agreed to pay National $9.0 million.*"

All the agreements are bilateral (i.e., between two firms) and involve 43 different firms. Most of these companies (25 out of 43) sign only one agreement, and a small number of them (five out of 43) sign more than three agreements.[14] Unfortunately, the financial terms of most agreements are not disclosed. In the few cases in which they are disclosed, the terms involve lump-sum payments and not royalties related to sales or future usage of intellectual property.[15]

I use multiple data sources to identify patent disputes between firms in my sample: the

---

[13] In their SEC filings, companies are required to disclose any contract upon which the registrant's business is substantially dependent, as in the case of "license or other agreement to use a patent upon which registrant's business depends to a material extent" (see Standard Instructions for Filing Forms under Securities Act of 1934- regulation S-K).

[14] Specifically, 25 firms sign one agreement, 12 sign two agreements, one signs three agreements, two sign four agreements, two sign five agreements and one signs six agreements.

[15] I do not include in the empirical analysis contracts that were explicitly related to a technology standard or those related to a specific geographic area. The focus on broad licensing deals is an important difference with the papers of Siebert and Von Gravenitz (2010a, 2010b) that look at a more comprehensive set of semiconductor licensing deals. Although their sample is much larger than mine (800 licensing deals in the period 1989-1999), they do not look at contracts associated with patent litigation and consider also narrower licensing deals involving few patents. In addition, they report that only 2 percent of their contracts (roughly 16 contracts) could have been categorized both as ex-ante and ex-post licensing (as a broad cross-license deals).

Litalert database (described in Lanjouw and Schankerman, 2001; 2004), firms SEC annual filings (10-K) and press releases on the companies' web pages. I identify 125 patent disputes filed in the period 1985-2005 in which both plaintiff and defendant are firms in my sample (i.e., all cases are semiconductor specific). Only 34.4 percent of the patent cases were filed in the decade 1985-1995, with the remaining 65.6 percent litigated in the decade 1995-2005. The disputes involve 97 distinct firm pairings.

Because my theoretical model studies broad cross-license negotiations between two firms and all the agreements in my dataset are bilateral, the unit of observation of my analysis is a firm pairing. I register some form of interaction (either a dispute or a cross-licensing agreement or both) for 115 distinct firm pairings.[16] I refer to these pairings as *interacting pairings*. To be consistent with my theoretical framework, in which the two firms face a long series of disputes that conclude in either litigation or cross-licensing, most of my analysis will focus on the sample of interacting pairings. In section 4.3 I discuss possible selection biases and investigate the robustness of the results using a larger sample.

The two main dependent variables used in my empirical analysis are:

Cross-License: A dummy variable defined for all interacting pairings that equals one if there is a broad-cross license agreement between the two firms.

Length: A variable defined only for firm pairings signing a broad cross-license agreement. It measures the duration (in months) of the litigation period preceding a cross-license agreement. It is equal to zero if the agreement is not preceded by litigation and it has a positive value if cross-licensing occurs as settlement of litigation between two companies. When I observe multiple suits (overlapping in time) between two parties before a broad cross-license I measure length from the first case filed.[17]

I construct two right-hand side variables to capture the main elements of the bargaining model (capital intensity and frequency of infringements). I also consider a range of control variables suggested by the existing literature on innovation in the semiconductor industry (in

[16] Mergers and acquisitions are another form of interaction that firms may use to solve their intellectual property disputes. Because it is difficult to evaluate whether an acquisition is carried out only to obtain access to intellectual property, I do not consider M&A in my analysis.

[17] I do not observe any patent litigation event between firm pairing after a broad cross-license agreement is signed. In section 5.2 I consider the possibility of truncation in the data and I construct the variable length for some pairings that litigate but do not cross-license. In this case the variable is equal to the period between the first litigation event and 2006. In the empirical analysis, I will treat these observations as right-censored duration spells.

particular by Hall and Ziedonis 2001; 2007).

Average Capital Intensity: The two firms' capital intensities are measured as the logarithm of the ratio of deflated (2001 dollars) net property, plant and equipment to the number of employees. I follow previous literature (Hall and Ziedonis, 2001; 2007; Ziedonis, 2003) and employ this ratio based on the idea that the losses incurred when halting or altering production processes are more detrimental for firms investing intensively in product specific manufacturing facilities.

Technological Closeness: A measure based on the 426 technology classes (n-classes) provided by the USPTO. Following Jaffe (1986, 1988) and Bloom, Schankerman and Van Reenen (2008) I use the average share of patents per firm in each technology class to construct the vector $c_i = (c_{i1}, c_{i2}, ..., c_{i426})$ describing the distribution of patents of firm i across technological classes (I use the patents granted to the firms in the 10 years before the interaction). The technological closeness measure is calculated as the uncentered correlation between all firm pairings:

$$\text{Technological Closeness}_{ij} = \frac{c_i' c_j}{(c_i' c_i)^{\frac{1}{2}} (c_j' c_j)^{\frac{1}{2}}}$$

and quantifies the proximity between two firm research activities in the 426-dimensional space generated by the USPTO n-classes. Following Lerner (1995), I use similarity in the technology areas of firm patent activities as a proxy for expected frequency of infringements.

Complementarity: A measure constructed from the set of patents obtained by the firm pairings in the 10 years before the first interaction. Having identified these patents for each interacting pairing, ij, I count firm i citations that referred to firm j patents and firm j citations referring to firm i patents and divide these two values by the total number of citations made by the two firms. More formally the measure is:

$$\text{Complementarity}_{ij} = \frac{\#\text{ of citations from } i \text{ to } j + \#\text{ of citations from } j \text{ to } i}{\#\text{ of citations made by firm } i + \#\text{ of citations made by firm } j}.$$

This measure is based on the idea that the greater the complementarity between the portfolios, the more likely that firm i will be engaged in inventions that build on firm j patents and vice versa.

Average Size: The firm size is measured as the logarithm of the number of employees.

Average R&D Intensity: The average of the logarithms of the ratio of R&D spending to employees. This variable captures the importance of knowledge assets to the firm independently on the number of patents obtained.

Average Patent Yield: The yield of each firm is measured as the logarithm of the ratio of its patent stock (constructed using annual patent granted and 15 percent depreciation rate) to R&D spending. It measures the importance of patents in the firm's technology strategy.[18]

Texas Instrument: It is a dummy for pairings involving this firm. This variable is introduced because Hall and Ziedonis (2001, 2007) found that Texas Instruments tends to patent and litigate more aggressively compared to other semiconductor firms.

Design Firm: A dummy equal to one if one of the firms in the pairing (or both) are specialized in design. I classify a firm as a design firm if press-releases or annual reports describe the company as "fabless."[19]

Size Asymmetry: The ratio between the sizes of the larger and the smaller firm in a pairing.

Time Dummies: Following Hall and Ziedonis (2007) I construct dummies for the periods 1985-92, 1993-1997 and 1998-01. I used a different dummy for each year in the period 2002-05.

The dataset is a cross-section and the variables are constructed at the time in which the first interaction is observed. Therefore, for firm pairings signing a cross-license without previous litigation, I used the year of the cross-license deal. For firm pairings signing a cross-license deal as settlement of patent litigation, I used the year in which litigation begins. Finally, for firm pairings litigating but not cross-licensing, I used the year in which the first case is filed.[20]

Table 1 provides summary statistics for the dataset involving the sample of interacting firms (115 pairings). While in the regressions I use a logarithmic representation, Table 1 reports the variables in absolute value. On average, for these pairings the investment in plant and equipment per employee is about $92,000 per firm, employment is approximately 9,400 employees per firm, R&D investment per employee is about $54,000 per firm and the patent

[18] The patent stock for firm i is constructed following Hall and Ziedonis (2007) using a 10 year inventory equation with declining balance depreciation. Specifically $PatentStock_{it} = \sum_{s=0}^{9}(1-\delta)^s PatentApplications_{i,t-s}$ where $PatentApplications_{i,t}$ is the number of (eventually granted) patents applied for in year t by firm i and the depreciation rate, δ, is chosen to be 0.15.

[19] Some of the firms gradually moved from operating their own fabrication facilities to specializing in chip design. This slow process makes it difficult to identify the exact year in which a firm becomes "fabless" and may be a source of measurement error (likely to induce some attenuation bias in the estimates).

[20] The interactions are distributed over time as follows: 23 (4 ended with a broad cross-license) in the period 1985-92; 20(8) in 1993-97; 31(13) in 1998-01; 24(10) in 2002-03 and 17(3) in 2002-05.

stock is 921 patents per firm.

Table 1 also provides similar summary statistics for the 38 cross-licensing firm pairings. Firm pairings involved in cross-licensing have greater average capital intensity and larger size. On average, pairings with cross-licensing deals spend about $123,000 per firm in plant and equipment per employee, whereas those litigating but not cross-licensing spend roughly $76,000 (the difference is significant at the 0.01 level). Moreover, pairings in a cross-licensing deal have on average about 13,000 employees per firm whereas those litigating but not cross-licensing have approximately 7,400 employees per firm (the difference is significant at the 0.01 level).

5.2 Findings

To test the predictions of the model, I use the following probit regression:

$$\begin{aligned} & \Pr(\text{Cross-License}|\ \text{Average Capital Intensity}_{ijt},\ X_{ijt}) \\ = \ & \Phi\left(\beta_1 \text{Average Capital Intensity}_{ijt} + \beta_2 X_{ijt}\right) \end{aligned}$$

where ij denotes the pairing that includes firm i and firm j, t is the time of first interaction between i and j, Φ is the cumulative standard normal distribution and X_{ijt} is a vector of control variables for other factors that affect the decision to cross-license. In this setting, implication 1 entails that $\beta_1 > 0$: broad cross-license agreements are more likely for pairings with high capital intensity. Table 2 reports parameter estimates for this probit model.

In column (1) I include only average capital intensity. The coefficient is positive and statistically significant at the 0.01 level, confirming implication 1: Broad cross-license agreements are signed by firm pairings with high capital intensities. At the mean value of average capital intensity the estimated parameter implies a marginal effect equal to 0.22. This means that a one standard deviation increase in average capital intensity (in logs) increases the probability of a broad cross-license deal by 14.5 percentage points, from 0.32 to 0.465.

In columns (2)-(4), I incrementally add control variables. Column (2) includes technological closeness and column (3) adds all the other controls (but not time dummies). There is almost no change in the estimate for average capital intensity that remains significant at the 5-percent level (p-value= 0.02).

Finally, in column (4) I introduce dummies for the time periods. The coefficient on average capital intensity drops by 7.7 percentage points but remains significant at the 0.05 level. Results from unreported regressions indicate that the choice of time period dummies does not affect results substantially.

The only other variable that appears to be correlated with the decision to sign a broad cross-license agreement is the average patent yield. The marginal effect for the coefficient in column (4) implies that a one standard deviation increase in average patent yield reduces the estimated probability of signing a broad cross-license agreement by 0.12 percentage points, from 0.25 to 0.13. This finding is consistent with the idea that broad cross-license agreements are more appealing when firm strategies do not extensively rely on patents. An alternative explanation is that firms with large patent portfolios face a lower litigation risk and therefore have fewer incentives to enter a broad cross-license agreement.[21],[22]

To test the other two predictions of the model, I follow Galasso and Schankerman (2010) and use the following proportional hazard model with exponential specification:

$$\begin{aligned} \ln h_{ijt} &= \alpha_0 + \alpha_1 \text{Average Capital Intensity}_{ijt} + \alpha_2 \text{Technological Closeness}_{ijt} + \\ &\quad \alpha_3 X_{ijt} + \varepsilon_{ijt} \end{aligned} \quad (4)$$

where ij denotes the pairing that includes firm i and firm j, t is the time of first interaction between i and j, X_{ijt} is a vector of control variables that affect the decision to cross-license and ε_{ijt} is a mean zero random error. In my baseline regression I allow for arbitrary hetheroskedasticity. (Results do not differ when I cluster across interaction periods.) A *negative* coefficient on a regressor in the hazard rate model means that the variable makes it less likely that negotiations end in a given period, which corresponds to a *longer expected delay.* The model implies the following predictions in this specification: Average capital intensity reduces bargaining delay ($\alpha_1 > 0$) and technological closeness increases delay ($\alpha_2 < 0$). The exponential specification imposes a constant (baseline) hazard rate, but the results are nearly identical for the more flexible Weibull specification which allows for an age-dependent hazard rate (Kiefer, 1988).[23] Table 3 reports parameter estimates for this exponential hazard regression.

[21] Lanjouw and Schankerman (2004) show that litigation risk is much higher for firms with small patent portfolios.

[22] In unreported regressions I substituted the patent yield variable with a citation weighted patent stock. The coefficients suggest that broad cross-licensing agreements are more common when firms have highly cited patent portfolios. However, the high correlation between firm size and patent stock makes it difficult to explore this issue more in detail.

[23] The Weibull is a two-parameter distribution with the (baseline) hazard function $h(t) = \lambda\gamma t^{\gamma-1}$. The exponential case arises when $\gamma = 1$. In the baseline econometric specification, the point estimate of γ is 1.36 (s.e. = 0.21), so I cannot formally reject the exponential restriction in favor of the Weibull with an increasing hazard rate.

In column (1) I include only average capital intensity and the measure of technological closeness. The coefficient on a capital intensity is positive and statistically significant at the 0.01 level, confirming implication 2: Broad cross-license negotiations are shorter for firm pairings with high capital intensities. The coefficient on technological closeness is negative and statistically significant at the 0.05 level, confirming implication 3: Broad cross-license negotiations are longer for firm pairings facing a high frequency of infringements.

In column (2) I show that, despite the small number of observations, results are robust to the inclusion of additional controls and time dummies. The estimates in column (2) imply a marginal effect of -11.14 for average capital intensity and a marginal effect of 20.21 for technological closeness. With these marginal effects a one standard deviation increase in average capital intensity reduces negotiation delay by eight months, and a one standard deviation increase in technological closeness increases delay by five months.

There is the concern that some of the firm pairings litigating during the sample period may have signed a broad cross-license agreement after 2005. This is especially likely for cases litigated in the last years of the sample period.[24]

To address this problem, in column (3) I run an exponential hazard regression treating cases litigated in 2004-05 as right-censored spells (Kiefer, 1988). Intuitively, for pairings litigating in the period 2004-05, I consider the litigation period as a truncated negotiation over a broad cross-license agreement and assume that an agreement will be reached after 2005. This exercise increases the number of observations from 38 to 52. I still find support for the two testable predictions: Both the coefficient on capital intensity and the coefficient on technological closeness remain significant and with the right sign.

The only other coefficient that shows up as statistically significant both in column (2) and in column (3) is the one on average R&D intensity. This negative coefficient implies that broad cross-license negotiations last longer when firm pairings have high R&D intensity. This is consistent with the idea that the value of an agreement is greater if knowledge assets are important for firms' technology strategies. In other words, because agreements are more valuable for firms with high R&D intensity, they have greater incentive to delay the agreement and learn more accurately the relative frequency of infringements.

[24] Another concern is that data can be left-censored in the sense that for some interactions I do not observe all the litigation events preceding the cross-license (e.g., an agreement is classified as immediate because I do not have data on previous litigation). If present, this truncation is likely to affect observations at the beginning of the sample period. To address this concern I dropped interactions in the period 85-89; the results from this restricted sample are very similar to those for the full sample.

In column (4) I treat all 2002-05 litigation events as right-censored spells. This exercise increases the number of observations to 66. I continue to find support for the testable predictions of the model but the coefficient on technological closeness is significant only at the 0.1 level. [25]

In Section 4.3 I argued that the theoretical model does not deliver a clear prediction about the impact of asymmetries in capital intensity. Nonetheless, in column (5) I explore the issue empirically introducing a variable (Capital Intensity Asymmetry) that captures the ratio between the larger and the smaller capital intensity in a firm pairing. The coefficient is positive and significant at the 0.1 level and suggests that negotiations are shorter when firms are asymmetric. This finding is consistent with the idea that agreement is easier to reach when there is little uncertainty about which firm suffers the most in case of infringement (i.e., there is less disagreement about which firm gains the most from signing the deal).[26]

Table 4 assesses the empirical validity of the fourth testable implication. In columns (1) and (2) I show that the complementarity measure is positively associated with the likelihood of signing a broad cross-license deal. Columns (3) and (4) examine the relation between complementarity and negotiation delay using an exponential hazard model. The coefficients are statistically insignificant suggesting no meaningful relationship between the measure and negotiation length. The estimates in column (2) imply a marginal effect of 9.39 for the portfolio complementarity measure. This marginal effect indicates that a one standard deviation increase in complementarity is associated with an increase in the probability of cross-licensing by about 0.2 percentage points.

Lerner (1994) argues that the USPTO classification may be noisy at the nclass level and not oriented to industry classification. To address this concern, I constructed a new closeness measure TECH_IPC that exploits the eight edition of the International Patent Classification scheme. Specifically, the distance measure is equal to the uncentered correlation between the vectors of patent shares constructed at the four-character IPC classification level. The other technological closeness measure and TECH_IPC appear highly correlated (the rough

[25] If I treat all the litigation events that do not end in a cross-license agreement as right censored (i.e., all 115 interacting pairs are used), the coefficient on average capital intensity remains positive and significant, whereas the technological closeness coefficient remains negative but loses its statistical significance.

[26] I also introduced separately the larger and the smaller capital intensities in the pair. The coefficient on the larger capital intensity is positive and significant at the 0.01 level. The coefficient on the smaller capital intensity is negative with p-value equal to 0.11. These findings also suggest that negotiations are shorter when there is asymmetry in capital intensity. I did not find any effect of capital intensity asymmetries in the probit regressions.

correlation is equal to 0.84). Unreported regressions indicate that results are qualitatively and quantitatively very similar if I replace the technological closeness measure with TECH_IPC.[27]

5.3 Robustness

This section describes a variety of extensions and robustness checks. The tables associated with these models are presented in the Appendix.

Endogeneity

A potential concern is that the measure of capital intensity may be endogenous. Specifically, it may be the case that firms increase their capital intensity because they anticipate that they will sign a cross-license deal. Moreover, their incentives to sink investments can be greater when they expect a deal not to be associated with long litigation. If this happens, capital intensity will affect the probability of cross-licensing and the length of litigation, but the perspective deals and litigation will also affect capital investments.

To address this concern, I re-estimate the probit and exponential hazard regressions using firms' capital intensities five years before the first interaction is observed. Because the average lifespan of a state of the art fabrication facility is roughly three years (SIA, 2006), it is unlikely that firms sink their investment because they anticipate a cross-license deal after five years. Column (1) in Appendix Table A1 reports parameter estimates for the probit regression. Despite the five-year lag, the coefficient of capital intensity is very similar to the one in column (4) of Table 1 and it is significant at the 0.1 level. Similarly, column (2) shows that the coefficients of the hazard regressions are very similar to those in column (2) of Table 3.[28]

The coefficient on patent yield may also be biased because patenting may be affected by the expectation of a broad cross-license agreement. For example, semiconductor firms may have an incentive to obtain a large patent portfolio before entering a broad cross-license negotiation (Hall and Ziedonis, 2001). Notice that the three testable implications derived from the theoretical framework do not involve the patent yield variable. Because the examination of the impact of a broad cross-licensing agreement on patenting (and R&D) decisions would

[27] Because of the likely measurement error in the design dummy, I re-estimated the econometric models dropping pairings involving design firms. Despite the substantial reduction in sample size, the results are qualitatively similar. Quantitatively, the effects of capital intensity are stronger in this subsample but the differences with the coefficients of the baseline regressions are not statistically significant.

[28] When the five-year lagged capital intensity was not available, I used the oldest available data in the five-year window.

require a different econometric framework (panel data analysis), I leave a more careful study of this topic for future research.[29]

Product Market Competition

A second concern is that capital intensity may be correlated with characteristics of the downstream product market in which these technologies are used that in turn may affect the odds of cross-license or the timing of negotiation. In particular, if firms with high capital intensities operate in product markets with low competitive pressure, these firms may be more likely to cooperate (sign a broad cross-license deal) and less likely to negotiate aggressively for their deals (shorter negotiation delay). To examine this concern, I follow Bloom, Schankerman and Van Reenen (2010) and develop a measure of product market closeness that uses data from the Compustat Segment Dataset. Specifically, I collected information on each firm's sales broken down into four-digit industry codes and constructed the vector $S_i = (S_{i1}, S_{i2}, ..., S_{i597})$, where S_{ik} is the share of sales of firm i in the industry code k. The measure of product market closeness for two firms i and j is then computed as the uncentered correlation:

$$SIC_{ij} = \frac{S_i' S_j}{(S_i' S_i)^{\frac{1}{2}} (S_j' S_j)^{\frac{1}{2}}}.$$

As the measure of technological closeness, the SIC measure ranges between zero and one depending on the degree of product market overlap.

In columns (3) and (4) of Appendix Table A1, I introduce this variable in the probit and the hazard regressions. There is no substantial change in the coefficients of interests (average capital intensity and technological closeness). The coefficients on the SIC measure are not statistically significant.[30]

Logs and Ratios

The main variables used in the empirical analysis have been constructed following previous studies on the patenting in the semiconductor industry (in particular Hall and Ziedonis, 2001; 2007). Most of these variables (capital intensity and R&D intensity, in particular) are

[29] The measure of technological closeness may also be endogenous because the expectation of a cross-license deal may affect firms' patenting decisions. The problem is partially attenuated by the fact that the measure is constructed using patents in a 10-year window. Moreover, results are very similar if I construct the measure dropping the patents obtained the year before the first interaction.

[30] For the semiconductor firms in my sample, the dispersion of sales across product classes is quite low. This suggests that, while the Compustat Segment Data can be very useful in assessing product market distance in samples comprising firms across multiple industry (as in Bloom, Schankerman and Van Reenen, 2010), the data do not offer the level of detail required for a deep exploration of market closeness within industries.

expressed as ratios. In addition, I used a logarithmic specification that is common in duration models because, under some parametric restrictions on the hazard function, the coefficients can be interpreted as elasticities (Kiefer, 1988). To address the concern that results are driven by these specification choices, I present in Table A2 results from regressions using alternative specifications. In columns (1) and (2), I replace capital intensity with net property, plant and equipment and replace R&D intensity with R&D expenditure. In other words, I use the total value of capital and R&D and not the amount per employee. The estimates provide qualitative support to the main testable predictions of the model. In columns (3) and (4) I drop the logarithmic specification. This specification also supports the testable implications of the model.

Sample Selection

Following the theoretical model, I focused the empirical analysis on the sample of firm pairings involved in some observable form of interaction (either litigation or cross-licensing). Selection bias may arise, however, because there may be unobservable characteristics affecting the decision to interact that are correlated with the unobservable characteristics affecting the decision to cross-license and negotiation length. For example, a plausible source of bias is related to firm pairings that do not rely substantially on patents to appropriate their returns to R&D. These firm pairings are more likely to enter a broad cross-license agreement (patent enforcement is not an important component of their strategy) and are less likely to negotiate the agreement aggressively (the low value at stake reduces their incentives to learn the pattern of future infringements). Because high capital intensity firms gain a lot from cross-licensing and low capital intensity firms gain only a little, low capital intensity firms that enter a broad cross-license may be those that do not rely substantially on patents. If this selection takes place, negotiations between firms with low capital intensity will be shorter and I will underestimate the true (negative) impact of capital intensity on broad cross-license negotiations.

Unfortunately, it is extremely likely that any factor that affects the decision to interact with another firm also affects both the choice between litigation and cross-license contract and the duration of broad cross-license negotiations. This makes it very difficult to think of suitable instruments that could be used in a Heckman selection model.

Nevertheless, in Table A3, I report results from two estimations using Heckman's selection correction including the same variables in the second stage and in the interaction first stage. This regression is identified only by functional form (the nonlinearity of the first stage Probit equation). I use a linear specification of the duration equation to avoid computational complexities associated with estimating selection models in a hazard function framework (Kiefer, 1988). Notice that the regressions in Table A3 use information on all semiconductor firm pairings, not just those involved in litigation or cross-licensing.[31], [32]

In column (1) I present the estimates of a Heckman selection model that consider both the likelihood of being an interacting pair (first stage) and the probability of signing a broad cross-license (second stage). The first stage results show that the likelihood of interaction increases with average capital intensity and technological closeness. The second stage results show that the main finding of Table 2 is robust to this alternative model: Broad cross-license agreements are signed by firms pairing with high capital intensity.[33]

In column (2) I present the estimates of a Heckman selection model that consider both the likelihood of signing a broad cross-license deal (first stage) and the length of the negotiation (second stage). As in the previous model, the first stage results show that in the extended sample broad cross-licenses are signed by firms that are similar in the technology space. The second stage results show that negotiation delay decreases with capital intensity and increases with technological closeness. The signs of these coefficients are consistent with those in Table 3 but, in this extended model, only the coefficient on average capital intensity is statistically significant.[34]

I also run a similar Heckman two-stage model using the variables that were insignificant in the structural equation estimations (Tables 2 and 3) to identify the selection equation. The results are very similar to those presented in Table A3.

---

[31] For non-interacting parings, the variables are constructed using 1995 data, or the available data closest to 1995.

[32] In this larger sample the correlation between average capital intensity and average R&D intensity is high (equal to 0.56). This high correlation is consistent with the data in Hall and Ziedonis (2007). In the sample of interacting firm pairings, the correlation between capital intensity and R&D intensity is much smaller (it is equal to 0.05). The correlation is even smaller once I focus on the 38 pairings entering a broad cross-license agreement (it is equal to 0.01).

[33] The first stage regression is a linear probability model and not a probit model. The magnitude and the statistical significance of the coefficient on average capital intensity is very similar to the one obtained estimating the model in column (4) of Table 2 with a linear probability model (it is equal to 0.150 and significant at the 0.1 level).

[34] Also in this case, the coefficients are similar to those obtained estimating the model in column (2) of Table 3 with OLS. The coefficient on average capital intensity is equal to -11 and significant at the 0.05 level whereas the coefficient on Technological closeness is equal to 13.42 and not statistically significant.

6 Conclusion

This paper investigates firm incentives to engage in broad cross-licensing and the duration of broad cross-license negotiations. I develop a model of bargaining with learning in which firms litigate over their patent disputes if they do not agree on a cross-license. The model predicts that the incentive to litigate decreases with firms' capital intensity. More precisely, I show that whereas firms with low capital intensity prefer not to sign a cross-license agreement, firms with high sunk costs are better off cross-licensing their intellectual property. In addition, the model predicts that broad cross-license negotiations will be shorter for firm pairings with high capital intensity and for those facing a low frequency of infringements. Using a novel dataset of broad cross-license agreements in the semiconductor industry, I provide empirical support for the predictions of the model.

The paper can be extended in various directions. First, the model does not consider R&D investment and the possible impact that broad cross-license agreements may have on innovation rates. To study this issue would require endogenising disputes and investigating the relationship between R&D investments and the likelihood of infringement. Second, the model considers only two firms. Extending the framework to more than two players may help understanding what externalities cross-licensing creates and what network of agreements one should expect to observe.

Empirically it may be valuable to consider the possible effects of property right fragmentation on firm incentives to sign a broad cross-license. To construct a fragmentation index similar to the one in Ziedonis (2004) would be central to exploring this effect. Finally, this paper documents a relationship between capital intensity and broad cross-license negotiations in the semiconductor industry where the hold-up problem is argued to be particularly severe. I leave for future research the study of broad cross-license agreements in different industries such as pharmaceuticals or chemicals.

References

[1] Arora A., M. Ceccagnoli and W. Cohen, 2008, *R&D and the Patent Premium*, International Journal of Industrial Organization **26**, 1153-1179.

[2] Bebchuck L., 1984, *Litigation and Settlement under Imperfect Information*, RAND Journal of Economics **15**, 404-415.

[3] Binmore K., Rubinstein A. and Wolinsky A., *The Nash Bargaining Solution in Economic Modelling*, RAND Journal of Economics **17**, 176-188.

[4] Bloom, N., Schankerman, M. and Van Reenen, J., 2010, *Identifying technological spillovers and product market rivalry*, NBER working paper 13060.

[5] Bureau of Economic Analysis, 2010, Value Added by Industry, Gross Output by Industry, Intermediate Inputs by Industry, the Components of Value Added by Industry, and Employment by Industry http://www.bea.gov/industry/iedguide.htm

[6] Chatterjee K. and W. Samuelson,1983, *Bargaining under Incomplete Information*, Operations Research, 835-851.

[7] Choi J., 2003, *Patent Pools and Cross-Licensing in the Shadow of Patent Litigation*, mimeo.

[8] Crampton P., 1984, *Bargaining with Incomplete Information: An Infinite-Horizon Model with Two-Sided Uncertainty*, Review of Economic Studies **51**, 579–593.

[9] Crampton P., 1992, *Strategic Delay in Bargaining with Two-Sided Uncertainty*, Review of Economic Studies **59**, 205–225.

[10] Fernandez R. and Glazer J., 1991, *Striking for a Bargain Between Two Completely Informed Agents*, American Economic Review **81**, 240-252.

[11] Fershtman C. and Kamien M., 1992, Cross Licensing of Complementary Technologies, *International Journal of Industrial Organization* **10**, 329-348.

[12] Galasso A. and Simcoe T., 2011, *CEO Overconfidence and Innovation*, *Management Science* **57**, 1469-1484.

[13] Galasso A. and Schankerman M., 2010, *Patent Thickets, Courts and the Market for Innovation*, RAND Journal of Economics **41**, 472-503.

[14] Gans J. and Inderst R., 2007, Introduction to Special Issue on 'Negotiations and Cooperative Arrangements in Industrial Organization, *International Journal of Industrial Organization* **25**, 879-883.

[15] Gans J. and Stern S.,2000, Incumbency and R&D Incentives: Licensing the Gale of Creative Destruction, *Journal of Economics and Management Strategy* **9**, 485-511.

[16] Gans J. and Stern S., 2003, The Product Market and the Market for 'Ideas': Commercialization Strategies for Technology Entrepreneurs, *Research Policy* **32**, 333-350.

[17] Grindley, P. and Teece D., 1997, *Managing Intellectual Capital: Licensing and Cross Licensing in Semiconductor and Electronics*, California Management Review **39**, 8-41.

[18] Hall, B. and Ziedonis, R., 2001, *The Patent Paradox Revisited: an empirical study of patenting in the U.S. Semiconductor Industry* 1979-1985, Rand Journal of Economics **32**, 101-128.

[19] Hall, B. and Ziedonis, R., 2007, *An Empirical Analysis of Patent Litigation in the Semiconductor Industry,* working paper.

[20] Jaffe, A.,1986, *Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value*, American Economic Review **76**, 984-1001.

[21] Jaffe, A.,1988, *Demand and Supply Influences in R&D Intensity and Productivity Growth*, Review of Economics and Statistics **70**,431-437.

[22] Kiefer, N., 1988, *Economic Duration Data and Hazard Functions*, Journal of Economic Literature **26**, 646-679.

[23] Lanjouw J. and Schankerman M., 2001, *Characteristics of Patent Litigation: A Window on Competition*, RAND Journal of Economics **32**,129-151.

[24] Lanjouw J. and Schankerman M., 2004, *Protecting Intellectual Property Rights: Are Small Firms Handicapped?*, Journal of Law and Economics **47**, 45-74.

[25] Lemley, M. and Tangri, R., 2003, *Ending Patent Law's Willfulness Game*, Berkeley Technology Law Journal **18**,1085-1125.

[26] Lerner, J., 1994, *The Importance of Patent Scope: An Empirical Analysis*, RAND Journal of Economics **25**, 319-333.

[27] Lerner, J.,1995, *Patenting in the Shadow of Competitors*, Journal of Law and Economics **38**, 463-495.

[28] Lerner, J. and Tirole J., 2004, *Efficient Patent Pools* , American Economic Review **94**, 691-711.

[29] Macher, J.T. and D.C. Mowery, 2003, *Managing' Learning by Doing: An Empirical Study in Semiconductor Manufacturing*, Journal of Product Innovation Management **20**, 391-410.

[30] Marjit, S., Mukerjee A. and Shi L.,2001, *Cooperation in R&D: the case of patent infringement agreements*, Journal of Economic Behavior & Organization **45**, 383-401.

[31] Monteverde K., 1995, *Technical Dialog As an Incentive for Vertical Integration in the Semiconductor Industry*, Management Science **41**, 1624-1638.

[32] Semiconductor Industry Association,2006, Association Annual Report, http://www.sia-online.org/cs/annual_report

[33] Shaked, A. and Sutton, J., 1984, *Involuntary unemployment as a perfect equilibrium in a bargaining model*, Econometrica **52**, 1351-1364.

[34] Shapiro, C., 2001, *Navigating the Patent Thicket: Cross Licenses, Patent Pools and Standard Setting* in Innovation Policy and the Economy, Volume I, Adam Jaffe, Joshua Lerner, and Scott Stern, eds., MIT Press, 2001.

[35] Shapiro, C., 2003, *Technology Cross-Licensing Practices: FTC vs Intel*, in The Antitrust Revolution, J. Kwoka and L. White eds, Oxford University Press.

[36] Shapiro C., 2007, *Patent Reform: Aligning Reward and Contribution*, in A. Jaffe, J. Lerner, and S. Stern, eds.,Innovation Policy and the Economy, National Bureau of Economic Research **8**,111-156.

[37] Siebert R. and von Graevenitz G., 2010a, *Jostling for Advantage or Not: Choosing Between Patent Portfolio Races and Ex Ante Licensing*, Journal of Economic Behavior and Organization **73**, 225-245.

[38] Siebert R. and von Graevenitz G., 2010b, *Does Licensing Resolve Hold Up in the Patent Thicket?*, working paper.

[39] Somaya D., 2003, *Strategic Determinants of Decisions not to Settle Patent Litigation*, Strategic Management Journal **24**, 17-38.

[40] Sutton J.,1986, *Non-Cooperative Bargaining Theory: An Introduction*, Review of Economic Studies **53**,709-724.

[41] Theron, D.C., J. Rayter, I. Paprotny, G.T. Mackulak, 1999, *A Quantitative Approach for Minimizing Fab Cost of Ownership through Systems Integration*, in the Proceedings of the 10th Annual IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop, Boston, MA.

[42] Turley J., 2005, *The Essential Guide to Semiconductors*, Prentice Hall Professional Technical Reference.

[43] U.S. International Trade Commission, 2006, Semiconductor Manufacturing Equipment, USITC Publication 3868

[44] Yildiz M.,2003, *Bargaining Without a Common Prior - An Immediate Agreement Theorem*, Econometrica **71**, 793-811

[45] Yildiz M.,2004, *Waiting to Persuade*, Quarterly Journal of Economics **119**, 223-248

[46] Ziedonis R., 2003, *Patent Litigation in the U.S. Semiconductor Industry*, in Wesley M Cohen and Stephen A Merrill, editors, Patents in the Knowledge-based Economy,180-215

[47] Ziedonis R., 2004, *Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms*, Management Science **50**, 804-820

Appendix

Equilibrium Characterization

I begin by looking at the impact of capital intensity on the trade-off between litigating and cross-licensing that firms face to protect their intellectual property. If a cross-license agreement is reached, firms share duopolistic profits and the total per-period surplus is $2\pi^D$. Conversely, if litigation occurs then total per-period surplus depends on the court decision and in expectation is given by:

$$\overline{u} + \underline{u}(k) = \frac{1}{2}(\pi^M - kf) + \frac{1}{2}2\pi^D - 2L.$$

It is easy to see that $2\pi^D > \overline{u} + \underline{u}(k)$ only if $kf \geq \pi^M - 2\pi^D - 4L \equiv \underline{k}f$, therefore a cross-license agreement generates a larger surplus only for large values of capital intensity. The next proposition shows that if there is no initial disagreement either the firms enter a broad cross-license immediately or the contract is never signed.

Proposition 1 *If $\overline{m}_1 = \overline{m}_2$ and $k \leq \underline{k}$ the firms never sign a cross-license agreement. If $\overline{m}_1 = \overline{m}_2$ and $k > \underline{k}$ the firms sign a cross-license agreement at $t = 0$.*

P roof. If $k \leq \underline{k}$ the outside options (litigation payoffs) exceed the cross-license surplus and agreement cannot be reached. If $k > \underline{k}$ the duopoly surplus exceed the litigation payoff and agreement is possible. When $\overline{m}_1 = \overline{m}_2$ then $p_0^1(m) = 1 - p_0^2(m)$ and the firms agree that in the future firm 1 will make offers with probability $p_0^1(m)$. Without loss of generality, we assume that firm 1 is making the offer at period 0 and denote $p_0^1(1) = p$. Following Sutton (1986) and Fernadez and Glazer (1991), we indicate with Z_2 the supremum of the continuation payoff that firm 2 can obtain in the game starting after a rejection of firm 1 offer. Similarly, let Z_1 be the supremum of the continuation payoff that firm 1 can obtain in the game if it rejects firm 2 offer. Then:

$$\begin{aligned} Z_2 &= \underline{u}(k) + \delta p Z_2 + \delta(1-p)\left(\frac{2\pi^D}{1-\delta} - Z_1\right) \\ Z_1 &= \underline{u}(k) + \delta(1-p) Z_1 + \delta p\left(\frac{2\pi^D}{1-\delta} - Z_2\right). \end{aligned}$$

These equations imply:

$$Z_2 = \frac{\underline{u}(k)}{1-\delta} + \frac{\delta}{1-\delta}(1-p)\left(\frac{2\pi^D - 2\underline{u}(k)}{1-\delta}\right).$$

Following Sutton (1986), it is possible to show that supremum and infimum coincide. This implies that there is a unique subgame perfect Nash equilibrium with immediate agreement. The firm making the first offer obtains a payoff of $2\pi^D/(1-\delta) - Z_2$ and the firm receiving the offer obtains Z_2. ■

Intuitively, if $k \leq \underline{k}$ the costs associated with litigation are not very large and firms find more profitable to protect their intellectual property than to engage in broad cross-licensing. When $k > \underline{k}$, firms' capital intensity renders litigation less attractive than broad cross licensing. Because in the absence of asymmetric beliefs firms'expectations about the pattern of future infringements coincide, agreement is reached immediately. The bargaining game with common priors is therefore characterized by a discontinuity in agreement time: there is either immediate agreement (no delay) or perpetual disagreement (infinite delay). The presence of initial disagreement generates an intermediate case in which (finite) delay is observed. This result is illustrated in the following proposition where I characterize the unique Markov perfect equilibrium in which firms' strategies only depend on how many times firm 1 patents have been infringed.

Proposition 2 *If $\overline{m}_1 > \overline{m}_2$ and $k \leq \underline{k}$ the firms never sign a cross-license agreement. If $\overline{m}_1 > \overline{m}_2$ and $k > \underline{k}$ firms enter a broad cross-license agreement at a uniquely predetermined settlement date $t^* \geq 0$.*

P roof. As in the case of no optimism, there will be agreement only if the duopolistic surplus exceeds the surplus from litigation: i.e $2\pi^D \geq \overline{u} + \underline{u}(k)$. For low values of k this inequality is not satisfied and there is equality at $\underline{k}f = \pi^M - 2\pi^D - 4L$. Define now $V_t^i(m)$ as the continuation value of i at (m,t) and S_t as total surplus $S_t = V_t^1(m) + V_t^2(m)$. From Yildiz (2004) we know that S_t is deterministic and does not depend on m. Define the agreement regime the case in which

$$\frac{2\pi^D}{1-\delta} \geq \overline{u} + \underline{u}(k) + \delta S_{t+1}. \tag{5}$$

In this case the player chosen by nature extracts the rent $2\pi^D/(1-\delta) - \overline{u} - \underline{u}(k) - \delta S_{t+1}$. We define the no agreement regime the case in which (5) is not satisfied. In this case the rent extracted is zero. We can therefore define the rent extracted in period t as:

$$R_t = \max\left\{\frac{2\pi^D}{1-\delta} - \overline{u} - \underline{u}(k) - \delta S_{t+1}, 0\right\}.$$

Moreover

$$V_t^i = p_t^i(m)\left[R_t + \overline{u} - \underline{u}(k)\right] + \underline{u}(k) + \delta E(V_{t+1}^i) = \sum_{s=t}^{\infty} \delta^{s-t}\left[p_t^i(m)\left(R_s + \overline{u} - \underline{u}(k)\right) + \underline{u}(k)\right]$$

where the second equality follows because the current continuation value is the infinite sum over expected future rents. This can be re-written as:

$$V_t^i(m) = p_t^i(m)\Lambda_t + \frac{\underline{u}(k)}{1-\delta}$$

and it implies that

$$S_t = (1+y_t)\Lambda_t + \frac{2\underline{u}(k)}{1-\delta}$$

where $\Lambda_t = \sum_{s=t}^{\infty} \delta^{s-t}\left(R_t + \overline{u} - \underline{u}(k)\right)$. In the agreement regime:

$$R_t = \frac{2\pi^D}{1-\delta} - \overline{u} - \underline{u}(k) - \delta S_{t+1} = \frac{2\pi^D}{1-\delta} - \overline{u} - \underline{u}(k) - \delta\left[(1+y_{t+1})\Lambda_{t+1} + \frac{2\underline{u}(k)}{1-\delta}\right].$$

In addition the definition of Λ_t implies that $\Lambda_t = R_t + \overline{u} - \underline{u}(k) + \delta\Lambda_{t+1}$ and this condition can be used to obtain the following difference equation:

$$\Lambda_t = \frac{2\pi^D - 2\underline{u}(k)}{1-\delta} - \delta y_{t+1}\Lambda_{t+1}.$$

Notice now that in the agreement regime:

$$\frac{2\pi^D}{1-\delta} \geq \overline{u} + \underline{u}(k) + \delta S_t = \overline{u} + \underline{u}(k) + \delta(1+y_t)\Lambda_t + \frac{2\delta\underline{u}(k)}{1-\delta}. \tag{6}$$

that can be rewritten as:

$$\Lambda_t \leq \frac{2\pi^D - 2\underline{u}(k) - (\overline{u} - \underline{u}(k))(1-\delta)}{(1-\delta)\delta(1+y_t)} \equiv D_t. \tag{7}$$

Let us define

$$PA \equiv \{t \in T | \Lambda_s \leq D_s \ \forall s > t\}$$

and take $t^* = \min PA$. By definition there is agreement at each $t \geq t^*$. Moreover, take the largest $t < t^*$ for which there is no agreement at $t+1$. Because δS_{t+1} is large we have that $\Lambda_t = \overline{u} - \underline{u}(k) + \delta\Lambda_{t+1}$ and

$$S_t = (1+y_t)\left(\overline{u} - \underline{u}(k) + \delta\Lambda_{t+1}\right) + \frac{2\underline{u}(k)}{1-\delta} > \delta\Lambda_{t+1}(1+y_{t+1}) + \delta\frac{2\underline{u}(k)}{1-\delta} = \delta S_{t+1}$$

so there is disagreement at t. This result implies that there cannot be agreement at any $t < t^*$.
■

In equilibrium, the timing of the agreements is determined by a trade-off between the benefit and the cost from waiting one period. As in Yildiz (2004), the bargaining game is based on a tension between learning and the discount factor. On the one hand, asymmetric beliefs induce both parties to delay the agreement because waiting reduces opponent's optimism and therefore decreases opponent's requested surplus share. On the other hand, the discount factor reduces the value of future payoffs and induces parties to agree immediately.

Proof of Implication 1

It follows immediately from the fact that a broad cross-licence agreement is profitable only if the duopolistic surplus exceeds the surplus from litigation: i.e $2\pi^D \geq \overline{u} + \underline{u}(k)$. This inequality is satisfied only if $kf \geq \pi^M - 2\pi^D - 4L \equiv \underline{kf}$.

Proof of Implication 2

In our framework, that extends Yildiz (2004), it is not possible to find an analytical solution for t^*. Nonetheless, as in Yildiz (2004) we can characterize an upper and lower bound for t^* and show that these bounds decrease in k. Specifically, from Yildiz (2004) lemma 6, lemma 7 and lemma 8, we know that the condition that specify the upper bound in agreement delay, t_u, is:

$$\overline{B}_t \equiv \frac{(2\pi^D - 2\underline{u}(k))}{(1-\delta)(1+\delta y_{t+1})} \leq \frac{2\pi^D - 2\underline{u}(k) - (\overline{u} - \underline{u}(k))(1-\delta)}{(1-\delta)\delta(1+y_t)} = D_t.$$

that can be rewritten as

$$y_t - y_{t+1} \leq \frac{1-\delta}{\delta} - \frac{(\overline{u} - \underline{u}(k))(1-\delta)(1+\delta y_{t+1})}{(2\pi^D - 2\underline{u}(k))\delta}. \tag{8}$$

Let us rewrite (8) as:

$$y_t - y_{t+1} \leq \frac{1-\delta}{\delta} - A(k)\frac{(1-\delta)(1+\delta y_{t+1})}{\delta}$$

where the function

$$A(k) = \frac{(\overline{u} - \underline{u}(k))}{2\pi^D - 2\underline{u}(k)} = \frac{\frac{1}{2}(\pi^M + kf)}{2L + kf + \pi^D}$$

is what differentiate our model from the one of Yildiz (2004) in which the outside options $\overline{u}$ and $\underline{u}(k)$ are not present (if $\overline{u} = \underline{u}(k) = 0$ formula (8) is equivalent to the condition in theorem 1 of Yildiz(2004)). Notice that $A(k)$ is decreasing in k and that $A(\underline{k}) = 1$.[35] Since $y_t - y_{t+1}$ is decreasing in t and approaches zero as $t \to \infty$ there exists some real number t_u such that $\overline{B}_t \leq D_t$ if and only if $t \geq t_u$. Because $A(k)$ decreases in k, t_u is decreases in k. Similarly, the condition that specifies the lower bound, t_l, is:

$$\underline{B}_t \equiv \frac{(2\pi^D - 2\underline{u}(k))}{(1-\delta)} \frac{(1 - \delta(y_{t+1} - y_{t+2}))}{1+\delta y_{t+2}} \geq \frac{2\pi^D - 2\underline{u}(k) - (\overline{u} - \underline{u}(k))(1-\delta)}{(1-\delta)\delta(1+y_t)} = D_t$$

that can be rewritten as

$$\delta(1+y_t)\frac{(1-\delta(y_{t+1} - y_{t+2}))}{1+\delta y_{t+2}} \geq 1 - A(k)(1-\delta). \tag{9}$$

The right hand side of (9) increases in k because $A(k)$ decreases in k. Yildiz (2004) shows that the left hand side of (9) decreases in t so there exists some real number t_l such that $\underline{B}_t \geq D_t$ if and only if $t \leq t_l$. Because $A(k)$ decreases in k, t_l decreases in k.

Proof of Implication 3

When condition (8) is satisfied, the right hand side of (8) decreases in δ. This implies that an increase in δ increases t_u. Rewrite (9) as

$$H(t,\delta) = \delta(1 - A(k) + y_t)\frac{(1-\delta(y_{t+1} - y_{t+2}))}{1+\delta y_{t+2}} - 1 + A(k) \geq 0$$

[35] The derivative of $A(k)$ is negative as long as $\pi^M - \pi^D > 2L$ that is satisfied whenever $\underline{k}f > 0$.

with $H(t_l, \delta) = 0$. From Yildiz (2004) we know that $\partial H(t,\delta)/\partial t < 0$. In addition $\partial H(t,\delta)/\partial \delta > 0$.[36] Total differentiation implies that $dt_l/d\delta > 0$.

Proof of Implication 4

The first part of the implication follows immediately from the fact that an agreement is signed only if the duopolistic surplus exceeds the surplus from litigation, i.e. $2\pi^D$ is large enough.

To show the second part of implication 4, I exploit a simplified version of the model and show that bargaining delay does not depend on complementarities. Take a two period version of the bargaining game with no discounting. Firm 1 is the patentee in the first period and the identity of the second period infringer is stochastic. Each firm has a prior belief about the probability of being the second period patentee: firm 1 believes that it will be the patentee with probability p_1 and firm 2 believes that it will be the patentee with probability p_2. This belief structure is assumed to be common knowledge.

I assume that because of portfolio complementarities $2\pi^D \geq \pi^M$. Because of these complementarities, if the patent is found infringed the patent holder will prefer to license the patent and will not require the infringer to shut its production down. I assume that the licensing fee is set by equally splitting the surplus from complementarities so that in the case of licensing the payoff of the patentee is $\pi^M + (2\pi^D - \pi^M)/2$ and the one of the infringer is $(2\pi^D - \pi^M)/2$. The expected payoffs of the patentee and the infringer in the case of litigation are

$$\overline{u} = \frac{1}{2}\left(\pi^M + \frac{1}{2}(2\pi^D - \pi^M)\right) + \frac{1}{2}\pi^D - L$$
$$\underline{u} = \frac{1}{2}\left(\frac{1}{2}(2\pi^D - \pi^M)\right) + \frac{1}{2}\pi^D - L.$$

The main difference between these payoffs and those in formulas (1) and (2) is that when the infringer is found liable there is no stop in production and the cost kf is not sustained. This is the reason why the payoff of the infringer does not depend on k when complementarities are present.

In the two period game, if there is no agreement in the first period, firm 2 second period payoff is going to be $2\pi^D - \underline{u}$ if it makes the cross-license offer or $\underline{u}$ if firm 1 is the patentee. Therefore, the first period offer of firm 1 that will make firm 2 indifferent between accepting or rejecting a cross-license is: $\underline{u} + p_2(2\pi^D - \underline{u}) + (1-p_2)\underline{u} = 2\underline{u} + p_2(2\pi^D - \underline{u})$. This offer is going to be profitable for firm 1 if and only if:

$$4\pi^D - 2\underline{u} - p_2(2\pi^D - \underline{u}) \geq \overline{u} + \underline{u} + p_1(2\pi^D - \underline{u}). \tag{10}$$

The left hand side of (9) captures firm 1 payoffs from a cross license deal: the duopoly profits for two periods $(2\pi^D + 2\pi^D)$ minus the payoff that makes firm 2 indifferent between accepting or rejecting the deal $(2\underline{u} + p_2(2\pi^D - \underline{u}))$. The right hand side captures the payoff of firm 1 in the case of litigation: the expected payoff of litigation as patentee in the first period $(\overline{u})$ plus the expected second period payoff $(p_1(2\pi^D - \underline{u}) + (1-p_1)\underline{u} = \underline{u} + p_1(2\pi^D - \underline{u}))$. Condition (9) can be rewritten as

$$\frac{2\pi^D - \overline{u} - \underline{u}}{2\pi^D - 2\underline{u}} \geq y \tag{11}$$

[36] The sign of the derivative is equal to the sign of $1 - \delta(y_{t+1} - y_{t+2}) - \delta\,(1 + \delta y_{t+2})\,(y_{t+1} - y_{t+2})$ that is positive because in an agreement regime:

$$\frac{1 - \delta(y_{t+1} - y_{t+2})}{1 + \delta y_{t+2}} \geq 1 - \delta \geq \delta(y_{t+1} - y_{t+2}).$$

where $y = p_1 + p_2 - 1$. Because $2\pi^D - \overline{u} - \underline{u} = 2L$ and $2\pi^D - 2\underline{u} = 2L + 0.5\pi^M$ the condition for delay in agreement (10) is not affected by π^D.

Table 1. Descriptive Statistics

	Mean	Median	Std. Dev.	Min	Max
All interacting firm pairings (115 obs.)					
Cross-License	0.33	0	0.25	0	1
Average Capital Intensity	92070.70	78841.24	62363.73	10432.33	350815.60
Technological Closeness	0.51	0.55	0.29	0	0.97
Complementarity	0.008	0.001	0.022	0.000	0.187
Average Size	9441.02	4473	12195.36	111.00	48907.00
Average R&D Intensity	53850.20	42632.22	42293.85	1915.54	224610.2
Design Firm	0.43	0	0.49	0	1
Size asymmetry	1.65	1.34	1.31	1	8.83
Cross Licensing firm pairings (38 obs.)					
Length	13.61	5	19.47	0	95
Average Capital Intensity	123017.20	97883.84	81156.63	14628.64	350815.60
Complementarity	0.015	0.002	0.035	0.000	0.187
Technological Closeness	0.57	0.55	0.25	0	0.97
Average Size	13392.25	8995	13848.29	200	48908
Average R&D Intensity	64212.94	46945.46	48269.16	19155.41	185815.9
Design Firm	0.42	0	0.50	0	1
Size asymmetry	1.34	1.23	0.28	1.01	2.11

NOTES: Unit of observation is firm pair. Variables are constructed at the time of first interaction. Cross-License =1 if pairing enters in a broad cross-license agreement. Length=duration of pre-agreement litigation in months. Capital Intensity = ratio of property plant and employees to number of employees. Technological closeness= correlation of patenting across technology areas. Complementarity: cites between firms divided by total cites made by firms. Size=number of employees. R&D Intensity= ratio R&D to employees. Design=1 if one firm in the pair is fabless. Size asymmetry =ratio of the log employees of larger firm to log employees of smaller. Figures in 2001 dollars. Absolute values (not logs) are reported in this table.

Table 2. Probit Regression - Dependent Variable = Cross-License

	(1)	(2)	(3)	(4)
Average Capital Intensity	0.571	0.541	0.551	0.474
	(0.220)***	(0.229)**	(0.230)**	(0.241)**
Technological Closeness		0.243	0.363	0.536
		(0.467)	(0.507)	(0.449)
Average Size			0.178	0.201
			(0.124)	(0.167)
Size Asymmetry			-0.636	-0.632
			(0.395)	(0.414)
Average R&D Intensity			-0.087	-0.175
			(0.178)	(0.199)
Average Patent Yield			-0.508	-0.604
			(0.235)**	(0.267)**
Texas Instruments			0.469	0.763
			(0.313)	(0.454)*
Design			0.580	0.588
			(0.368)	(0.368)
Period 93-97				0.439
				(0.523)
Period 98-01				0.721
				(0.474)
Year Dummies 02-05				YES
Observations	115	115	115	115

*NOTES: Robust standard errors are reported in parentheses. Statistical significance: *10%, **5%, ***1%. Constant term not reported. Unit of observation is firm pairing. Variables are constructed at the time of first interaction. Cross-License =1 if pairing enters in a broad cross license agreement. Capital Intensity = log of ratio of property plant and employees to number of employees. Technological closeness= correlation of patenting across technology areas. Size=log of employees. R&D Intensity= log of ratio R&D to employees. Patent Yield: log of ratio patent stock to R&D. Design=1 if one firm is fabless. Size asymmetry =ratio of the log employees of larger firm to log employees of smaller. Figures in 2001 dollars.*

Table 3. Exponential Hazard Regression- Dependent Variable: Length

	(1)	(2)	(3)	(4)	(5)
Average Capital Intensity	1.199	0.819	0.958	1.118	0.827
	(0.231)***	(0.319)**	(0.343)***	(0.418)***	(0.322)***
Capital Intensity Asymmetry					0.030
					(0.018)*
Technological Closeness	-1.442	-1.485	-1.769	-1.727	-1.459
	(0.707)**	(0.707)**	(0.904)**	(1.004)*	(0.721)**
Average Size		0.178	0.067	0.511	0.222
		(0.265)	(0.292)	(0.332)	(0.264)
Size Asymmetry		1.575	-0.073	-0.126	1.678
		(0.715)**	(0.238)	(0.246)	(0.705)**
Average R&D Intensity		-0.558	-0.353	-0.267	-0.506
		(0.178)***	(0.212)*	(0.217)	(0.201)**
Average Patent Yield		1.067	0.652	-0.244	1.077
		(0.487)**	(0.562)	(0.810)	(0.506)**
Texas Instruments		-1.612	-1.062	-1.729	-1.757
		(0.825)*	(0.681)	(0.955)	(0.819)**
Design		0.375	0.112	0.496	0.241
		(0.287)	(0.473)	(0.436)	(0.295)
Period 93-97		0.750	0.676	-0.126	0.502
		(0.667)	(0.783)	(1.058)	(0.676)
Period 98-01		0.211	-0.012	-0.695	0.041
		(0.502)	(0.726)	(0.956)	(0.498)
Year Dummies 02-05		YES***	YES***	YES***	YES***
Observations	38	38	52	66	38

*NOTES: Robust standard errors are reported in parentheses. Statistical significance: *10%, **5%, ***1%. Constant term not reported. Unit of observation is firm pairing. Variables are constructed at the time of first interaction. Length=duration of pre-agreement litigation in months. Capital Intensity = log of ratio of property plant and employees to number of employees. Technological closeness= correlation of patenting across technology areas. Size=log of employees. R&D Intensity= log of ratio R&D to employees. Patent Yield: log of ratio patent stock to R&D. Design=1 if one firm is fabless. Size asymmetry =ratio of the log employees of larger firm to log employees of smaller. Figures in 2001 dollars. In column (3) 2004-05 litigating pairings are treated as censored spells. In column (4) 2002-05 litigating pairings are treated as censored spells.*

Table 4. Portfolio Complementarities and Cross-Licensing

	(1)	(2)	(3)	(4)
Estimation Method	Probit	Probit	Expon. Hazard	Expon. Hazard
Dependent Variable	Cross-License	Cross-License	Length	Length
Average Capital Intensity	0.739	0.675	1.294	0.816
	(0.272)***	(0.279)**	(0.249)***	(0.345)**
Technological Closeness	-0.318	0.012	-1.659	-1.476
	(0.619)	(0.556)	(0.853)*	(0.841)*
Complementarity	16.954	28.293	3.665	-0.285
	(7.378)**	(11.978)**	(4.411)	(8.690)
Average Size		0.245		0.173
		(0.163)		(0.279)
Size Asymmetry		-0.352		1.557
		(0.239)		(0.894)*
Average R&D Intensity		-0.182		-0.559
		(0.230)		(0.199)***
Average Patent Yield		-0.782		1.069
		(0.298)***		(0.478)**
Observations	115	115	38	38

*NOTES: Robust standard errors are reported in parentheses. Statistical significance: *10%, **5%, ***1%. Columns (2) and (4) control for design firms, Texas Instruments and time dummies. Unit of observation is firm pairing. Variables are constructed at the time of first interaction. Cross-License =1 if pairing enters in a broad cross license agreement. Length= duration of pre-agreement litigation in months. Capital Intensity = log of ratio of property plant and employees to number of employees. Technological closeness= correlation of patenting across technology areas. Complementarity: cites between firms divided by total cites made by firms. Size=log of employees. R&D Intensity= log of ratio R&D to employees. Patent Yield: log of ratio patent stock to R&D. Design=1 if one firm is fabless. Size asymmetry =ratio of the log employees of larger firm to log employees of smaller. SIC=correlation of sales across four digits SIC codes. Figures in 2001 dollars.*

Table A1. Robustness - Lags and SIC distance

	(1)	(2)	(3)	(4)
Estimation Method	Probit	Expon. Hazard	Probit	Expon. Hazard
Dependent Variable	Cross-License	Length	Cross-License	Length
Average Capital Int. (5 year lag)	0.432	0.765		
	(0.245)*	(0.344)**		
Average Capital Intensity			0.480	0.813
			(0.243)**	(0.315)***
Technological Closeness	0.616	-1.308	0.470	-1.512
	(0.475)	(0.609)**	(0.467)	(0.780)**
SIC Distance			-2.731	-0.357
			(3.448)	(3.769)
Average Size	0.247	0.228	0.189	0.171
	(0.179)	(0.250)	(0.168)	(0.266)
Size Asymmetry	-0.544	1.920	-0.706	1.541
	(0.409)	(0.575)***	(0.394)*	(0.768)**
Average R&D Intensity	-0.165	-0.465	-0.172	-0.554
	(0.197)	(0.159)***	(0.200)	(0.190)***
Average Patent Yield	-0.582	1.050	-0.563	1.081
	(0.244)**	(0.429)**	(0.269)**	(0.463)**
Observations	115	38	115	38

*NOTES: Robust standard errors are reported in parentheses. Statistical significance: *10%, **5%, ***1%. All columns control for design firms, Texas Instruments and time dummies. Unit of observation is firm pairing. Variables are constructed at the time of first interaction. Cross-License =1 if pairing enters in a broad cross license agreement. Length= duration of pre-agreement litigation in months. Capital Intensity = log of ratio of property plant and employees to number of employees. Technological closeness= correlation of patenting across technology areas. Size=log of employees. R&D Intensity= log of ratio R&D to employees. Patent Yield: log of ratio patent stock to R&D. Design=1 if one firm is fabless. Size asymmetry =ratio of the log employees of larger firm to log employees of smaller. SIC=correlation of sales across four digits SIC codes. Figures in 2001 dollars.*

Table A2. Robustness - Logs and Ratios

	(1)	(2)	(3)	(4)
Estimation Method	Probit	Expon. Hazard	Probit	Expon. Hazard
Dependent Variable	Cross-License	Length	Cross-License	Length
log (Average Property Plant and Equipments)	0.471**	0.619**		
	(0.227)	(0.315)		
Technological Closeness	0.439	-1.651	0 .454	-1.811
	(0.492)	(0.642)***	(0.626)	(0.578)***
log (Average R&D expenditure)	-0.193	-0.259		
	(0.145)	(0.104)**		
log (Average Size)	-0.084	-0.090		
	(0.309)	(0.474)		
Average Capital Intensity (no log)			0.001	0.001
			(0.000)***	(0.000)***
Average R&D Intensity (no log)			0.001	-0.002
			(0.005)	(0.004)
Average Size (no log)			0.001	0.001
			(0.001)	(0.000)***
Observations	115	38	115	38

*NOTES: Robust standard errors are reported in parentheses. Statistical significance: *10%, **5%, ***1%. All columns control for design firms, Texas Instruments Paten Yield, size asymmetry and time dummies. Unit of observation is firm pairing. Variables are constructed at the time of first interaction. Cross-License =1 if pairing enters in a broad cross license agreement. Length= duration of pre-agreement litigation in months. Capital Intensity = log of ratio of property plant and employees to number of employees. Technological closeness= correlation of patenting across technology areas. Size=log of employees. R&D Intensity= log of ratio R&D to employees. Patent Yield: log of ratio patent stock to R&D. Design=1 if one firm is fabless. Size asymmetry =ratio of the log employees of larger firm to log employees of smaller. Figures in 2001 dollars.*

Table A3. Sample Selection

	(1)	(2)
Second Stage OLS		
Dependent Variable	Cross-License	Length
Average Capital Intensity	0.169	-12.222
	(0.092)*	(6.392)*
Technological Closeness	0.220	18.264
	(0.186)	(15.667)
Average Patent Yield	-0.174	-6.811
	(0.075)**	(5.946)
First Stage Probit		
Dependent Variable	Interaction	Cross-License
Average Capital Intensity	0.361	0.531
	(0.171)**	(0.196)***
Technological Closeness	2.096	1.648
	(0.323)***	(0.380)***
Average Patent Yield	-0.442	-0.529
	(0.143)***	(0.171)***
Observations	23653	23653

*NOTES: Standard errors are reported in parentheses. Statistical significance: *10%, **5%, ***1%. All columns and both stages control for Size, Size Asymmetry, Texas Instruments and time dummies. Unit of observation is firm pairing. Variables are constructed at the time of first interaction and in 1995 for non-interacting pairs. Cross-License =1 if pairing enters in a broad cross license agreement. Length= duration of pre-agreement litigation in months. Capital Intensity = log of ratio of property plant and employees to number of employees. Technological closeness= correlation of patenting across technology areas. Patent Yield: log of ratio patent stock to R&D. Interaction=1 if the pairing engages in either litigation or broad cross-license.*