Distribution Agreement

Signature:

______________________________ ______________

David Tan Date

Status Ambiguity and Conflict in the Market

By

David Tan
Doctor of Philosophy

Business

Peter W. Roberts, Ph.D.
Advisor

Kira R. Fabrizio, Ph.D.
Committee Member

Christopher I. Rider, Ph.D.
Committee Member

Anand Swaminathan, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

Date

Status Ambiguity and Conflict in the Market

By

David Tan
BSBA, Creighton University, 2004

Advisor: Peter W. Roberts, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy
in Business
2009

Abstract

Status Ambiguity and Conflict in the Market

By David Tan

My main theoretical contention is that status hierarchies provide a source of guidance to firms for resolving disputes. A status hierarchy implies a system of deference rules among firms. When disputes arise, deference rules can provide a basis for shared expectations and protocols of conduct about how technical ambiguities should be resolved. In many contexts, technical merit is difficult to assess. However, deference rules can operate as social conventions to which firms default, helping to align potentially incompatible expectations.

Within this general framework, my dissertation examines how competitors in the semiconductor industry manage uncertain and frequently overlapping patent rights. In practice, patent rights are highly imperfect legal instruments when it comes to demarcating each firm's contributions to innovation in the industry. Patent disputes arise because of the ambiguity this creates about how much of the collective market returns to innovation each firm is entitled to receive. Despite the prolific patenting and propensity for disputes, the industry has remarkably not ground to a halt from runaway litigation. Litigation events, while significant, are rare.

I suggest that this degree of order is, at least partly, attributable to status processes. Status can operate as a stabilizing force in the market, helping to generate orderly competition in the face of disputes. To examine whether this is the case in the semiconductor industry, I theorize that disputes are less easily resolved when the parties

involved face greater status ambiguity, i.e. are less clearly differentiated from one another in status. Under status ambiguity, deference rules lose the rule-like, universal quality that makes them persuasive in resolving disputes. This has two consequences. First, firms facing low status ambiguity are less likely to be involved in patent litigation than are firms facing high status ambiguity. Patent litigation events represent failures to resolve patent disputes out of court. Second, firms facing low status ambiguity increase their product line sizes more than do firms facing high status ambiguity. The threat of difficult-to-resolve patent disputes represents a cost that can deter firms from bringing products to market.

Status Ambiguity and Conflict in the Market

By

David Tan
BSBA, Creighton University, 2004

Advisor: Peter W. Roberts, Ph.D.

A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy
in Business
2009

Acknowledgements

During my doctoral education, I have been lucky to have had many kind-hearted, hard-working individuals as my mentors and role models. I have reached this milestone because of the lessons they have taught me.

First and foremost, I am grateful to my dissertation chair, Peter Roberts. Throughout my doctoral education, Peter's ongoing lesson to me was to pursue the best version of my ideas and have faith in the research process to take me in a worthwhile direction. To this end, Peter patiently endured countless revisions as I struggled to give structure to my ideas. Peter's doctoral seminar gave me my first exposure to research in economic sociology. Though, looking back, I had little real understanding at the time of what economic sociology was, I was motivated by Peter's encouragement and, above all, his own enthusiasm for the subject to learn more. Over the next several years, Peter remained unwaveringly supportive as I developed my dissertation. Above all, he never allowed me to give up on my ideas without a fight.

I am also indebted to Kira Fabrizio, Anand Swaminathan, and Chris Rider, who gave so generously of their time and wisdom as members of my dissertation committee. I consider myself extremely lucky to have committee members who come from such diverse backgrounds and also share such a consistent attitude towards research. Above all, they tried to impart to me their commitment to doing careful, meticulous work. The exacting demands they place on their own work will continue to be a source of motivation for me for years to come. Kira instilled in me a taste for parsimonious explanations, which has helped to temper my taste for imaginative ones. I was constantly humbled and motivated by Anand's expansive knowledge and devotion to training doctoral students. I am grateful to Chris for the frequent conversations and advice that were invaluable to my development from student to scholar.

During my years at Emory, many other faculty members contributed to my growth as a scholar. I learned much from my interactions with faculty, including Giacomo Negro, Maura Belliveau, and Rich Makadok. I am especially grateful to Bob Drazin and Rob Kazanjian, who provided me with invaluable learning opportunities from my first day at Emory and helped me to feel a part of the community. I am grateful to Monica Worline, who instilled in me an enthusiasm for the craft of research.

I am fortunate to have belonged to a supportive community of colleagues in the doctoral program. I want to thank Scott Hayward, Chad Navis, Dave Kryscynski, Adina Sterling, Laura (McClelland) Nelson, Jung Lee, and Chirag Kasbekar for their friendship and the lessons I have learned from them. Each faced unique challenges as they went through the program. They taught me much through their example and contributed to my growth as an individual. From my cohort, I am grateful to have had Chad as both a friend and role model. I benefited greatly from our mutual encouragement as we went through the program. I hope we can continue to encourage one another for years to come. I want to thank Scott for the frequent conversations which provided a forum for sharing ideas as well as a welcomed break from the solitude of dissertation work. Finally, I want to thank Adina for helping me to rediscover my wonderment with science and providing a constant reminder of what I love most about our profession.

Above all, I am indebted to my parents for encouraging me to pursue a PhD and supporting me to finish. Some of my earliest childhood memories were of my father when he was finishing his PhD. Despite the pressure of raising a family and hardships along the way, my father never lost his love of scholarship. Before I was old enough to understand science, I remember his stories about scientists and scientific discovery. In this and many other ways, my father fostered my inquisitiveness about the workings of the world. He instilled in me the sense that there are no mundane things in the world, only things for which we have yet to discover the interesting questions. My mother, in addition to providing love and support as I was growing up, was also my first teacher. While my father was finishing his dissertation, my mother was teaching me math and literature. Over the years, she had to put on many other hats as well. By her example, my mother taught me to take pride in my work, whatever it happens to be.

Finally, I owe special thanks to my sister. In addition to offering encouragement and good humor when I was in my grumpiest of moods, she helped me with the laborious process of collecting my dissertation data. I hope I get the chance to return the favor. Otherwise, there are not enough pies in the world.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Motivation

In many social contexts, actors are differentiated by status, i.e. differences that are socially-constructed through honorific deference as opposed to meritoriously earned. As widely established in empirical studies, actors of higher status enjoy certain material advantages over actors of lower status (e.g. Podolny, 1993; Podolny and Stuart, 1995; Podolny, Stuart, and Hannan, 1996; Benjamin and Podolny, 1999; Jensen, 2003; 2006). Yet the interactions between them are more often orderly than disorderly. For casual intuition, consider, for instance, the relative infrequency with which the day-to-day interactions between people of high and low status prompt violent confrontations despite gross resource inequality (Gould, 2003). In the context of interactions between firms, consider the relative infrequency with which business disputes escalate into costly litigation (Macaulay, 1963).

Conflict ultimately is not about the division of resources but a failure to correctly discriminate among actors that ought to be treated differently. While resource competition and inequality are prevalent in many social settings, conflict remains relatively rare. Income inequality and other material grievances thought to generate civil war rarely actually do (Fearon and Laitin, 2003). Racial job segregation and inequality do not lead to higher levels of racial conflict (Olzak, 1992). Even violence-prone 'feuding societies' rarely see disputes escalate into intergroup violence (Gould, 1999). Conflict, instead, often arises as a consequence of some deeper form of ambiguity. For instance, disputing individuals may have ambiguous social status in one another's eyes, making them unwilling to show the deference that each expects from the other (Gould, 2003). Disputing groups may be unable to credibly demonstrate their commitment to fight and must instead demonstrate it through actual fighting (Gould, 1999). Governments may have difficulty distinguishing insurgents from civilians, leading to indiscriminate violence that fuels civil war (Fearon and Laitin, 2003). Ambiguity in general and ambiguous identity in particular provides an important structural condition for resolvable disputes to escalate into conflict. I apply this intuition to understanding conflict in market settings.

A market-based model of status ambiguity and conflict

In this dissertation, I develop the concept of organizational status ambiguity and examine its influence on conflict in market settings. Organizations face status ambiguity when there is room for them to disagree about how they sit relative to one another in the status hierarchy. Position in the status hierarchy signals unobservable qualities that make an organization more or less attractive as a transaction partner (Podolny, 1993). Organizations care about status differences because they want to correctly discriminate in their dealings with one another. Status ambiguity generates friction because it leads organizations to treat one another differently from how they expect to be treated.

In most kinds of transactions considered in status research, organizations discriminate among higher- and lower-status organizations primarily through their choices of transaction partners (e.g. Podolny, 1994). If misjudgments of status generate friction between organizations, then deals are simply not consummated—the parties walk away unscathed. In the context of settling a dispute, however, the parties cannot simply walk away. Misjudgments of status can delay settlement, prolonging the interaction between two parties while at the same time adding friction to these interactions. Status ambiguity therefore provides a structural condition for resolvable disputes to escalate into more costly conflict.

To build a theory of how status ambiguity generates conflict, I draw on two ideas: the idea that status is a signal of quality (Podolny, 1993) and the idea that a dispute is an occasion for a business transaction (Coase, 1960; Calabresi and Melamed, 1971). In a market setting, an organization’s status reflects the extent to which its outputs are endorsed by other organizations as being high in quality (Podolny, 1993). The level of endorsement an organization receives from its peers both enables and constrains it to produce at a correspondingly high level of quality. In this way, status operates as a signal of quality. The precise meaning of quality depends on the nature of the transaction being considered. For instance, in the sale of consumer goods such as clothing, quality may refer to durability. In the sale of a professional service such as legal representation, quality may refer to diligence. In general, quality refers to a property of a good or service that is desirable to transaction partners and, hence, commands better terms of exchange.

In many settings, ranging from (but not limited to) consumer goods to professional services, quality is difficult to assess in advance of the transaction. Status becomes important because it operates as a signaling mechanism for organizations to differentiate themselves and to discriminate in their dealings with one another (Podolny, 2001; Spence, 1974). Higher status organizations demand and receive better terms of exchange. Lower status organizations neither demand nor receive attractive terms of exchange. That every organization knows its place makes for orderly interactions all around in the face of uncertainty.

This becomes important for settling business disputes. Settling a dispute is seemingly different from other business dealings in that it involves compensation for damages rather than exchange of goods or services. However, to the extent that damages can be priced and exchanged in the same manner as an economic good, disputes can be resolved through market mechanisms (e.g. contractual agreements). In the context of settling a dispute, quality refers to a desirable property (or absence of a negative property) that influences what each party would be willing to pay to or accept from one another in a settlement. For instance, in Coase's (1960) example, quality may refer to the degree of care (or negligence) that a rancher takes to keep his herd from trampling his neighbor's farm. As in the exchange of a good or service, providing higher quality requires more costly investment, something which the providing party hopes will be reflected in the settlement of a dispute. And as in the exchange of a good or service, quality is often difficult to assess, making it difficult for parties to correctly discriminate in their settlement offers. The problem of settling a dispute therefore becomes analogous to the problem of two parties trying to discover one another's preferred terms of exchange. Just as misjudgment of bargaining thresholds can prevent transactions from happening, so it can also prevent the settlement of disputes. In the former case, the parties simply walk away. In the latter case, failure to find solutions within market mechanisms forces the parties to seek solutions outside of market mechanisms, such as going to court. I characterize the escalation of disputes to domains outside of the market as instances of conflict.

Combining these two ideas, I argue that when organizations occupy ambiguous positions relative to one another in the status hierarchy, they are less able to appropriately discriminate in their dealings with one another and, therefore, more likely to become embroiled in conflict.

Dividing returns to innovation

I use this framework to explain when litigation occurs in the course of dividing returns to innovation. In many fields of technology, a firm's costly investments in innovation generate unintended knowledge flows that benefit its competitors and vice versa (Cohen & Levin, 1989). Though firms prolifically patent their innovations, these knowledge flows between competitors rarely result in infringement litigation (Lanjouw & Schankerman, 2004). Instead, firms often arrange licensing agreements to exchange access to one another's patents, giving each other the freedom to make products without fear of litigation. Since, firms do not all contribute equally to innovation, they naturally care about offsetting these perceived differences when setting licensing terms (Grindley & Teece, 1997). This can be difficult to do. Patents provide a poor indicator of a firm's true innovative activities. At the same time, both high and low contributors stand to gain from representing themselves as high contributors. Being unable to correctly discriminate between high and low contributors can cause licensing negotiations to fail and force firms to take their disputes to court.

In this context, a firm's unobserved 'quality' corresponds to its costly contributions to the development of a technology. In order to earn high enough licensing royalties to recoup these costs, firms must convince other firms that they in fact have a credible commitment to invest in advancing a field of technology. This echoes the general case in which producers must convince consumers that they have a credible commitment to invest in quality in order to charge high enough prices to recoup costs. A status hierarchy, by channeling rewards and recognition away from low-status firms and toward high-status firms, provides a basis for demonstrating such commitments. It provides a shared understanding of the roles that firms occupy in the innovative activity of the industry and also induces firms to behave according to roles conferred upon them.

In doing this, a status hierarchy helps competitors resolve an important dilemma in dividing returns to innovation. If they adhere to a strict legal interpretation of patent rights, then the legal threat of patents potentially allows some firms to demand licensing fees far in excess of their true contributions to innovation. The resulting litigation to assert and fight such demands could be detrimental to all firms in the market. However, in order to divide returns to innovation in a way that reflects each firm's true contribution, they must first all agree about what each firm's true contribution is. This is no easy task. Status hierarchies can provide a basis for this kind of mutual understanding and help firms avoid costly litigation.

Roadmap

In the next three chapters, I develop my proposed framework. Chapter 2 examines anecdotal evidence from patent litigation in the semiconductor industry to make the case that patent litigation is primarily a matter of ambiguity. Specifically, I argue that litigation centers on two related problems: difficulty in tracing and measuring knowledge flows between companies, and ambiguity about the extent to which a given company is genuinely committed to innovation.

The flow of knowledge between companies generates complex patterns of competitive interdependency which are difficult to manage through licensing solutions. Attempts to resolve disputes break down in the face of this complexity because, on the one hand, ‘reasonable’ solutions are usually not the same as what is implied by legal obligations, but on the other hand, companies are prone to disagree about precisely what is ‘reasonable’. In addition, because some firms may be more committed to and invest more in ongoing innovation, there is the potential for licensing partners to benefit asymmetrically from a licensing agreement. This possibility induces firms to be discriminating in setting licensing terms, but at the same, however, firms have difficulty distinguishing whether one another’s demands are meant to provide funding for ongoing innovation or an easy revenue stream that can be maintained without investment in innovation. Hence, litigation reflects failed attempts to discriminate in the face of ambiguity.

Chapter 3 examines anecdotal evidence from the early history of microprocessor technology to reveal the role that structural forces play in the context of innovation. The complex reality of invention contrasts sharply with the much simpler understandings which firms actually use to characterize one another’s roles in the innovative activity of the industry. The case of the microprocessor illustrates that while historical details behind innovation are highly contestable, they are also easily suppressed.

Historical details may be forgotten or recounted selectively by industry leaders. However, biased accounts may go unchallenged because they nevertheless provide sensible guidance to firms for the ongoing task of dividing returns to innovation. The case illustrates how a firm can become locked (even reluctantly) into ongoing commitment to the development of a technology. This suppresses the importance of historical details which might suggest alternative ways of dividing credit for innovation. In other words, structural forces induce firms to behave according to the positions they occupy in the industry. When such commitments are clear to all firms involved in dividing returns to innovation, historical details simply do not arise as points of contention. This builds on the argument from chapter 2 that, when firms do get into conflicts about how to divide returns to innovation, it is because there is ambiguity about their commitments to innovation.

Chapter 4 synthesizes the insights from chapters 2 and 3 into a general framework for understanding status ambiguity and conflict. A status hierarchy can help firms settle disputes by providing occupants of different positions with a mutually-recognized signal of quality differences between them. To be effective, however, there must be sufficiently strong differences in the opportunities that different positions provide to their occupants to produce high quality output. Without sufficiently strong contrast between positions, firms are less able to appropriately discriminate in their dealings with one another and, hence, less effective at settling disputes.

This reasoning generates two sets of testable predictions. The first set of predictions concerns the risk of conflict. Since contrast in opportunities and constraints is stronger between distant positions in the status hierarchy than between nearby positions, the degree of status ambiguity and hence risk of conflict decreases with the degree of status difference between firms. This effect is weakened by the presence of inconsistencies in status difference when endorsements of quality come from multiple subpopulations of peers. The second set of predictions concerns growth in firms' product lines. Since status ambiguity increases the risk of conflict with competitors, it also imposes a cost to product line expansion. Bringing products to market exposes a firm to infringement disputes. For firms facing high status ambiguity these disputes are less easily resolved and more costly. This deters firms from bringing products to market. Firms with high status difference from others in the industry face less ambiguity and increase their product line sizes more than do firms with low status difference. This effect is weakened by the presence of inconsistencies in status difference when lower-status firms receive endorsements of quality from higher-status firms.

Chapters 5 and 6 present empirical tests of these predictions. In chapter 5, I test the first set of predictions concerning the risk of conflict. I construct a case-control dataset that matches competitors sued by a patent owner for infringement to other competitors making the same product that were not sued. This yields a sample of dyadic observations in which some instances of infringement resulted in litigation while other instances did not. Using this dataset, I show that, conditional on infringement, two firms which are different in status are much less likely to litigate than firms which are similar in status. I also show that the effect of status difference is attenuated by the degree to which these status differences are consistent across different subpopulations in the industry that are responsible for conferring status.

In chapter 6, I test the second set of predictions concerning growth in firms' product lines. I construct a dataset consisting of all products introduced between 1977 and 2001 in the EEPROM (electrically-erasable programmable read-only memory) segment of the semiconductor industry. Using this dataset, I show that a firm's rate of product line expansion increases with its degree of status differentiation. I also show that the effect of status differentiation becomes weaker in the presence of inconsistencies due to endorsements from higher-status firms to lower-status firms.

Chapter 7 concludes by discussing how the framework presented here relates to other relevant bodies of research.

Chapter 2: Ambiguity in the context of patent litigation

Introduction

Patent litigation (and in fact business litigation in general) remains a poorly understood phenomena. As with any other social phenomenon, the way we characterize litigation has important implications for how we explain it and how we use it to understand the circumstances behind it. For instance, because of its association with patents, patent litigation is commonly viewed as a mechanism for capturing returns to innovation. This is, indeed, its formal function from a legal standpoint—it serves as recourse against infringement. Similarly, it is tempting to characterize other kinds of business litigation as being fundamentally about the legal concepts that are formally at issue. In this view, the filing of lawsuits is as perfunctory a part of doing business as the writing of contracts.

Sociological studies suggest, however, that litigation reflects something more. There is a stigma in business communities around invoking arm's-length legal obligations in the course of doing business (Macaulay, 1963). Business people recognize that complexities of doing business are difficult to capture fully in legal contracts (Macneil, 1978). They rely instead on unwritten mutual understandings to adjust terms of trade in response to unforeseen circumstances (Uzzi, 1997; 1999). Hence, there is a stigma when businesses depart from these understandings and invoke legal obligations when it is to their advantage. In this view, business disputes that must be resolved are commonplace. Litigation is not. The fact that a dispute escalates to litigation implies that something more fundamental has broken down, and there is more to the dispute than simply the material stakes or formal legal concept for which a lawsuit is ostensibly filed.

In this chapter, I examine anecdotal accounts of patent litigation in the semiconductor industry. I focus in particular on the changes in the semiconductor industry starting in the 1980s with respect to firms' stances toward intellectual property. From the 1980s through the 1990s, firms in the semiconductor industry took an increasingly strong stance on protecting intellectual property. During this period, the number of patent lawsuits between firms in the industry increased dramatically. Several important changes were happening in the industry at the time to which popular press accounts have attributed the apparent 'litigation explosion'. Previous accounts of this period emphasize the strengthening of patent rights following the creation of the Court of Appeals for the Federal Circuit in 1982 and increasing competition in a maturing industry. In the following sections, I use anecdotal information to try to disentangle the possible mechanisms linking these changes to litigation. The result of doing this suggests that while the strengthening of patent rights and increased competition may indeed have made patent rights more valuable during this period, this alone was not the 'cause' per se of litigation.

Based on the anecdotal evidence, I suggest instead that patent litigation often arises from two related factors. First, it arises from the difficulty in tracing and measuring knowledge flows between companies. These generate complex patterns of competitive interdependency which are difficult to manage through licensing solutions. Attempts to resolve disputes break down in the face of this complexity because, on the one hand, 'reasonable' solutions are usually not the same as what is implied by legal obligations, but on the other hand, companies are prone to disagree about precisely what is 'reasonable'.

Second, I suggest that litigation arises from ambiguity about the extent to which a given firm is genuinely committed to innovation and what licensing demands are appropriate in light of this. Because some firms may be more committed to and invest more in ongoing innovation, there is the potential for licensing partners to benefit asymmetrically from a licensing agreement. This possibility induces firms to be discriminating in setting licensing terms, using royalty levels, for instance, to offset perceived disparities in their ongoing innovative contributions. At the same, however, firms cannot distinguish based on the aggressiveness of one another's demands whether these are meant to provide funding for ongoing innovation or an easy revenue stream that can be maintained without investment in innovation. Hence, licensing breakdowns fundamentally reflect failed attempts to discriminate in the face of ambiguity.

In suggesting that patent litigation arises from these two factors, I suggest that patent litigation is generally not simply about the narrow legal matter of patent infringement. As I hope to demonstrate in this chapter, the fundamental issues behind patent litigation events are often broader in scope than the particular patents and products that are ostensibly the subject of infringement suits. In illustrating this, I hope to motivate the idea that conflict is rooted in structural rather than idiosyncratic causes and the idea that litigation provides a window on these structural causes. I conclude the chapter by discussing how this portrayal of patent litigation is consistent with insights from the broader literature on patenting.

Creation of the CAFC and heightened competition, 1980s through 1990s

In the semiconductor industry in the 1960s and 1970s, companies were primarily focused on pushing innovative new products out the door and expanding the market for microelectronics. As a part of this focus, technology was widely licensed for only nominal royalties (Tilton, 1971; Wiegner, 1987). Often, companies simply traded licenses without assessing royalties at all (Richards, 1990; Kehoe, 1991).

This ‘era of benign neglect’ came to an end in the 1980s, however, when companies began paying closer attention to the value of intellectual property (Electronic Buyers News, 1990). The business press described the situation as “a bit like aging rock stars who realize that the hits are becoming harder to come by” with companies like Intel and Texas Instruments starting to look beyond the sales of their own products to capture returns to innovation (Wiegner, 1987). Ben Anixter, vice president at AMD observed, “Intellectual property has become much more valuable than everybody thought it was going to be 20 years ago. You have to protect your intellectual property because you put so much money into it” (Richards, 1990). Richard Agnich, general counsel for Texas Instruments stated, “There was a change in the industry culture. We recognized that our intellectual property assets were very valuable, and that we were not getting a good enough return on those assets” (Kehoe, 1991).

A number of factors were thought to have contributed to the apparent shift in IP strategy. The most prominent was the creation of the Court of Appeals for the Federal Circuit (CAFC) in 1982, which was thought to have in effect strengthened patent rights. Prior to 1982, patent cases were appealed to various federal appeals courts. However, these courts varied widely in their tendencies to uphold patent claims (Electronics Times, 1986). Parties to a suit often fought over the venue (between courts with pro- and anti-patent reputations) in order to improve their chances of a favorable ruling (Thompson, 1988). This discouraged the filing of patent cases and contributed to overall weakness in patent protection. The Court of Appeals for the Federal Circuit centralized appeals for patent cases. The CAFC earned a reputation for being ‘pro-patent’ in the sense of being more uniform and in tending to decide in favor of patent owners in cases passed up to it (Hall and Ziedonis, 2001). This resulted in trial courts adopting similar stances (Henry and Turner, 2006). Carl Silverman, chief counsel of Intel, stated, “Depending on whom you talk to, this new court has either made life more uniform or is more pro-patent” (Wiegner, 1987). John McDonnell, assistant general counsel for intellectual property at AT&T, stated, “You now get a lot of consistency even at the trial court [level]” (Thompson, 1988).

In addition, the 1980s saw increased competition, especially from Japanese semiconductor companies. Within the span of a decade, Japanese competitors had overtaken many leading American semiconductor companies in market share, especially in memory segments. In 1988, NEC (Nippon Electric Corporation) led the industry in sales with $4.65 billion (Steinert-Threlkeld, 1989). In contrast, Texas Instruments ranked fifth, with $2.75 billion in sales. Intel’s Andy Grove observed, “This industry lost a lot of its youthful innocence in the last few years. We had gone through an exuberant age when we believed we were in the lead and would stay there. We never had to face wholesale extinction” (Wiegner, 1987). James Pooley, a Palo Alto attorney in high technology patent law, observed “Back in the late 1950s and early 1960s, there was plenty of room for everybody. These days, the market is much, much more competitive, therefore you have to protect your technology edge much more carefully. That may be the most of what you have to offer" (Electronic Times, 1986). Many popular press accounts and accounts of semiconductor managers attributed the increased importance of patent rights in the industry and the apparent 'litigation explosion' to the confluence of these two changes, the creation of the CAFC and increased competition.

However, after closer consideration, it becomes less clear whether the industry had actually become more litigious or whether the heightened competition and strengthened patent rights contributed to litigiousness. A strengthening of patent rights under the CAFC undoubtedly altered the division of returns to innovation between net contributors and net users of innovation, but it is not clear that this in and of itself increased the propensity for litigation. While necessary, it is likely not on its own sufficient.

Anecdotal evidence suggests that in many cases, firms on both sides of licensing negotiations adjusted their expectations to the post-CAFC era. The business press observed that while it might appear that the semiconductor industry was "on the verge of tearing itself apart from within", the changes need not "lead to endless judicial snarls if the industry can come to understand the change and learn to adapt to it" (Electronic Buyers News, 1990).

First, there did not appear to be a shift to a greater use of patents for exclusivity. Tom Dunlap, vice president and general counsel at Intel suggested, "The only thing [the change] really does is cut out the free lunch. A lot of people think the industry has decided there will be less licensing or no licensing. I don't think that'll happen. It's a trend to looking at the value of the license. Don't license it if I'm not going to get something of value" (Thompson, 1988). William Sick, executive vice president at Texas Instruments stated "We don't really care which [licensing or exclusivity]. What we want is a reasonable return on our research and development investment, one way or another" (Kehoe, 1986). After filing a 1986 lawsuit against several Japanese electronics companies over DRAM patents, TI notified other producers of DRAMs of its company's willingness to enter into licensing agreements (PR Newswire, 1988).

Second, while average royalty rates increased, perhaps the most important change was that companies were not simply raising demands across the board but were instead becoming more discriminating in their licensing agreements. Glen Madland, chairman of the Integrated Circuit Engineering Corporation, observed that while royalty rates typically fell within a narrow range of 2% to 5 % of sales in the early 1980s, this had broadened to a range of 2% to as high as 40% by the late 1980s (Harbert, 1990). An important source of variation in licensing terms was the relative strength of two companies' patent portfolios. For instance, in a company statement, Fujitsu claimed that by using its own patent portfolio as a bargaining chip, it was able to negotiate a cross-license with Texas Instruments at "substantially lower royalty payments than Fujitsu has paid in the past" (Whiting, 1992). Similarly, Texas Instruments' earned much less in royalties after settling its suit against Hitachi than it did against other Japanese electronics companies it had sued in the past due to the strength of Hitachi's patent portfolio (Wall Street Journal, 1987).

Importantly, while there were several prominent examples of litigation when royalty negotiations broke down, most of the adjustment to the post-CAFC era was likely to have been accomplished quietly without litigation. The industry press observed that even without going to court, semiconductor companies were demanding and receiving royalties several times what was typical a decade before (Resnick, 1991).

In addition to the anecdotal evidence, quantitative evidence suggests that the propensity for litigation did not increase during this period. An important thing to consider is that during this time, the total number of patents granted each year was increasing dramatically (Kortum and Lerner, 1999). The 1980s may have seen a ‘litigation explosion’ in the sense of an increase in absolute *number* of lawsuits filed each year. However, when this is considered in light of the increasing number of patents that could be litigated, it is not clear that the *propensity* for litigation had increased. In addition to an increase in number of patents, this period also saw increasing overlap between firms in product markets. If it is the case that firms were encountering more competitive overlap with one another in product markets, then this would naturally increase the occurrence of patent disputes to be litigated. In the most systematic study to date of patent litigation, Lanjouw and Schankerman (2001) find that while the number of lawsuits filed increased dramatically, this increase was largely driven by the dramatic increase in number of patents granted. Accounting for the number of patents available to be litigated, the propensity for litigation remains constant over time. Looking at litigation relative to competitive overlap yields a similar conclusion. The number of lawsuits in the industry increases monotonically over time. However, so does the amount of competitive overlap.

Another piece of evidence to consider is the kinds of companies that were on the sending and receiving ends of lawsuits. A perceived effect of the CAFC was to contribute to litigation in the form of firms with large patent portfolios preying on firms with small patent portfolios. Competitors with large patent portfolios are presumably better able to enter into cross-licensing agreements with one another. However, Lanjouw and Schankerman (2004) find that firms with large patent portfolios are not more prone to litigation than firms with small patent portfolios. In fact, the rate of litigation, as implied by the proportion of total patents litigated, actually decreases significantly with the size of a firm's patent portfolio. Similarly, the press observed:

> "Big companies aren't always the aggressors. Last month a federal court jury, finding that Advanced Micro Devices Inc. had copied the chip design belonging to a much smaller San Diego firm [Brooktree], ordered AMD to pay $26 million." (Richards, 1990)

Hence, it was not simply a case of large firms versus small firms. Litigation often occurred between firms with large patent portfolios, as in the example of Samsung versus Fujitsu.

The anecdotal and empirical evidence therefore suggest that we be careful about how we characterize litigation and how we characterize the effect of changes in the semiconductor industry. The creation of the CAFC and heightened competition may have led patent owners to assert patent rights more aggressively. However, this alone does not necessarily lead to a higher propensity for litigation. In principle, firms adapt to changes in the legal regime by adjusting contracting terms. Literature in the law and economics tradition suggests that in a world where firms are free to contract with one another, changes in legal entitlements—the right to and extent of compensation for infringement, for instance—will influence contracting terms but not fundamentally alter firm behaviors (Calabresi and Melamed, 1972; Cooter, 1982; Merges, 1994). The creation of the CAFC has been criticized for promoting the granting and litigating of weak patents (Jaffe and Lerner, 2004). The available anecdotal evidence (though scarce because negotiations are confidential) suggests, however, that firms take this into account when negotiating licensing terms by weighting patents according to expected value and validity (Grindley and Teece, 1997; Sherry and Teece, 2004). The creation of the CAFC did seem to have influenced firm behavior in terms of inducing a ramp up in the rate of firm patenting (Hall and Ziedonis, 2001; Ziedonis, 2004). This is consistent with the idea that the CAFC strengthened patent rights, thereby increasing the value of obtaining and enforcing patents. However, it would suggest something else also. The evidence suggests that, despite the increasingly dense thicket of patent rights, firms were generally able to adjust their mutual understandings of the environment accordingly and manage the thicket in an orderly manner.

In the preceding discussion, I have portrayed the creation of the CAFC as an example of an increase in the value of patent rights that did not lead to an increased rate of patent litigation. The example is meant to illustrate Coase's (1960) theorem that firms can use contractual solutions to price and exchange damages like any other kind of good and, in doing so, provide motivation for viewing litigation events as bargaining failures. On its own, the creation of the CAFC is problematic as an example because it may also have reduced uncertainty about litigation outcomes at the same time that it increased the value of patent rights. If the overall effect of the CAFC was really a composite of these two countervailing forces, then looking at the rate of litigation alone would not provide a window on the magnitude of either force individually (Galasso & Schankerman, 2008).

The general idea that litigation represents a bargaining failure is, however, well-established in the theoretical literature (Bebchuk, 1984; Reinganum & Wilde, 1986; Meurer, 1989). In the next sections, I present additional anecdotal evidence to motivate this idea.

Litigation as a bargaining failure

If it is the case that firms did not, on average, become more litigation-prone, despite changes in the intellectual property regime, what do we make of the litigation events that did occur? Correctly characterizing litigation is important for explaining why it happens. Litigation is ultimately a two-sided event. One side must make a demand that the other side refuses. This is more of a barrier to litigation than is generally recognized. As long as litigation is costly to both firms involved, there is potentially a positive bargaining zone to be shared between firms that settle.

Litigation is financially costly. Alfred Stein, chairman and chief executive of VLSI Technology stated, “The cost of legal fees, and the time that senior executives are spending in meetings over these lawsuits is running into millions of dollars” (Kehoe, 1991). In addition, litigation is costly in terms of time and energy diverted away from core operations. T.J. Rodgers of Cypress complained that “I have had to spend one day per week for the past year working on legal issues” (Resnick, 1991). Stephen Cordial, CFO at Sierra Semiconductors, complained of having to divert engineers’ time from developing new products to documenting patent claims (Resnick, 1991).

In addition to litigation being financially and operationally costly, it is a risky approach that can backfire for both parties. After settling a lawsuit with National

Semiconductor, Robert Swanson, president of Linear Technology, said of the attorney and court fees, “neither of us would like to make that public, as it would be too painful. I can assure you it was more than the amount of the settlement” (Wirbel, 1989). In a case between Motorola and Hitachi, the federal district court judge, Lucius D. Bunton, called the case a “travesty of justice… replete with bickering and petty insults”, declaring both companies guilty of infringement and enjoining both from selling the infringing products (Norman, 1990). Bunton stated in his ruling:

> “The parties would have saved time, money, feelings and relations had they curbed their emotions and sat down to settle their difference out of court. In short, this suit is not the sort of thing federal courts should spend time and energy upon.” (Yoder, 1990)

Hence, many companies discover that the first-best option is to compromise. Neil Steinberg, associate patent counsel for Samsung, commented, “Eventually the parties come together and say ‘This is crazy’” (Wade, 1997). Charles Donohoe, vice president and general patent counsel for Samsung, said of litigation with Fujitsu, “It’s in the mutual interest of both companies to resolve our differences in a friendly fashion” (Saul, 1996). W.J. Sanders III, chairman and chief executive of AMD said of the company’s settlement of litigation with Cypress, “Cross-license agreements that expand market opportunities and provide value-for-value exchange are beneficial to all parties. We are very pleased that the settlement agreement with Cypress will enhance market opportunities for both companies while serving the interests of customers” (Business Wire, 1992). Similarly, Cypress CEO T.J Rodgers stated, “I think both companies [Cypress and AMD] will be better off tending to business and keeping out of the courtroom” (Business Wire, 1992).

Given this, why do companies litigate in the first place? In other words, why do they not discover these mutually beneficial compromises before litigation?

Litigation often occurs when licensing negotiations breakdown from disagreement over licensing terms. Joseph Hinchey, chief financial officer of Analog Devices suggested that the industry has not yet reached consensus about the true value of patent rights (Kehoe, 1991). Richard Agnich, general counsel and senior vice president of Texas Instruments, said about the company's suit against Micron Technology, "TI has been in detailed discussions with Micron for more than six months, but Micron has been unwilling to accept a license under terms which will provide us a fair return on our investment. This has left us with no recourse except litigation" (PR Newswire, 1988). Mark O'Molesky, SGS-Thomson's vice president of business development, said of the company's lawsuit against Hyundai, "We spent over 18 months attempting to conclude a license [agreement] with Hyundai which would provide just compensation for Hyundai's use of our intellectual property. Despite our efforts, and contrary to other companies we have approached, Hyundai has repeatedly refused to accept a license on reasonable terms" (Dunn, 1990). In 1993, National Semiconductor sued Mitsubishi after two-year long cross-licensing negotiations broke down. Mary Ann McKay of National Semiconductor said of the company's lawsuit against Mitsubishi, "We've been very successful in negotiating cross-licensing agreements in the past with other Japanese companies, and we would still prefer to do that here as well" (Newsbytes News Network, 1993). When Samsung and Fujitsu traded lawsuits in 1997, spokespeople for both companies stated that cross-licensing negotiations broke down because both companies felt entitled to royalty payments from the other (Wade, 1997). Michael Moore, general counsel for Fujitsu Microelectronics, stated “Fujitsu felt we had the stronger patent portfolio, and we were not getting a reasonable response from them” (Wade, 1997). Neil Steinberg, associate patent counsel for Samsung, stated “Fujitsu’s expectations were not consistent with Samsung’s. We informed Fujitsu we did not use any of their technology, and now it looks like we will have to prove it in court” (Wade, 1997).

Perhaps the best example illustrating litigation as a bargaining failure is the case of Texas Instruments versus Hitachi. In 1989, a mere two years after Texas Instruments sued Hitachi for patent infringement, the two companies announced an alliance for exchanging and jointly developing memory technology. In the midst of the U.S.-Japan trade disputes at the time, American semiconductor companies were taken aback. Dan Hutcheson, president of VLSI Research, stated that from the U.S. perspective it was as if “you rip open their shirts, and there’s a red circle on their T-shirts” (Steinert-Threlkeld, 1989). Nevertheless, as Daniel L. Klesken, a semiconductor analyst with Montgomery Securities, described, the two companies acquired intimate knowledge of one another’s R&D activities as a result of having to disclose these to settle their patent litigation (Steinert-Threlkeld, 1989). As the case of Hitachi and Texas Instruments illustrates, information is revealed to both parties in the course of litigation that would have led them to discover compromises if they had known this information from the start.

Litigation reflects a bargaining failure also in the sense that it reflects two parties’ misjudgment of one another’s ‘bottom lines’. Even for firms like Texas Instruments with so called ‘litigious strategies’, the aim of asserting patents is ultimately to earn royalties. The first-best outcome is to identify the maximum royalty level to which they can push a potential licensee without having to litigate. When a potential licensee is pushed so far that it faces going out of business, it would rather commit all its resources to litigation and ‘go down fighting’. In these cases, litigation is fundamentally a result of misjudging bargaining thresholds. Jack Menache, general counsel of Integrated Device Technology, stated, “When you take the [potential] demands of the licensors and aggregate them, it’s impossible to stay in business” (Richards, 1990). Alfred Stein, chairman and CEO of VLSI Technology, stated “If one company wants 5% and another company wants 5%, pretty soon that starts to add up. We just don’t have the money to pay those kinds of royalties” (Willett, 1990).

This reveals an important problem in dividing the returns to innovation. There is only so much to go around. As Grindley and Teece (1997) describe in their case study of licensing practices, it is important for a licensor to be wary of competing demands that a licensee faces from other licensors in order to avoid unrealistically burdensome royalties. This is difficult in practice because licensors have little information on how much other licensors are demanding. If they bargain too hard, then they risk pushing the licensee to litigate rather than pay.

This makes for complex interdependencies in negotiating dyadic licensing agreements. As the next section elaborates, licensing royalties are seen as ways to offset differences in firms’ relative contributions to innovation in light of the returns they enjoy in the market for these innovations. Bargaining failures happen when one party views the other as ‘not paying its dues’ while the other believes it is. Ambiguity about how a firm is positioned in the complex network of knowledge flows and licensing obligations contributes to the risk of litigation.

Structural changes behind litigation

While the creation of the CAFC and heightened competition were the most salient changes in the industry during the 1980s, it is important to recognize the other structural changes that were taking place at this time also. The natural course of technological and industry evolution were shifting existing classes of competitors and creating new ones, such that competitive interdependencies were becoming more complex and, consequently, more difficult to manage. The following anecdotes suggest that litigation often arises from difficulty in tracing knowledge flows between companies and the ambiguity about licensing obligations that this generates. These generate complex patterns of competitive interdependency which are difficult to manage through licensing solutions. Attempts to resolve disputes break down in the face of this complexity because on the one hand, ‘reasonable’ solutions are usually not the same as what is implied by legal obligations, but on the other hand, companies are prone to disagree about precisely what is ‘reasonable’. Hence, when companies fail to reach solutions for managing these interdependencies, litigation occurs.

Changing industry structure: ascension of Japanese semiconductor companies

The first important change during this period was that companies which had previously relied on licensing technology from others were becoming innovative in their own right. In the early years of the industry, IP rights were rarely enforced, which helped the industry to grow. The lax IP enforcement resulted in wide diffusion of knowledge that generated entire classes of competitors. The most salient example is the Japanese semiconductor industry. In the 1960s and 1970s, Japanese companies relied heavily on licensed technology from American semiconductor companies to manufacture products. In the 1980s, Japanese companies' prowess in manufacturing resulted in significant market share gains at the expense of American companies, including the ones that heavily licensed technologies to Japanese companies.

In the midst of the broader Japan-US trade disputes at the time, the tension between US and Japanese semiconductor companies became framed as a simple story of those who produce versus those who steal innovations. John McDonnell, assistant general counsel for intellectual property at AT&T, stated, "I guess what happened is that Japan Inc. went to work. We had to protect what we were good at—and what this country has always been good at is research" (Thompson, 1988). Richard Agnich, general counsel for Texas Instruments, stated, "We saw [intellectual property] as a dividing line between those who had contributed to the technology and those who hadn't" (Richards, 1990).

In 1986, Texas Instruments launched a highly-publicized salvo of lawsuits against major Japanese semiconductor companies. Initially, these were widely supported by other American semiconductor companies. Michael Mayback of Intel declared that the Texas Instruments suits "are not the last" (Kehoe, 1986). Daniel L. Klesken, a semiconductor industry analyst with Montgomery Securities in San Francisco, stated:

> "I applaud their [TI's] aggressive move in January, and I'm glad that, from the sound of the agreement, the Americans have won out. Texas Instruments' actions (in January) were precedent-setting, and if they settle as successfully with the other companies in this suit as they appear to have done with Fujitsu and Sharp, it will set an even stronger precedent" (Duke, 1987).

The Semiconductor Industry Association gave "full support" for Texas Instruments' actions, stating that "The semiconductor industry's R&D investment as a percentage of sales is 10 per cent, the highest rate of any industry in the US. Earning a return on this investment is fundamental to our industry's continued ability to sustain technological leadership" (Kehoe, 1986). As these statements suggest, the growing Japanese dominance in the mid-1980s was widely portrayed as a simple case of Japanese firms copying American technology outright.

A closer look suggests that the nature of the conflict was not so simple. The problem was not so much the flow of knowledge, which had been happening since the 1960s through licensing, but the fact that Japanese firms had begun to contribute their own innovation above and beyond what was licensed from American firms. This follow-on innovation (and not just naïve copying) was responsible for Japanese firms' ascension. However, the question arose as to how much 'credit' was due to the Japanese companies that invested in the follow-on innovation and how much was due to the American companies that made it all possible in the first place.

The more calculative (though difficult) problem of dividing returns to innovation was aggravated by the feeling that American companies had unwittingly laid the foundation for their own downfall. Texas Instruments chairman Mark Sheperd Jr. launched the company's litigious strategy during a time when it was losing market share in DRAMs to Japanese competitors even though Texas Instruments held many basic patents for DRAM fabrication (Alster, 1988). The business press described how "As American companies see it, they have been pouring millions of dollars into sophisticated research and development, only to have their creations unfairly copied by overseas competitors" and portrayed aggressive moves by Texas Instruments as a "crack down" on illegal use of intellectual property (Duke, 1987). Given increased competition and decline in dominance, it is not surprising that a company like Texas Instruments would seek to earn more revenue from its patents. Before litigation Texas Instruments approached companies like Fujitsu, Toshiba, and Mitsubishi asking to increase royalty rates by five to ten times what the companies had been paying before.

However, the fact that Texas Instruments' demands were received so poorly by the Japanese licensees suggests that some more fundamental form of misjudgment had occurred. As a representative for a Japanese company described, the Japanese companies were "mad as hell" about the patent suits; as an official for another Japanese company stated, "We wanted to settle this in a friendly manner, but they have shot a pistol at us. We don't like it" (Kehoe, 1986). The reality was that these Japanese competitors had for a long time been making products under technology licenses from (and paying royalties to) American companies like Texas Instruments. At the same time, they had made substantial investments in research and development in their own right. The fact that they had neither been sitting idly nor relying passively on American technology contributed to their indignation at Texas Instruments' demands.

Between the 1970s and 1980s, Japanese semiconductor companies were undergoing a shift which, at the time of the Texas Instruments suits, was just starting to be apparent to American companies. Previously, Japanese companies were known for their ability to manufacture commodity products at low cost. Tomihiro Matsumura, senior vice president and director of NEC, acknowledged that "Generally, Japanese companies' products have been oriented toward memories and other easy products"

(Pollack, 1984). Few American companies expected that their Japanese competitors would translate this experience into developing more sophisticated, design-intensive products. In 1983, NEC surprised the industry by announcing a new internally-developed family of microprocessors that rivaled or exceeded American-designed microprocessors in performance. Previously, Japanese companies, if they made microprocessors at all, did so with licenses on American designs, licenses for which they paid royalties (Kehoe, 1984). In a dramatic role-reversal, Zilog Inc., which had just a year before accused NEC of copying its designs (and whose founder, Federico Faggin, is commonly credited as a co-inventor of the microprocessor), agreed to take out a license from NEC, producing NEC-designed chips as a second source (Pollack, 1984).

Such role-reversals, however, did not happen without friction. American companies balked at the prospect of paying for technologies they believed were basically derivatives of their own. Japanese companies similarly balked at the prospect of paying additional royalties for technology in which they had invested heavily. There was undoubtedly also a sense among Japanese firms that after their risky investments had turned out fruitful, American companies were trying to swoop in to capture the returns. This story played out with near disastrous consequences in the case between Motorola and Hitachi. Hitachi had previously made microprocessors under license from Motorola. When Hitachi announced its own line of microprocessors that rivaled Motorola's in performance, Motorola accused Hitachi of infringement. Hitachi maintained that "we developed our own H Series microprocessors with our own original and proprietary architecture" and, in fact, accused Motorola's new line of microprocessors of infringing Hitachi patents (Norman, 1989). As mentioned in the previous section, the case ended in near disaster for both companies when the federal district court judge upheld both sides' claims, declaring each guilty of infringing the other's patents and barring both from the market (Norman, 1990). The two companies subsequently negotiated an agreement, and the judge, having made his point, lifted the injunction barring the two companies from the market.

Changing industry structure: ascension of 'second-generation' semiconductor companies

In the preceding examples, the problem leading to litigation was fundamentally one of disagreement about relative contributions to innovation. The industry had reached a stage when this issue had just started to be important. The litigation between American firms and Japanese firms reflected the more complex interdependencies in knowledge flows and licensing obligations arising not just between American and Japanese firms but also between different generations of American firms. In addition to Japanese companies, a newer generation of US semiconductor companies had benefited from the initial seeding of knowledge by the first-generation US companies. These more specialized startups were developing cutting edge innovations that were successful and linked only at broad levels to the foundational technologies of first-generation companies.

Consider the example of Advanced Micro Devices and Brooktree Corporation. In the early 1980s, AMD developed a 'color palette' chip, used to control the display of color graphics on computer screens. AMD's chip was implemented in bipolar technology. During this time, chip design in the industry was shifting away from bipolar and towards CMOS, which provided design and manufacturing advantages over bipolar. In the late 1980s, Brooktree, a San Diego-based firm specializing in designing graphics chips, developed a color palette chip implemented in CMOS and received a patent for it. At around the same time, AMD also successfully designed a CMOS version of its own chip and introduced it to the market. Brooktree promptly filed suit against AMD for infringement. As Mikio Ishimaru, director of technology law and associate general counsel at AMD, describes, AMD was at the time focused on obtaining patents on broad areas of semiconductor technology rather than specific implementations (Whiting, 1992). AMD countersued for Brooktree's infringement of these broader patents. An AMD spokesman stated, "We want to send them [Brooktree] a message. We're ready to cross-license and settle these things" (Electronic Engineering Times, 1991). Eventually, the two companies did settle. As James Bixby, Brooktree's president and CEO, described:

> "All of us at Brooktree welcome this settlement, and we're very happy with the positive spirit shown by both AMD and Brooktree in settling these issues. We have buried the hatchet. We are hopeful that the resolution of this landmark intellectual property case may signify a new era of cooperation among technology companies who respect and value each other's intellectual property" (Business Wire, 1993).

Of course, this newfound respect came only after five years of expensive legal maneuvers against one another. Both continued to supply the chips in question after the settlement. In describing what was achieved through the settlement, Richard Previte, president and chief operating officer of AMD, alluded to the problem of managing upstream and downstream innovations:

> "We are pleased to bring an end to this lengthy and complex litigation. The agreement reached is fair to all parties involved. From our perspective, we are

> especially pleased that the terms of the settlement will assure that our color palette customers will continue to have access to devices that incorporate all essential technology" (Business Wire, 1993). [Emphasis added]

Unobserved motives in licensing negotiations

The friction between companies like AMD and Brooktree is not unique. A closer look at the circumstances behind lawsuits suggests that most were not targeted, legal routines executed to prevent naïve infringement of specific patents and products. Rather many were fundamentally disagreements about how to divide returns to cumulative innovation on a much broader scale and scope. As firms added increasingly more follow-on innovations on top of foundational innovations, two things happened. First, market returns shifted towards the improved products of follow-on innovators. Second, foundational innovators incorporated improvements into their products. At this point, each side had a legitimate claim on a different stage of the technology needed to make products.

From a technological point of view, each side is dependent on the other. A 'reasonable' solution would be for early and later stage innovators to divide the returns to reflect their relative contributions. However, this is difficult to do in practice because the patents that each side holds on its stage of the technology allow it to get not just a 'reasonable' share of the returns but the full amount by threatening to shut out the other party completely. As the case of the American and Japanese firms illustrates, a patent is a right to exclude, not to use, so it is possible for patent owners at different stages of a cumulative stream of innovation to block one another (Merges and Nelson, 1990; Green and Scotchmer, 1995). From a bargaining standpoint, if the two parties reach any division at all, however asymmetric, both are better off compared to the case of shutting each other out of a market. Clearly, this represents a substantial bargaining surplus.

However, dividing returns to innovation through licensing was difficult in practice in light of the fact that firms had room to be suspicious of the unobserved motives behind one another's royalty demands. On the one hand, there were plausible and legitimate motives for firms to ask for higher royalties. Early pioneers of the semiconductor industry contributed much to the growth of the industry, both in developing foundational technologies and, more importantly, in allowing these to diffuse widely without stringent intellectual property controls. As Carl Silverman, chief counsel for Intel, stated, "We are the company that brought EPROMs to the world. We believe it's our responsibility to our shareholders to absolutely protect that investment" (Harbert, 1990). From its litigation settlements with several Japanese companies, Texas Instruments earned an apparent windfall of $191 million in 1987. However, William P. Weber, president of Texas Instruments' worldwide semiconductor operations, called these royalties mere "peanuts" compared to what the Japanese companies got from Texas Instruments: the invention of the integrated circuit (Steinert-Threlkeld, 1989). As Texas Instruments spokesman, Stan Victor, reminded critics, "this created an entire industry" (Shannon, 1988). With respect to the company's technology more generally, Victor, stated, "We think we made the investment and the inventions over the years that have made possible a lot of these businesses" (Fisher, 1990). AT&T can make a similar claim; the company is estimated to have received only $20 million is license fees for the invention of the transistor in 1947 (Pollack, 1988). However, in 1988, Texas Instruments ranked only fifth in the world in semiconductor sales, with $2.75 billion out of a total of $50.2 billion in industry sales (Steinert-Threlkeld, 1989). The top-ranked company was NEC Corporation with $4.65 billion in sales.

In addition to this ‘moral’ justification, there were other legitimate reasons behind raising royalties. An important reason was that the true costs of innovation were rising, making it difficult to continue funding R&D without recouping a larger share of past investments. Referring to Texas Instruments’ suit against Micron Technology, Texas Instruments’ Richard Agnich, commented, “TI has invested heavily in the research and development that generated our DRAM technology… We must obtain an adequate return to justify further R&D investments, and we gain that return from selling DRAMs and from royalties paid by others who use our patents” (PR Newswire, 1988). Between the period when the first generation of semiconductor companies were founded and the period of the 1980s, marginal returns to R&D investment had decreased, while fixed costs had increased. Thomas Dunlap, general counsel for Intel, observed “If you go back to the late 1970s and early ‘80s, a new design cost about $1 million” (Weber, 1991). Intel’s Andy Grove observed that while an early Intel random access memory in 1972 cost a few hundred thousand dollars to develop, a comparable leap in technology in the late 1980s would cost tens of millions of dollars (Wiegner, 1987). Mel Thomsen, analyst with Dataquest, stated “The protection of intellectual property is becoming much more important. You’re not going to replace your current technology as rapidly. You want to be much more protective of the technology you have. So all these people are dusting off their patents” (Electronics Times, 1986).

In addition to rising development costs, equipment costs were rising, even while product life cycles were shortening. A Dataquest Inc. report stated that while the marginal cost for adding a new wafer fabrication machine was $12,000 in 1968, the cost had risen to $1.2 million over the course of twenty years (Thompson, 1988). The total cost for a state of the art wafer fabrication facility had risen to $100 million. The same Dataquest report observed that while typical semiconductor logic families introduced between 1965 and 1982 had life spans of 5.6 years, microprocessor families, first introduced in 1972, have life spans of 4 years, and application specific integrated circuits, first introduced in 1983, have life spans of only 1.3 years.

On the other hand, however, there were also plausible and illegitimate motivations behind firms' royalty demands. In the semiconductor industry, infringement is more often incidental than malicious. For instance, Neil Steinberg, associate patent counsel for Samsung, observed that "We don't evaluate every patent, we just develop our own processes and chips, and sometimes people stumble upon each other's technology" (Wade, 1997). However, it is possible for companies to exploit this to their advantage. When Texas Instruments sued Micron Technology in 1988, it sought punitive (in addition to compensatory) damages 'for willful infringement' on the part of Micron (Electronic Engineering Times, 1988). Richard J. Agnich stated that "TI has taken this action to prevent unauthorized use of its technology" (PR Newswire, 1988). In contrast to Texas Instruments' accusations, the prevailing norm in the industry up to that point was for technologies to be licensed liberally and for patent owners to be lenient in the case of (inevitable) incidental infringement. By the standards of the time, Micron Technology had done little that was out of the ordinary. A Texas Instruments spokesman indirectly acknowledged this, stating that it is “difficult, if not impossible, to build a DRAM without violating one or more of these very basic patents” (Electronic Engineering Times, 1988). While from a ‘common sense’ point of view this undermines Texas Instruments’ accusation that Micron’s infringement was of a malicious nature, the accusation of ‘willful infringement’ provides a patent owner with a way of inflating damage awards in court. Licensing negotiations between Texas Instruments and Micron broke down after Texas Instruments asked for triple the typical royalty rate since this is how much it could (hypothetically) win in court by invoking ‘willful infringement’, and not because Micron was maliciously stealing Texas Instruments technology, as many US companies at the time believed Japanese companies were doing. Litigation occurred as a result of this licensing breakdown.

More generally, licensing was difficult because of each firm’s suspicion that the other side’s demands were based on legalistic considerations rather than technological ones. The risk of incidental infringement was so pervasive in the semiconductor industry that firms could exploit the law to their advantage should they choose. Ronald Laurie, San Francisco computer attorney suggested that “If you have good patents, litigation is a better way of making money than selling products” (Pollack, 1988). The fact that the innovations developed by second-generation semiconductor companies were linked only at broad levels to the foundational technologies of first-generation companies fueled suspicions that first-generation companies were asserting distantly-related patents to extract easy revenue rather than for legitimate technological reasons. Bruce Entin, vice president for investor relations at LSI Logic, called Texas Instruments’ tactics “tantamount to extortion” (Weber, 1991). James Menache, general counsel of Integrated

Device Technology, accused companies like Texas Instruments and AT&T of being "opportunists in exploiting a flaw in our patent laws" (Menache, 1991).

Credibility of a firm's commitment to ongoing innovation

For firms like LSI Logic and Integrated Device Technology, suspicion of licensor motives extended well beyond the particular patents and products in question. This can be seen in the stigma associated with being accused of using the legalistic patent strategy.

Invoking legalistic, rather than technological concepts to inflate royalty demands from patents was often seen as a disqualification of a firm's commitment to innovation. T.J. Rodgers, president and chief executive of Cypress, stated "TI is a bunch of lawyers out of control. They have a gun in your stomach and are saying, 'Give me your wallet.'" (Weber, 1991). LSI Logic vice president, Bruce Entin, commented that "It's a pity we're at the point where TI is litigating instead of innovating" (Weber, 1991) and accused Texas Instruments of relying on it as a "substitute for innovation to produce profits" (Fisher, 1990).

The targets of high royalty demands contrasted themselves with companies they viewed as having 'fallen from grace' as genuine innovators. Stephen Cordial, vice president and chief financial officer of Sierra Semiconductor, stated, "We'd rather spend money on R&D than on litigation—We think of ourselves as an old-fashioned, technology-driven company that makes money by building a better mousetrap" (Resnick, 1991). Jerry Rogers, president and chief executive of Cyrix Corporation vowed not to play the intellectual property 'game', stating "We're just going to out-innovate our competitors" (Whiting, 1992). Even Intel chief counsel, Carl Silverman, commented, “Only counting on lawyers, and legal concepts such as ‘intellectual property,’ is not a winning position” (Wiegner, 1987). Ironically, doubt about Texas Instruments’ underlying credibility as an innovator led some to reinterpret its earlier lawsuits against Japanese firms. For instance, T.J. Rodgers of Cypress Semiconductor criticized companies like Texas Instruments that “wrapped themselves in the flag and accused the Japanese of unfair competition” (Kehoe, 1990). The challenge to Texas Instruments’ commitment to innovation became salient enough that Texas Instruments chairman, Jerry R. Junkins, defended the company’s actions, stating that royalty revenue was going right back into research and development (Steinert-Threlkeld, 1990).

It is important to recognize that even firms on the receiving end of royalty demands did not disagree fundamentally with the idea that licenses ought to be taken out and royalties paid to firms that were indeed genuine innovators. As LSI Logic’s Bruce Entin pointed out, “We agree that intellectual property should be protected. We don’t agree that you should lose perspective on it and use it as an extortion tactic to get excessive amounts of money” (Weber, 1991).

The fundament point of disagreement was about the extent to which a given licensor was genuinely committed to innovation and what royalty level was appropriate in light of this. Because some firms may be more committed to and invest more in ongoing innovation, there is the potential for licensing partners to benefit asymmetrically from a licensing agreement. This possibility induced firms to be discriminating in setting licensing terms, using royalty levels to offset perceived disparities in their ongoing innovative contributions. However, firms could not distinguish based on the aggressiveness of one another’s demands whether these were meant to provide funding for ongoing innovation or an easy revenue stream that could be maintained without investment in innovation. Hence, as discussed in the previous section, licensing negotiations often broke down over disagreement about what constitutes 'reasonable royalty' levels.

At this point, it is important to clarify more explicitly the way I have portrayed licensing negotiations. This will have bearing on how we interpret litigation events. Formally, licensing agreements give firms access to one another's patents. In an ideal world, the only things that should have bearing on licensing terms are the values of the underlying technologies and boundaries around intellectual property rights as defined by the text of firms' patents. In practice, a number of factors cause the reality of licensing to differ from the ideal world.

Patent rights provide imperfect and possibly misleading indicators of firms' true contributions to innovation. The patent system has been criticized for its lax standards in granting patents. At the same time, firms have been criticized for taking advantage of lax patenting standards by padding their patent portfolios with low-value or worthless patents (Moore, 2005). If firms were to adhere strictly to the patent record in setting licensing terms, the direction and level of royalty payments would likely violate commonsense notions of fair return for investment in innovation. The anecdotal evidence in the preceding sections is meant to illustrate this. In the case of a cumulative technology such as semiconductors, firms recognize that they benefit from one another's investments in innovation in ways that do not map well to how extensively they infringe on one another's patents from a legal standpoint. Licensing negotiations can serve as occasions for firms to equalize differences in their contributions to the cumulative streams of innovation from which they all benefit.

In trying to equalize these differences, firms care about signals of one another's commitment to innovation because their true innovative contributions are difficult to assess. Patents provide a poor indicator. Assessing the actual innovative activities behind patents is rarely done both because it is costly and because it requires scrutiny of internal information that firms usually keep confidential. Objective assessments of firms' innovative activities are rarely available during private licensing negotiations. In fact, usually the only time this is truly done is in the discovery phase of litigation. As illustrated in the preceding anecdotes, this information asymmetry can generate suspicion between the negotiating parties that the opposing side is over-representing its innovative contributions and demanding excessive royalties. Firms care about signals of one another's ongoing commitment to investing in innovation because this helps them to discriminate between genuine demands (fair return on innovative contributions) and bluffs (excessive returns).

Another reason that firms care about signals of one another's ongoing commitments to innovation is that licensing agreements are at least partly forward-looking. In the semiconductor industry (and many others also), R&D spillovers between firms are widespread and inevitable. By granting access to one another's patents, licensing agreements are essentially patent truces, allowing spillovers to happen without threat of litigation. Access to a firm's patent portfolio is only as good as the innovations behind it. A firm stands to overpay if it agrees to low royalties from a licensing partner that subsequently invests and contributes little in innovation—the partner benefits more from the firm than the firm benefits from the partner. Another reason that firms care about one another's ongoing commitment to innovation is that technologies obsolesce faster than patents expire. In the case of cumulative technology, the value of an initial invention depends on costly investments into follow-on and complementary innovations. A firm stands to overpay if it agrees to pay high royalties to a licensing partner that makes little further investment but continues to enjoy royalty rent. Firms are therefore reluctant to lock in licensing terms when they doubt one another's ongoing commitments to innovation. For a firm to maintain high-quality output into the future is not effortless but requires ongoing investment. This is a reason why, even with objective information about a firm's past innovations, licensing partners care about credible signals of commitment to future innovations. As the preceding anecdotes illustrate, these were actual concerns voiced by firms in the semiconductor industry.

To this point, I have used anecdotal evidence to support my portrayal of licensing and litigation. In the conclusion to this chapter, I discuss how my portrayal of licensing and litigation is consistent with existing literature on patenting.

Discussion

Based on the anecdotal evidence, I have suggested that litigation often centers on two related problems: the difficulty in tracing and measuring knowledge flows between companies, and ambiguity about the extent to which a given company is genuinely committed to innovation. The flow of knowledge between companies generates complex patterns of competitive interdependency which are difficult to manage through licensing solutions. Attempts to resolve disputes break down in the face of this complexity because, on the one hand, ‘reasonable’ solutions are usually not the same as what is implied by legal obligations, but on the other hand, companies are prone to disagree about precisely what is ‘reasonable’. In addition, because some firms may be more committed to and invest more in ongoing innovation, there is the potential for licensing partners to benefit asymmetrically from a licensing agreement. This possibility induces firms to be discriminating in setting licensing terms, but at the same, however, firms have difficulty distinguishing whether one another’s demands are meant to provide funding for ongoing innovation or an easy revenue stream that can be maintained without investment in innovation. Hence, litigation reflects failed attempts to discriminate in the face of ambiguity.

This suggests that another way to view patent litigation is that it provides firms with an expensive way to overcome this ambiguity. The cost of patent litigation is typically on the order of millions of dollars, a sharp contrast to the thousands of dollars it costs to obtain a patent (Moore, 2005). What firms receive in return is a way to verify one another’s claims in a dispute. After an infringement suit is filed, the case goes into the discovery phase in which both parties are required to disclose evidence supporting their claims. During the discovery phase, the disputing parties get perhaps their first detailed look at the broader R&D activities behind one another’s patents. Among disputes that reach litigation, few make it past pretrial procedures. The vast majority (95 percent) are settled, usually in the form of a licensing or cross-licensing agreement (Lanjouw & Schankerman, 2001). Hence, economic models of litigation often portray it as a bargaining failure, arising from the inability of the parties involved to credibly signal hidden information that could facilitate settlement (Bebchuk, 1984; Meurer, 1989; Schweizer, 1989).

The semiconductor industry is remarkable in that instances of litigation are relatively rare considering the propensity for infringement and bargaining hazards that make settlement difficult. Economists and legal scholars have raised concern that in highly crowded technological spaces, patent rights have become so densely packed that it may be impossible for competitors to avoid infringing one another's patents (Heller & Eisenberg, 1998; Lemley & Shapiro, 2007; Ziedonis, 2004). The problem of 'patent thickets' is particularly pronounced in cumulative technologies such as semiconductors, since final products embody systems of related technologies patented by different competitors (Merges & Nelson, 1990; Shapiro, 2001). In these settings, firms face the prospect of having to seek out numerous competitors to negotiate licensing agreements in order to avoid infringement litigation (Ziedonis, 2004). Anecdotal evidence suggests that competitors arrive at these agreements quite regularly. Competitors at most risk of infringing on one another's patents frequently enter into broad cross-licensing agreements, so that when disputes do arise, these do not escalate to litigation (Grindley & Teece, 1997).

While it is not surprising that competitors arrive at these solutions, it is remarkable that they do so in a way that accounts for unobservable differences in the quality of their patent portfolios. Anecdotal evidence suggests that firms often set balancing payments to offset perceived disparities in their patent portfolios (Grindley & Teece, 1997). Firms may also account for disparities in other ways such as by licensing only select portions of their patent portfolios to a given competitor. When firms attempt to even out these perceived differences, as opposed to simply calling litigation truces, all the hazards associated uncertain patent value, holdup threat, and hidden motives come to bear on the bargaining process.

In cumulative technologies such as semiconductors, bargaining covers much broader and more poorly defined technological and market spaces than it does in discrete technologies. Cross-licensing negotiations center on anticipated infringement, which in the case of cumulative technology may involve products that have yet to be developed (Grindley & Teece, 1997). In addition, technological progress happens as a function of many interrelated discoveries, whose individual contributions to overall technological progress are difficult to trace and separate. Hence, in the case of cumulative technology, firms are generally not bargaining over lost sales due to simple product imitation. Rather, they are bargaining over how to divide the oligopolistic pie in accordance to their innovative contributions to the industry (Cohen, Nelson, & Walsh, 2000).

These factors make bargaining difficult, since firms generally have hidden information about the R&D activities behind their patents. Some firms may contribute little serious innovation to the industry and instead rely on the holdup threat of their patent portfolios to inflate royalty demands against competitors. Other firms may contribute genuine innovation to the industry and genuinely wish to appropriate some of the returns to these contributions. The problem is that while a firm knows its own intentions and the true quality of its patents, its competitors do not. So when a firm negotiates licensing terms with competitors, it must somehow justify its demands in the face of hidden motives.

These attributes of the semiconductor industry exacerbate the hazards normally associated with bargaining over patent rights. It is difficult for competitors to bargaining over patent rights because each patent's underlying value is difficult to assess and, yet, each patent represents an equal threat to litigation. Economists have long observed that there is high variation in the economic value of patents (e.g. Trajtenberg, 1990). A small proportion of patents turn out to be extremely valuable, while most turn out to have little or no economic value. Lemley and Shapiro (2005) have portrayed patents as 'lottery tickets'. Individually, each patent costs little to obtain and has low odds of being valuable, so firms acquire as many as possible to improve their chances of hitting the jackpot. In the semiconductor industry, firms engage in the practice of patent 'harvesting', essentially maximizing the patent yield of their R&D activity (Hall & Ziedonis, 2001). The various elements of any given invention will usually end up being covered by a field of interrelated patents. Moreover, because innovation in semiconductor technology is cumulative, these patents are likely to be owned by many different competitors (Ziedonis, 2004). A side-effect of the patent harvesting strategy is that competitors flood the industry with low-value patents covering closely related content, and these can dilute the effect of patent ownership as a signal of genuine innovativeness.

Bargaining over patents is hazardous because each patent represents a right to litigate, regardless of the value of its underlying technology. Given the volume of patenting in industries such as semiconductors, it is virtually impossible for firms to evaluate their products explicitly for infringement of every potentially relevant patent. Whenever they invest in facilities and equipment for a product, firms face the risk of litigation by competitors whose patents they had not considered or licensed in advance.

Where high sunk costs are involved, a patent need not be extremely valuable to pose a 'holdup' threat (Hall & Ziedonis, 2001; Lemley & Shapiro, 2007). A patent on a relatively minor component of a final product could just as easily stall production as a patent on a major component, so long as facilities and equipment are difficult to modify. This holdup threat distorts the value of a patent in the sense that it potentially allows the patent owner to extract higher royalties than the 'true' value of the underlying technology. More broadly, the practice of patent harvesting may allow firms to inflate the value of their patent portfolios by padding them with low-valued but equally-threatening patents.

These features make this context ideal for studying the effects of status. When disputing firms make unverifiable claims about the quality of their innovations, the presence of strong structural constraints may help restrict their claims to non-overlapping quality ranges and align their expectations about reasonable settlement terms. In other words, structural constraints can operate as a mutually-understood signal of the relative quality of two firms' innovations. In the semiconductor industry, structural constraints can be quite strong. Inventions in semiconductor technology usually begin far from commercialization. Many inventions that have not gained acceptance may be potential breakthroughs. However, they require substantial follow-on R&D to realize their full potential, and this is often beyond the capability of any single inventor (Green & Scotchmer, 1995; Merges & Nelson, 1990). More generally, semiconductor technology is often subject to network externalities (Katz & Shapiro, 1985). The more firms there are doing R&D to improve a technology, the more valuable it becomes. For instance, the success of a microprocessor design depends heavily on the simultaneous development of peripheral components with which it must operate and alternative suppliers of compatible replacements (Wade, 1995, 1996). These properties make semiconductor technology susceptible to the kinds of self-fulfilling prophecies characteristic of status hierarchies (Podolny & Stuart, 1995).

Chapter 3: Structural forces in the context of innovation

Introduction

In the previous chapter, I suggested that firms care about each other's ongoing commitment to investing in innovation because this provides a basis for judging the credibility of each other's claims of innovative contribution. Firms care about dividing returns to innovation in ways that correspond to their actual innovative contributions rather than to a strictly legal reading of their patents. However, the reality of each firm's innovative contributions is difficult to discern. In this chapter, I use anecdotal evidence to illustrate more precisely why this is the case.

I also use the anecdotes in this chapter to introduce the idea that structural forces can lock firms into positions of ongoing investment in a technology. By structural forces, I mean that social consensus about which firms are technological leaders and, hence, which firms' technologies are worth adopting can become a self-fulfilling prophecy. Firms that attract more investment in follow-on and complementary innovation by others in the industry are more likely to see their innovations become truly valuable. Firms in these positions also have a greater incentive to invest in innovation because they enjoy the promise of better rewards. And because a firm's position is conferred by the group, the position is both publicly-recognized and beyond the firm's direct control. Therefore, a firm's position becomes a credible signal of its commitment to investing in innovation and, hence, a signal of the actual, unobserved innovative contributions behind its patents.

To motivate this argument, I examine the circumstances surrounding the invention of the microprocessor. Perhaps the most familiar part of the story of the microprocessor is that in November of 1971, Intel introduced the 4004, the world's first microprocessor, the work of Intel engineers, Marcian 'Ted' Hoff and Federico Faggin. Intel subsequently became the world's leading producer of microprocessors. The less familiar parts of the story are the circumstances leading up the introduction of the 4004 and the more influential 8008. It involved companies and individuals whose names have never been associated with the microprocessor but who played key roles in its inception.

Examining the details of this story illustrates a number of points that will be important in subsequent chapters. First, the case illustrates how digging into the facts often reveals a reality that is more complex than industry participants subsequently recognize and that would make dividing credit for innovation difficult. Second, the case illustrates how a firm can come to occupy a leadership role in the industry through fortuitous circumstances rather than superior vision or innovative ability. Finally, the case illustrates how such a firm is subsequently induced to 'play the part' appropriate for its role.

The historical facts of how a technology was pioneered become lost over time (if they ever came to light at all in the first place). In dividing credit for innovation, firms rarely dig into these details and instead rely on a simplified but nonetheless sensible picture of the world. I suggest this picture consists of firms' shared understanding of one another's levels of commitment to the ongoing development of a technology. It substitutes for the more complex technical and legal relationships that link firms in the context of cumulative innovation. This chapter in conjunction with the previous chapter provides anecdotal evidence that this is a reasonable characterization of how firms divide returns to innovation in practice. In the next chapter, I conceptualize these positions and shared understandings as elements of a status hierarchy among firms. This allows me to generate testable predictions which provide empirical support for my characterization of how firms divide returns to innovation in practice.

Where did the impetus for the microprocessor come from?

The creation of the microprocessor in the early 1970s was in many ways inevitable. It was a matter of who would be in the right place at the right time to do it. The idea of putting a computer on a single chip had been around since at least the mid-1960s and was well-known by the end of the 1960s when the events leading up to the invention of the microprocessor took place. At the time, there was an ongoing debate between two paradigms for calculator design: one advocating the use of custom circuits and one advocating the use of standard circuits (Faggin, 1992). The majority of designers at the time supported the use of custom circuits. The argument for using custom chips was that general-purpose calculator chips would need to include too many hardwired options (that are not always utilized) and so the resulting calculator design would be bigger and less cost-effective than one using chips tailored to the needs of the particular calculator. The argument for standard chips was that they could be made versatile and cost-effective if, instead of hardwiring functions, they were designed instead as programmable computing chips. Federico Faggin, commonly credited as one of the fathers of microprocessor, observed that the leading semiconductor companies of the time had already made advancements in implementing the idea of a CPU on a chip:

> "Fairchild had already done pioneering work in this area, developing a 1-bit serial CPU architecture, as had Rockwell, where Michael Ebertin and his coworkers designed a more sophisticated CPU." (Faggin, 1992)

At the same time, broader developments in design and manufacturing led to rapidly increasing levels of integration. As Gordon Moore observed in 1965, the number of components that could be integrated onto a single chip had doubled every year since the invention of the integrated circuit in 1959 (Moore, 1965). In the mid-1960s, central processing units (CPUs) used circuit boards that already included chips of small- to medium-scale integration (tens to hundreds of components, respectively). Everyone knew that large-scale integration (thousands of components per chip) was coming, and "it was just a matter of time until a CPU on a chip appeared" (Faggin, 1992).

In the late 1960s, several individuals at different corners of the globe, working on different problems, arrived at the common idea of integrating the various components of a CPU onto a single chip as the solution to their problems. Here I discuss the individuals and companies whose involvement, directly and indirectly, culminated in the creation of the first microprocessor at Intel. These anecdotes illustrate how the primary impetus for the invention of the microprocessor originated in a variety of individuals, organizations, and circumstances rather in Intel alone.

Sharp and Busicom

One of the less-known individuals who played a role in the creation of the microprocessor was Tadashi Sasaki, an engineer and senior manager at Sharp. Before joining Sharp, Sasaki had been an electrical engineer at Fujitsu who had established

Fujitsu's semiconductor research program after a meeting and several communications with future co-inventor of the transistor, John Bardeen of Bell Laboratories (Aspray, 1997). Sasaki joined Sharp in 1964 as a senior manager. In the 1960s, Sharp was a leading supplier of domestic appliances in Japan. Under Sasaki's urging, Sharp entered the calculator business and produced the first all-transistor (as opposed to vacuum-tube) calculator.

As a result of increasing integration, the number of chips needed to implement the functions in a calculator dropped predictably through the 1960s, reaching as little as four chips in 1968. Sometime during this period, one of Sasaki's engineers suggested to him that if the trend continued, the number of chips in a calculator would eventually drop to a fractional number. By fractional, the engineer apparently meant that there would be functionally differentiated regions within a single chip (the distinguishing feature of modern CPUs). Sasaki liked the idea and began looking for a semiconductor company to design and produce a chip containing functional divisions. He first approached Rockwell. At the time, Rockwell was the exclusive supplier of semiconductor devices for Sharp's appliance business. However, Rockwell refused, "according to Sasaki because Rockwell was already earning high profits with its other semiconductor devices and did not want this distraction" (Aspray, 1997).

Then in 1968, Sasaki was approached by Robert Noyce. Intel had been founded just a few months earlier, and Noyce was on a trip to build up the struggling startup's client list (for memory devices). Sasaki had admired and drew on Noyce's work at Fairchild and so wanted to help the new startup (Aspray, 1997). However, citing its exclusivity agreement, Rockwell refused to allow any purchases, however small, from

Intel. This led Sasaki to approach Busicom, a Japanese calculator company whose president at the time was a graduate of the same university department as Sasaki. Sasaki secretly provided 40 million yen under the agreement that Busicom would contract with Intel to make a four-division chip. In September 1969, Busicom sent Intel a contract for ten custom circuits for use in a low-cost desktop printing calculator (Faggin, 1992).

Computer Terminal Corporation

The other stream of events that culminated in Intel's creation of the microprocessor began with Austin Roche, a former NASA engineer who helped found the Computer Terminal Corporation (CTC) in San Antonio, Texas. The concept of building a programmable personal computer had been Roche's obsession. In 1968, Roche reportedly drew out the plans for what would ultimately become the Intel 8008 on tablecloths in a private club in San Antonio (Wood, 2008). However, the concept was still unfamiliar at the time the company was founded in 1967. So instead the company secured financing on the pretense of designing teletype machines.

In order to ensure a place for their product in existing markets, CTC founders promoted their computer as a replacement for the IBM 029 punch machine (Wood, 2008). This had important implications for CTC's subsequent contract with Intel. In order to maintain aesthetic similarity to the IBM 029 in the eyes of customers, designers tried to reduce the overall size of the monitor and casing for their computer. This compactness, however, had the side effect of generating heat problems. To overcome this, designers were forced to look for ways to reduce the number of chips used in the design, including in the CPU board (Wood, 2008). In December of 1969, Vic Poor, vice president of research at CTC, approached Intel to develop a processor for the new computer with the goal of integrating the processor's roughly one hundred logic circuits onto as few chips as possible (Manners, 1996).

Creation of the microprocessor: reluctance and delay at Intel

In late 1969, Busicom and Computer Terminal Corporation set in motion the sequence of events leading up to the creation of the first microprocessor at Intel. As the following sections illustrate, however, Intel as an organization played a consistently reluctant role in this process. That these events catapulted Intel to success was largely a product of fortuitous circumstances rather than foresight and vision on Intel's part.

Work on both projects at Intel languished from the start. Intel's management was initially resistant to taking the projects and afterwards remained resistant to diverting manpower to work on them. Both Busicom and CTC contracts were for the design of custom circuits, a common line of business in the semiconductor industry but something well-beyond what Intel management felt was their primary business of developing memory chips.

CTC's Roche had initially presented his conception of the CPU chip to Noyce with a suggestion that Intel design the chip at its own expense (Wood, 2008). Roche felt that Intel would jump at being handed the potentially revolutionary idea. By designing the chip at its own expense, Roche suggested, Intel could keep the design proprietary and subsequently sell the chip to other customers besides CTC. Vic Poor recalled, "Intel would have none of it. They felt it was beyond the state of the art, that it would consume too much of their resources, they weren't in the business of doing 'specials', etc., etc." (Gardner, 1990). In addition, CTC's John Frassanito recalled:

> "Noyce said it was an intriguing idea, and that Intel could do it, but it would be a dumb move. He said that if you have a computer chip, you can only sell one chip per computer, while with memory, you can sell hundreds of chips per computer" (Wood, 2008).

However, in 1969, Intel was a struggling startup needing cash to fund further development of its core memory business. As Ted Hoff recalled:

> "When you develop a new memory chip of your own design… You are ready to make them by the millions and people are just buying them by the ones and twos to try them out. So the business of trying to get around that is doing custom silicon." (Aspray, 1997)

With no steady customers and no money coming in, Intel reluctantly took up the Busicom contract with the hope of building up its client base. Similarly, CTC used a lot of semiconductor memories in its products and had to wield its clout as a large customer to convince Intel to take the project (Gardner, 1990). Noyce also rejected the idea of developing the chip at Intel's expense, even in return for the rights to the design. Intel and CTC ultimately agreed to a $50,000 development contract after which CTC would keep the rights to the design (Wood, 2008).

In addition to Intel's reluctance to commit to the projects, Intel at the time also lacked the manpower to complete the projects to the two companies' specifications and timelines. Federico Faggin recalled thinking at the time Intel accepted the contracts, "Intel was in no position to bid on the contract. The company had no in-house expertise in random-logic design, and it would have taken too many engineers to do the work" (Faggin, 1992). This realization came as a surprise to both Busicom and CTC. Masatoshi Shima, Busicom's engineer in charge of managing the project at Intel, estimated that Busicom's design team had already completed 80 to 90 percent of the work in laying out the architecture, leaving Intel the straightforward task of translating the designs into circuits. As Shima recalled later:

> "In the beginning of the meeting[s] we thought that once we showed the logic schematic to them, they could understand what we wanted to do. But after several meetings, we found out that they didn't have a logic engineer to understand our logic schematics and they didn't have any circuit engineer to convert our logic schematics to circuit schematics… Also they didn't know about desktop calculators themselves." (Aspray, 1997)

Little work was done, and Intel management continued to focus on the company's memory business. Faggin recalled, "Shima was furious when he found that no work had been done since his visit approximately six months earlier" (Manners, 1996). Similarly, CTC's Vic Poor recalled that Intel was so behind schedule that CTC was forced in the meantime to design a version of the chip in-house in order to stay on schedule with the rollout of its new computer line (Gardner, 1990).

Fortuitous side-effects

Both delays, however, had unforeseen side-effects that ultimately worked in Intel's favor. The architectural decisions leading up to the invention of the microprocessor arose (at least in part) from design tradeoffs necessitated by Intel's inability to fulfill Busicom's specifications. In addition, Intel's eventual ownership of rights for its microprocessor resulted from missing its deadline for CTC.

In the Busicom project, the complexity of implementing Busicom's design, coupled with Intel's shortage of logic and circuit engineers, forced Ted Hoff to rethink a simpler design strategy—one that would become standard for modern microprocessors. Hoff had no experience with calculators but did have previous experience using Digital Equipment Corporation's PDP-8 computer to do his research (Aspray, 1997). The PDP-8 relied on a small instruction set, which made for a simpler architecture but, as a tradeoff, demanded more memory. Based on this example, Hoff reportedly saw a chance to both relax the workload demands on Intel's design staff and take advantage of the company's strength in making high-capacity, low-cost memories (Aspray, 1997). The lack of design expertise at Intel also had the fortunate side-effect of forcing Intel to hire Federico Faggin from Fairchild. Although Faggin was the most technically-qualified choice to bring in for the project, Les Vadasz, then head of Intel's design team, was initially reluctant due to friction between the two when they worked together at Fairchild (Aspray, 1997). Nevertheless, in the spring of 1970, Federico Faggin was hired from Fairchild for the express purpose of completing what was by then a delinquent contract for Busicom (Manners, 1996). Working furiously, Faggin finished a working chip, the 4004, in February 1971 and delivered full kits to Busicom in March 1971. The 4004 would become commonly recognized as the world's first microprocessor.

In the CTC project, the delay at Intel forced CTC to develop an in-house version of the CPU to use in its new computer line. In March 1970, CTC's Vic Poor built a less ambitious but working model of the CPU board using bipolar logic components instead of an integrated CPU chip (Gardner, 1990). CTC made the first sales of the Datapoint 2200 desktop computer in May 1970 (Wood, 1008) and delivered the first production-line model to Pillsbury Foods' chicken farm plant in El Dorado, Arkansas in April 1971 (Gardner, 1996). By this time, Faggin had finished work on the Busicom project and was put to work to resume the CTC project (Faggin, 1992). However, the economic recession of 1970 had dropped the prices of bipolar components. This made CTC's in-house design economically more attractive than the prospective Intel design. CTC was no longer interested in completing the project. Intel nevertheless decided to complete the project after Seiko expressed interest in using the chip for a scientific calculator (Wood, 2008). CTC agreed to give Intel the rights to the chip's architecture in return for Intel waiving the $50,000 development fee (Wood, 2008). The chip was renamed and introduced in April 1972 as the Intel 8008 (Faggin, 1992). Although technically introduced after the 4004, the two projects began at nearly the same time, and the 8008 architecture, much more so than the 4004, became the foundation for subsequent Intel designs.

What exactly was Intel's contribution?

The preceding sections suggest caution in characterizing Intel's role in the inventive process. Broader trends in semiconductor technology had made the general concept behind the microprocessor common knowledge and the eventual implementation of the concept inevitable. Moreover, the impetus for the work leading up to the microprocessor not only came from outside Intel but was met with resistance from Intel management.

Once the projects began, a substantial amount of work continued to be done by the two client firms. Since both projects began as custom-chip contracts on which the client firm would keep the rights, there was substantial sharing of designs between the client firms' designers and Intel designers. As a result, much of the design work behind the architecture of the Intel 4004 and Intel 8008 was done by Busicom and CTC designers, respectively. Vic Poor observed that the 8008 was an "absolute copy" of CTC's 2200 design (Gardner, 1996). Moreover, since Intel's subsequent microprocessor families developed from the 8008, Poor observed that Intel's 80386 contains the same fundamental architecture as CTC's design (Gardner, 1990).

Structural forces behind Intel's rise

To be fair, the inventive step taken at Intel, while cumulative, was an important one above the foundation that the two client firms had laid. Engineers at both Busicom and CTC did a great deal of work in laying out the architecture for the chips that became the 4004 and 8008, respectively. The task that remained, integrating the various functions onto a single chip, was still a nontrivial one. It took a substantial amount of further work, much of it done by Federico Faggin, to accomplish this. As Faggin stated, "An invention requires reduction to practice. It can't just be an idea. The idea of a CPU on a chip had been around since the mid-60s…but making it work was the trick" (Manners, 1996).

The question this raises, however, is what impact (if any) this experience had on Intel's internal expertise and commitment to the microprocessor market. Clearly, Intel subsequently became the industry leader in microprocessors. Through what mechanism did the company's initial experience translate into this subsequent position?

There are two broad answers to this question. The first is that, while initially lacking in expertise and resistant to the concept, Intel's fortuitous involvement with the invention of the microprocessor nevertheless led to the birth of superior internal expertise and vision, which Intel exploited by committing to the new market. The second is that the overwhelming opportunity created for Intel by its fortuitous involvement in the invention of the microprocessor induced the organization to commit to the new market *in spite of* its lack of superior expertise and vision. The evidence suggests the latter rather than the former was the case.

The first piece of evidence to consider is that Intel management remained reluctant to commit to the market even after the successful creation of the 4004 and 8008. After completing the 4004, Faggin expected that Intel management would jump at the opportunity to market it for a broad range of applications (Faggin, 1992). However, Intel management believed it was only good for calculators. Faggin recalled, "I was kicking and screaming for management to do more with the 4004" and had to look for opportunities to demonstrate its broader applications (Gardner, 1990). An additional obstacle that Faggin pushed to overcome was to obtain rights to the 4004 from Busicom. The 4004 project began as a custom job for Busicom, wherein Busicom had exclusive rights to the completed chip. Faggin recalled, "I suggested that perhaps Intel could trade some price concessions for nonexclusivity. I had heard from Shima that Busicom was hurting in the marketplace and needed a lower price to effectively compete" (Faggin, 1992). However, many in the company doubted the scale of the market. Hoff recalled:

> "One marketing guy said, 'the total sales of minicomputers is 20,000. We are latecomers to the business, so we will be lucky to get 10% [of the business]. 2,000 computers is not worth all this.'" (Aspray, 1997)

Intel management still viewed itself as primarily a memory business. One argument that helped to sway management was that since Intel's microprocessor required a lot of supporting memory, promoting microprocessors could be a way to grow the market for Intel's memory products (Aspray, 1997). Noyce was ultimately able to negotiate rights from Busicom for non-calculator applications of the 4004 (Gardner, 1990). Intel introduced the 4004 family in November 1971.

The launch of the 4004, however, was not the end of Intel management's reluctance to commit to the new market. For instance, Intel management, as late as 1973, feared that if the company put a CPU chip in its catalog, Intel's computer vendor customers would see Intel as a competitor and go elsewhere for memory chips (Wood, 2008). In addition, Faggin met with resistance when in 1972, he began lobbying for resources for the next generation of chips he had in mind—what ultimately became the phenomenally successful 8080. As Faggin recalled, "Intel management wanted to see how the market would respond to the MCS-4 and, later, to the MCS-8 introduction before committing more resources. I thought we were wasting time" (Faggin, 1992). Nevertheless, Faggin eventually received the go-ahead and work proceeded on the 8080.

What ultimately swayed Intel management to commit fully to further microprocessor development was the irresistible growth in Intel's market position as a result of the success of the 8080. The 8080 was, according to Faggin, "the one that opened the floodgates" (Manners, 1996). The chip was introduced in April 1974 at a price of $360. Intel recouped its development costs within the first five months of shipments. By 1974, the microprocessor business accounted for 30 percent of the Intel's sales (Aspray, 1997).

The preceding evidence suggests that, as an organization, Intel's commitment followed from rather than preceded the spectacular market success of the microprocessor. Although the 4004 and 8008 designs had demonstrated the technical advantages of the microprocessor concept (and revealed the market potential to Faggin), Intel management did not come around until this potential was realized with the 8080. As CTC's Vic Poor observed, "In reality, Intel was pulled kicking and screaming in the business [of making microprocessors]" (Gardner, 1996).

A second piece of evidence to consider is that while Intel took an important inventive step in creating the first microprocessor, the ideas and expertise contributed by the individuals at Intel were largely not unique to Intel. As mentioned before, the concept of a CPU on a chip was widely known, as was the design concept of a simple but flexible general-purpose computer. For instance, while Ted Hoff is widely credited with devising the architecture of the 4004, Hoff later acknowledged that he had been aware of similar ideas that engineers had been working on at other organizations including SRI, IBM, and RCA (Aspray, 1997). More concrete evidence can be seen in the advanced state of Intel's competition at the time it entered the market. As mentioned earlier, leading semiconductor companies in the 1960s such as Fairchild and Rockwell had been working to develop a CPU on a chip well before Intel's involvement (indeed, before Intel's founding). When the Intel 8080 was introduced in March 1974, it was introduced in the midst of competing products. Motorola introduced the 6800 six months after Intel's 8080.

Rockwell had introduced the PPS-4 in 1972 and the PPS-8 shortly after Intel's 8080. While the Rockwell PPS-8 was not a close match, Faggin acknowledged that, "In many ways, the 6800 was a better product" (Faggin, 1992). Faggin attributed the success of Intel's design to superior timing and market adoption.

A final piece of evidence to consider is that, if Intel's fortuitous experience resulted in the birth of any unique expertise or commitment to microprocessor development, this was lost when key members of the design team left Intel to found their own microprocessor company. While the 4004 and 8008 were wholly derived from projects with Busicom and CTC, the 8080 was Intel's first microprocessor intended from the beginning as an internal, proprietary design. It may be viewed as the first attempt within Intel to make use of the expertise acquired from its involvement with the two previous projects. As mentioned previously, this was not a broad effort but heavily advocated (and executed) by Faggin.

By the time that sales from the 8080 had made it obvious to Intel management that the microprocessor was a breakthrough, Faggin had become restless. He left Intel in October 1974, seven months after the introduction of the 8080, to found Zilog. Zilog was the first startup dedicated to the newly created microprocessor market. Describing his reasons for leaving, Faggin recalled, "Intel, in those days, was not really committed to microprocessors—they were committed to memories. My work wasn't appreciated" (Manners, 1996). As mentioned before, Intel management had reluctantly hired Faggin in 1970 to complete what it regarded as a peripheral project. Shortly after founding Zilog, Faggin hired Masatoshi Shima, who had previously moved from Busicom to Intel to work with Faggin on the 4004 project.

Evidence of the expertise that followed Faggin to Zilog may be seen in the creation of the Zilog Z80 microprocessor. Faggin envisioned the chip in December 1974 and completed it after Shima arrived to help with the design. The Z80 went on to become one of the longest-lived and best-selling microprocessors of all time. Introduced in 1976, the Z80 continued to sell by the tens of millions a year into the 1990s (Manners, 1996).

Adopting the role

As the preceding evidence suggests, Intel was in many ways pulled reluctantly into its position as a leader in the microprocessor market. After the takeoff of the 8080, the value of the new microprocessor market and, more importantly, Intel's superior position in it could no longer be ignored.

Subsequently, structural forces took over to induce Intel to play the role in which it found itself. Clearly, Intel did go on to commit substantial resources to the ongoing development of microprocessor technology, and its internal expertise undoubtedly became genuinely superior as a result. However, it is important to recognize how this was both enabled and compelled by the position in which Intel found itself. First, the windfall sales from the first generations of microprocessors provided Intel with the material resources to out-invest competitors in the development of subsequent generations. Second, Intel's sudden prominence among prospective adopters gave it a greater expected return and, hence, greater incentive to invest in the development of subsequent generations. Intel's superior prominence among prospective adopters is illustrated by IBM's choice of Intel's design over Zilog's for IBM's first personal computer (Manners, 1996). The effect of these structural forces can be seen in the fact that despite the technical expertise that followed Faggin to Zilog, Zilog was ultimately not able to match Intel in terms of research funding or product adoption.

In addition to increasing its tangible investment in microprocessor technology, Intel also ‘played the part’ by reinforcing an Intel-centric account of the invention of the microprocessor. Intel’s company histories subsequently celebrated Hoff as the ‘father of the microprocessor’ while downplaying the roles of others (Gardner, 1990). Although Hoff undoubtedly played an important role by making key architectural decisions at the start of the projects, most of the remaining work was completed by others. As mentioned earlier, engineers Busicom and CTC did a substantial amount of design work behind what became the Intel 4004 and 8008. Given that the projects began as custom contracts, design work done at the two client companies was shared with the engineers at Intel in hopes of pushing the projects along.

In the case of CTC, a working CPU was completed and put into production while waiting for Intel to make progress. The rights to the architecture were subsequently given to Intel, against the objections of CTC’s Austin Roche (Wood, 2008). In the case of the Busicom project, work stopped on the project after Hoff’s initial discussions with Busicom when the contract was accepted. Work did not resume until Shima returned to check on the progress and Faggin was hired to Intel to pick up slack on the delinquent contract. Both Faggin and Shima had been led to believe that the project was much farther along, and both were surprised at the work that remained to be done when they arrived (Aspray, 1997). According to Faggin, “Hoff was away on business and thought his job was finished” (Faggin, 1992). Years later, accounts suggested that “Faggin is still smarting today over what he regards as lack of credit for his work in microprocessors at

Intel" (Gardner, 1990). Others suggest that "Intel's powerful public relations machine has done little to disabuse the notion that people who remain close to the company (Hoff) deserve more credit than others who moved on to rival firms (Faggin)" (Aspray, 1997).

Discussion

This chapter contrasts the complex reality of invention with the much simpler understandings that firms actually use to characterize one another's roles in the industry. The historical facts of how a technology was pioneered become lost over time. In dividing credit for innovation, firms rarely dig into these details and instead rely on a simplified but nonetheless sensible picture of the world. I suggest that this picture consists of firms' shared understanding of one another's levels of commitment to the ongoing development of a technology. It substitutes for the more complex technical and legal relationships that link firms in the context of cumulative innovation. While highly contestable, such details are rarely contested. (I develop these ideas in more detail in chapter 4.)

The question then is: why not. Moving forward, it is not satisfactory to say simply that firms tend toward simple rather than complex accounts of reality. These accounts prescribe how returns to innovation ought to be divided. So naturally, they may be challenged, and firms are prone to creating self-serving accounts when money is at stake. For a simplified account of reality to be shared among firms and taken-for-granted as if it were a reality, it must be somehow robust to this explicit contestation even when it implies inequality.

The case of the microprocessor illustrates that while historical details are highly contestable, they are also easily suppressed. Historical details may be forgotten or recounted selectively by the dominant player in the industry. These biased accounts may go unchallenged because they nevertheless provide sensible guidance to firms for the ongoing task of dividing returns to innovation. Though Intel began with neither the vision nor technical expertise typically associated with the pioneer of a market, it was subsequently induced to genuinely behave *as if* it did.

Structural forces induce firms to behave according to the positions they occupy in the industry. As illustrated in the case, a firm can become locked in (even reluctantly) to ongoing commitment to the development of a technology. This suppresses the importance of historical details which might suggest alternative ways of dividing credit for innovation. The fact that Intel began with neither the will nor ability to pioneer the microprocessor market does not detract from the fact that over the years it made genuine contributions toward advancing microprocessor technology. When such commitments are clear to all firms involved in dividing returns to innovation, historical details simply do not arise as points of contention. In contrast, when firms do get into conflicts about how to divide returns to innovation, it is often because there is ambiguity about firms' commitments to contribute to advancing a technology. (I developed this idea in chapter 2.)

An important issue not addressed in this chapter was the role of patents in the invention of the microprocessor. Patents are meant in principle to document precisely the kinds of historical details presented in this chapter so as to minimize subsequent conflict over credit. In principle, formal patents, not informal understandings among firms, should indicate how returns to innovation should be divided. In practice, however, patents do a poor job of this. As a simple illustration, neither Ted Hoff, Federico Faggin, Austin Roche of CTC, nor any of the individuals and organizations described in the preceding story holds the patent for the first microprocessor. Instead, years later, the contest for first patent on the microprocessor ended up happening between Texas Instruments and an unknown independent inventor named Gilbert Hyatt (Dunn, 1996).

Patents, rather than provide clarity, often simply make for additional layers of complexity that must be overcome in the process of dividing returns to innovation. When patent disputes go to litigation, lengthy discovery procedures aim to reveal precisely the kinds of complex historical details found in the preceding story. Issues such as the work done by various individuals, the sharing of knowledge among individuals, and the extent to which an idea was common knowledge in the industry are examined to determine how much credit is due to the various parties involved. In fact, many of the historical facts used in this chapter surfaced in the wake of Texas Instruments' dispute with Gilbert Hyatt over credit for microprocessor. As the facts suggest, the reality of invention had weak correspondence to the patent record.

Luckily, in practice, firms generally recognize that patent rights provide an unreasonable basis for dividing returns to innovation. As chapter 2 illustrated, dividing returns to innovation happens in a more orderly way when firms share such an understanding instead of adhering to formal interpretation of patent rights.

Chapter 4: Status ambiguity and conflict

Introduction

In the preceding chapters, I have suggested that the complex reality of invention contrasts sharply with the much simpler understandings that firms actually use to characterize one another's roles in the industry. The historical facts of how a technology was pioneered may become lost over time. In dividing credit for innovation, firms rarely dig into these details and instead rely on a simplified but nonetheless sensible picture of the world. I suggest that this picture consists of firms' shared understanding of the roles they play in the innovative activity of the industry. If firms agree about their own and one another's roles in the industry, then this substitutes for the more complex technical and legal details that actually link firms in the context of innovation. To make the case that this is a reasonable story of how firms divide returns to innovation in practice, I presented anecdotal evidence in chapters 2 and 3. In this chapter, I take a step towards more systematic evidence by building a more abstract theoretical version of my story and generating empirically testable predictions from it.

I have suggested that while highly contestable, the technical and legal details behind innovation are rarely contested. To merely assert that firms adhere to simpler, mutual understandings does not explain why. Since these understandings are informal and have bearing on dividing monetary returns to innovation, they are vulnerable to being challenged. Under what conditions do firms agree about one another's roles even when these leave them with unequal shares to returns from innovation?

I argue that structural forces can lock firms into positions of ongoing investment in a technology. Social consensus about which firms are technological leaders and, hence, which firms' technologies are worth adopting can become a self-fulfilling prophecy. Firms that attract more investment in follow-on and complementary innovation by others in the industry are more likely to see their innovations become truly valuable. Firms in these positions therefore have a greater incentive to invest in innovation because they enjoy the promise of better rewards. Moreover, because a firm's position is conferred by the group, the set of incentives associated with the position is both publicly-recognized and beyond the firm's direct control. As a result, a firm's position can become a credible signal (in the eyes of other firms) of its commitment to investing in innovation and, hence, a signal of the actual, unobserved innovative contributions behind its patents.

Since the central feature of these positions is that they indicate which firms ought to have priority over which other firms when it comes to dividing up returns to innovation, I conceptualize them in terms of positions in a status hierarchy (e.g. Podolny, 1993; Gould, 2002; 2003). Research has argued that status hierarchies operate as signaling mechanisms in market settings where true quality is difficult to assess (Podolny, 1993). I argue that the same is true in the context of innovation. Firms' actual innovative contributions are difficult to assess. As illustrated in the preceding chapters, firms must sift through a complex web of technical and historical details before they can begin to quantify their true innovative contributions. This makes for a highly contestable reality when firms try to determine how much they owe one another for the benefits they derive from one another's innovative contributions.

Status hierarchies can suppress contestation over these complex technical details by providing a simpler basis for differentiation. Status differences signal differences in the incentives that firms face when investing in innovation. When a firm is more widely-perceived as being a technological leader than its competitors, it has a structural advantage over these competitors when investing in innovation. This structural advantage can serve as a signal of its greater innovative contribution over its competitors even if the reality of their relative contributions is difficult to verify. When there is a clear status difference between two firms, the idea that one has contributed more to innovation than another appears to both sides as a taken for granted reality. As a result, initial terms that the firms offer to one another for dividing returns to innovation are more likely to be accepted, and further contestation over technical details behind their true contributions becomes less likely. On the other hand, when the status difference between two firms is ambiguous, the idea that one side has contributed more to innovation than the other appears to both sides as a tenuous, malleable reality. As a result, initial terms that firms offer to one another are less likely to be accepted, and further contestation over technical details becomes more likely as firms try to counter one another's claims.

This contestation has two important and observable consequences. First, it increases the likelihood that attempts to find a market solution (licensing) break down, forcing firms to turn to litigation. I treat litigation as an indicator of conflict. Second, it imposes a cost to operating in the market, which can deter firms from bringing products to market in the first place. This can be seen in a firm's product line size.

In the next sections, I use these two variables to develop empirically testable predictions. I argue that the status difference between two firms becomes clearer, less ambiguous as the magnitude of the difference becomes larger. In other words, firms of similar status face greater ambiguity than firms of different status. We would expect litigation to be more likely between firms of similar status than between firms of different status. We would also expect firms to maintain larger product lines when they are more differentiated in status from their competitors.

In addition, I argue that large status differences only reduce ambiguity to the extent that they are consistent. I identify two sources of inconsistency, i.e. internal contradictions in status hierarchies, which are both likely to be present in real-world status hierarchies and which we would theoretically expect to make status differences less meaningful. The first is the presence of multiple subpopulations of peers, not all of whom agree in their endorsements of every given firm. The second is the presence of endorsements from higher-status firms to lower-status firms. In the presence of these kinds of inconsistencies, two firms can come to different (self-serving) conclusions about who is of higher and lower status by looking at the underlying network of endorsements differently. This makes even large status differences more ambiguous and therefore contestable. In the presence of inconsistencies, large status differences do not reduce the likelihood of litigation and do not make firms more wiling to maintain large product lines. The following sections develop these predictions in more detail.

How do status hierarchies generate order?

In a market setting, an organization’s status reflects the extent to which its outputs are endorsed by other organizations as being high in quality (Podolny, 1993). Status is important in markets where true quality is difficult to assess because it operates as a signaling mechanism for organizations both to differentiate themselves and to discriminate in their dealings with one another (Podolny, 2001; Spence, 1974). Higher status organizations demand and receive better terms of exchange. Lower status organizations neither demand nor receive attractive terms of exchange.

That every organization knows its place makes for orderly interactions in the face of uncertainty. Organizations can avoid the uncomfortable situation of making or receiving an insulting offer. On the other hand, when status is more ambiguous, organizations are less able to differentiate themselves from one another. When forced to deal indiscriminately with one another, organizations become systematically prone to misjudging one another’s true quality and, therefore, one another’s preferred terms of exchange. Interactions between status-ambiguous organizations are consequently less orderly and prone to conflict.

Defining dispute

Conflict arises from the inability to settle disputes that arise in the course of doing business. Disputes are commonplace in settings where actors must divide material resources. In the course of a dispute, two actors make unverifiable claims about their relative quality and contest the obligations that these would imply. If they accept one another’s claims, then the dispute ends. If not, then the dispute escalates into conflict. By conflict, I mean that two actors assert their claims against one another using means outside of the original setting in which a dispute arose. Under this definition, resolving business disputes through litigation (a non-market mechanism) corresponds to conflict, whereas resolving business disputes through contracts (a market mechanism), corresponds to the absence of conflict.

In the context of settling a dispute, quality refers to a desirable property (or absence of a negative property) that influences what each party would be willing to pay to or accept from one another in a settlement. For instance, in Coase's (1960) example, quality may refer to the degree of care (or negligence) that a rancher takes to keep his herd from trampling his neighbor's farm. As in the exchange of a good or service, providing higher quality requires more costly investment, something which the providing party hopes will be reflected in the settlement of a dispute. And as in the exchange of a good or service, quality is often difficult to assess, making it difficult for parties to correctly discriminate in their settlement offers. The problem of settling a dispute therefore becomes analogous to the problem of two parties trying to discover one another's preferred terms of exchange. Just as misjudgment of bargaining thresholds can prevent transactions from happening, so it can also prevent the settlement of disputes. In the former case, the parties simply walk away. In the latter case, failure to find solutions within market mechanisms forces the parties to seek solutions outside of market mechanisms, such as going to court. I characterize the escalation of disputes to domains outside of the market as instances of conflict.

What distinguishes a dispute from other situations in which two parties must divide some pool of resources between them is that a dispute arises involuntarily and the disputing parties cannot simply walk away until they can reach an agreement about how the pool is to be divided. For instance, contrast the sale of an automobile with an automobile accident. In both cases, there are potential gains from trade. In the former situation, there is a bargaining surplus if dealers reveal the lowest price they would be willing to take and buyers reveal the highest price they would be willing to pay. In the latter situation, there is a bargaining surplus if both parties admit their relative faults and avoid costly legal inquiries to determine fault. In the former case, if bargaining breaks down, then no sale takes place. In the latter situation, if bargaining breaks down, then the dispute escalates until costly legal inquiries reveal relative fault and impose a net compensation.

In the event of a dispute, a shared prior understanding can help to limit the possibility that the opposing parties both make reasonable claims—i.e. consistent with reasonable beliefs about what the opposing party would be willing to accept—that nevertheless lead them to competing conclusions about who owes whom. The fact that quality is difficult to assess leaves room for both sides to make self-serving representations of quality. In the course of bargaining, both sides can make reasonable offers that turn out to be incompatible with one another. Whether bargaining fails or succeeds is influenced by the availability of a signaling mechanism through which each side can credibly differentiate its own quality and correctly discriminate among quality levels of the opposing side. A status hierarchy can help actors to avoid conflict by providing such a signaling mechanism.

A model of status in the context of disputes

Status hierarchies operate as signaling mechanisms in the sense that they generate predictable variation in quality across positions the hierarchy. Actors in high-status positions enjoy collective endorsement by their peers. They get the benefit of the doubt from audiences when they make high-quality claims about their output, and so they tend to receive disproportionately high recognition and reward for producing high-quality output. Actors in low-status positions, on the other hand, lack endorsement by their peers. Their claims about the quality of their output are discounted by audiences, and so they tend to receive little if any recognition or reward for producing high-quality output.

Over time, the expectations associated with a position tend to be self-fulfilling. Occupants of low-status positions become locked in a vicious cycle of low-quality output. Realizing their efforts will go unrewarded, they begin to produce at the level expected of their position. Similarly, occupants of high-status positions become locked in a virtuous cycle of high-quality output. Because they enjoy the benefit of the doubt, high-status actors receive ample opportunity to repeatedly demonstrate their quality, and the rewards make their continued efforts more than worthwhile. Essentially, by channeling recognition and reward toward higher-status positions and away from lower-status positions, a status hierarchy generates the conditions for actors to produce at the quality level expected of their positions. The consequence is predictable and stable variation in quality across positions in a status hierarchy.

In this way, a status hierarchy operates as a signaling mechanism (Podolny, 1993). A high-status actor's position in the status hierarchy allows it to convince audiences that it has a credible reason to invest in quality. This allows high-status actors to command higher prices and recoup costly investments in quality. A low-status actor's position in the status hierarchy prevents it from being able to credibly claim investment in quality. This prevents low-status actors from making costly investments that cannot be recouped.

From the perspective of an external audience evaluating actors' quality claims, status processes ensure that quality levels vary predictably across positions in the hierarchy and that these quality differentials remain stable over time. Stated differently, status processes force separation between producers of high and low-quality output in spite of quality claims being difficult to verify in advance (Spence, 1974).

A status hierarchy therefore performs an important function in the market in that it aligns the two necessary sides of a business transaction. It simultaneously restricts the range of quality at which actors are willing to produce on the one hand and the quality claims that audiences are willing to accept on the other.

These same mechanisms become important for averting conflict because they also imply deference rules between occupants of different positions (Gould, 2002) Just as status restricts the quality claims that audiences are willing to accept from occupants of a given position, so it also restricts the claims of superiority that occupants of different positions are willing to accept from one another. This has important implications for the conduct of disputes. By deference, I mean simply one actor's acknowledgement of another actor's superiority in terms of quality. By deference rule, I mean that the pattern of endorsements between occupants of a status hierarchy is stable; once in place, no single actor, whether in a high or low status position, would be willing to deviate from it (Gould, 2002). In this sense, the system of deference behaviors operates as a self-reinforcing social convention and, hence, takes on a rule-like quality.

What makes deference an interesting social phenomenon is that while, on average, higher status actors produce at a higher quality level than do lower status actors, the magnitude (and even possibly the direction) of the true quality difference between a particular high status actor and a particular low status actor in the context of a particular dispute will vary. Suppose that at every position in the status hierarchy, there is a distribution of quality levels at which occupants of that position are enabled and compelled to produce. While low-status actors are on average of lower quality than high-status actors, this average quality differential may not necessarily reflect the reality of each particular dispute.

Given this, the assumption that quality is difficult to assess is necessary for the concept of status in general and deference in particular to be theoretically meaningful. When quality is difficult to assess, each actor has room to represent its own quality level in self-serving ways. For the same reason, each actor can also reasonably suspect that others are representing their quality levels in self-serving ways. This contestability naturally makes it difficult for actors to reach mutually-agreeable terms of exchange. Deference means essentially that actors adhere to the rules governing their *positions* without taking the costly steps needed to assess the reality of the particular situation. For lower-status actors to acquiesce to the demands of higher-status actors represents a sensible (rather than technically accurate) rule of conduct.

Testable prediction: Status difference, status consistency, and conflict

In order to test the preceding model of how status hierarchies suppress conflict, I develop two predictions about conditions under which the mechanisms described above provide weak guidance for resolving disputes. When actors' positions in the status hierarchy provide weak signals for differentiating their quality levels from one another, they face status ambiguity. First, I consider how the difference in status between two actors influences the degree of ambiguity they face. I then consider how this effect varies with the degree of inconsistency in status difference.

Status difference

For a status hierarchy to provide a strong signal of quality, an actor's ability and incentive to produce high quality output must depend heavily on the level of endorsement it receives from its peers. This is because socially-imposed constraints are both commonly-recognized and beyond the direct control of an individual actor. Powerful structural forces overwhelm other factors that might influence an actor's ability and incentive to produce high quality output. In doing so, it provides a credible signal that there are few factors within the control of individual actors that they can manipulate secretly to profitably misrepresent their own quality.

Analogously, for a status hierarchy to provide strong guidance in a dispute, the two actors in a dispute must depend heavily on the endorsements of their peers in the hierarchy. The reason is that if their abilities or incentives to produce high quality output involve too many factors beyond the influence of their peers (but for which they have private knowledge and control), then their positions in the status hierarchy would only serve as noisy signals of their commitment to quality. In its strongest form, a status hierarchy suppresses all disputes. Arbitrarily small differences in social status generate sharp differences in the abilities and incentives of actors to produce high quality output (Podolny, 1993). Hence, arbitrarily small status differences between parties in a dispute are enough to signal differences in their obligations towards one another.

In practice, arbitrarily small differences in social status are not enough. Actors are generally influenced by at least some factors beyond the influence of their peers and for which they have private knowledge and control. Whether good or bad, these factors cannot be credibly signaled through status. For instance, a low-status actor might possess some hidden quality advantage, and this allows the actor to sustain a high level of quality despite lack of immediate recognition and reward. Similarly, a high-status actor might possess some hidden motive to permanently exit from the hierarchy and 'cash out' on its status before it does.

Therefore, the probability of disputes escalating will depend on the magnitude of the status difference. The logic is that if two disputing parties are sufficiently close in status, then reasonable beliefs about the nature of one another's hidden motives could lead both parties to competing claims about their relative quality and hence competing conclusions about who owes whom in a settlement. When status is a noisy signal of quality, deference rules provide much weaker guidance in disputes between actors of similar status. This is consistent with Gould's (2003) observation that serious conflict is often the result of apparently minor disagreements between parties of similar status.

Intuitively, deference rules governing interaction between nearby positions are more contestable than rules governing interaction between distant positions. Deference rules in a status hierarchy depend on contrast between the opportunities and constraints of different positions. This contrast is naturally weaker between nearby positions (Podolny, 1993). In practice, this means that minor points of disagreement in a dispute, instead of simply shifting the magnitude of one actor's obligation to another, will tip the direction of the obligation. Therefore, we would expect that disputes between actors of different status are less likely to escalate than disputes between actors of similar status.

Hypothesis 1: The larger the status difference between two parties to a dispute, the lower is the likelihood that the dispute escalates.

This prediction is broadly consistent with Podolny's (1994) argument that firms of different status receive asymmetric benefits from transacting with one another, and as a result, transactions are more likely to happen between firms of similar status than between firms of different status. For fundamentally the same reason, disputes between firms of different status are easier to resolve than disputes between firms of similar status. The fact that firms of very different status are not willing to transact with one another also makes disputes easy to resolve—it is clear which side benefits more than the other. The fact that firms of similar status are willing to transact with one another also makes disputes difficult to resolve—it is unclear which side benefits more than the other.

Attenuating effect of inconsistency in status difference

Status differences are effective in helping firms to resolve disputes to the extent that they impose clear differences in opportunities and constraints for producing high-quality output. High-status firms must not only have sharp incentive to invest in quality but also have a sharp disincentive to divest in quality. Similarly, low-status firms must have a sharp incentive to divest in quality and sharp disincentive to invest in quality. The fact that these contrasting incentives derive from the collective endorsements of a common population of peers makes the incentives both mutually-understood and beyond the direct manipulation of either firm. In this way, status differences provide meaningful signals of quality differences.

For this reason, the existence of subdivisions in a population weakens the overall power of status as a signaling mechanism and, therefore, as a mechanism for suppressing conflict. While in status theory, firms are often assumed to derive their status from a single population of peers, in practice, firms are likely to derive status from multiple subpopulations, not all of which agree with one another. The lack of agreement among subpopulations makes it possible, in principle, for firms to earn high endorsement in one subpopulation even though they receive low endorsement in another subpopulation. Firms in these positions face competing incentives that may cause them to behave differently from both clearly higher- and clearly lower-status firms. On the one hand, this provides a potential chance at upward mobility. By eroding the clear contrast in opportunities and constraints, these 'intermediate' positions provide a way for firms to break out the rigid discipline normally imposed by status hierarchies. On the other hand, firms in these positions are less able to credibly demonstrate their commitment to a high quality level for one subpopulation because they simultaneously face competing incentives from another subpopulation. The fact that firms can occupy these kinds of 'intermediate' positions means that status differences between firms may be inconsistent. The relative standings of two firms may vary and even reverse across subpopulations. Naturally, this makes status differences more contestable and less helpful in resolving disputes.

Being of higher or lower status, i.e. status difference, matters because of the sharp contrast in opportunities and constraints between actors that are more and less highly endorsed by their peers. Peer endorsements, in turn, are so powerful in shaping opportunity and constraint because peers have social influence over one another. Social influence generates a cascading effect of collective endorsement. In a variety of settings, social influence induces mimetic behavior among peers (Bothner, 2003; Greve, 1996, 1998, 2000; Wade, 1996). In the context of a status hierarchy, the cascading effect of peer endorsement contributes to the virtuous and vicious cycles that ensure higher quality output at higher status positions. Empirical research finds evidence of this mechanism; the total amount of endorsement an actor receives from its peers contributes multiplicatively to the attractiveness for any single peer to endorse the actor (Podolny & Stuart, 1995). Formal analysis of status hierarchies also confirms that social influence is the primary mechanism responsible for generating a stable system of differentiated positions (Gould, 2002).

Where endorsements come from several subpopulations with little or no social influence with one another, the population's capacity to collectively reward and discipline its members weakens. We would, therefore, expect the presence of inconsistencies to weaken any effect that status differences would have in helping firms to resolve dispute. Stated one way, even large status differences may have little effect in suppressing disputes if there is sufficient inconsistency. Stated another way, even small status differences may have a strong effect in suppressing disputes if there is sufficient consistency. We would therefore expect inconsistency to attenuate the effect of status difference in preventing the escalation of disputes.

Hypothesis 2: The less consistent the status difference between two parties to a dispute, the weaker is the negative effect of status difference on the likelihood that the dispute escalates.

Testable prediction: product proliferation

The risk of conflict is not merely a way to test a theory of status. Because conflict is costly, it has real consequences for firms (and ultimately for the status hierarchy as a whole). If a firm's position in the status hierarchy is systematically prone to conflict, then it is in a position that the firm cannot sustain. The diversity of positions and quality levels in a status hierarchy is stable over time only if each firm gets the right recognition and reward to produce at its present quality level (and only at its present quality level).

Conflict reflects instances when firms must fight for rather than automatically receive the recognition and reward they need to make their investments in quality worthwhile. In a general market setting, consider a firm that tries to raise the quality level of its products but still gets the same low offers from customers. In the context of dividing returns to innovation, consider a firm that raises its investment in innovation but is still asked to pay the same high licensing fees for its perceived reliance on competitors' innovations. In each case, the firm's position cannot be consistently reproduced over time, whether due to unrecovered investments in quality or the cost of fighting to recover these investments.

The viability of a firm's position can be seen in the expansion or contraction of firms' product lines. The number of different products offered by a firm reflects its coverage of the various pockets of consumer preferences that make up the market. This is particularly important in the context of innovation. A firm can spread the cost of innovation over a larger market if it introduces more product variants embodying the same innovations (Cohen & Klepper, 1992).

As firms in a market expand their product lines, however, they are also more likely to encounter one another as competitors. In order to manage this competitive overlap, they must be able to differentiate themselves on quality. Status hierarchies provide a mechanism for firms to accomplish this. As discussed in the previous section, a status hierarchy induces firms to restrict themselves to mutually-exclusive quality levels rather than try to invade nearby quality levels. Research has argued that this kind of self-enforced, mutually-recognized differentiation can substantially increase the carrying capacity of a market (Carroll and Hannan, 1999). Empirical research on ethnic conflict, for instance, suggests that job markets can sustain high levels of inequality, free of conflict, so long as positions in the occupational hierarchy can be kept clearly segregated by ethnicity (Olzak, 1992). Finally, research has argued that to the extent that economic actors mutually recognize the boundaries between positions, their behaviors serve to reinforce the opportunity structure implied by the system of differentiated positions (White, 1981).

Effect of status differentiation on product proliferation

As described in the preceding section, status hierarchies in reality differ from the status hierarchies of theory because they have limited granularity. In the ideal status hierarchy portrayed in theory, powerful structural forces generate sharp contrast in the opportunities and constraints between actors of different status so that even small status differences produce clear differences in quality. In other words, the status hierarchy is fine-grained. Extremely fine-grained status hierarchies begin to approach the continuum of differentiated positions portrayed in theoretical models (Podolny, 1993; White, 1981; Gould, 2002).

In reality, however, structural forces are rarely strong enough to wipe out the effect of other factors for which actors have private knowledge and control. Social status influences the average quality level at which actors are willing and able to produce. These other factors generate variation in actual quality within a given position. As a consequence, small differences in peer endorsements do not necessarily signal that two actors are producing at different quality levels. For differences in peer endorsement to generate clear differences in quality, the magnitude must be sufficiently large in magnitude. The resulting status hierarchy is course-grained.

A consequence is that in reality there is some overlap in quality levels at which firms in nearby positions are willing and able to produce. For reasons which they cannot signal through their status, firms may find it attractive to ‘invade’ nearby positions. For instance, a higher-status firm may try to leverage its price premiums to produce lower quality products. A lower-status firm may invest in quality improvements to try and produce higher quality products.

In the context of innovation, a firm that is not widely viewed as a technological leader may invest heavily in what it believes to be a breakthrough technology. A firm that is widely viewed as a technological leader may leverage returns from innovation to subsidize non-innovative competition, such as cost-reduction.

The risk of conflict arises because firms do not automatically recognize one another's costly investments in quality. Factors that are beyond the influence and knowledge of a firm's peers also cannot be signaled through status. Firms must instead fight to overcome doubt about their investments in quality. The reason is that in addition to the 'legitimate' strategies described above, there are also 'illegitimate' strategies that are outwardly indistinguishable. A higher-status firm, instead of producing at two different quality levels and using one to subsidize the other, may instead just drop quality across the board and continue to demand high-quality prices. Similarly, the lower-status firm, instead of investing in quality improvement, can simply mimic the higher-status firm and demand higher prices.

In the context of innovation, a firm that has not been widely viewed as a technological leader may insist that it is now investing in innovation and demand more favorable licensing royalties in return. Similarly, a firm that has been widely viewed as a technological leader may revert entirely to a non-innovative strategy but continue to demand favorable licensing terms. In either case, if a firm is suspected of investing little in innovation and simply padding its portfolio with worthless patents, then its claims are unlikely to carry much weight in licensing negotiations (Grindley & Teece, 1997). The fact that firms are uncertain about one another's actual investments in innovation makes it difficult for them to reach similar beliefs about what constitutes reasonable offers. Bargaining becomes less predictable and more prone to breaking down. This friction in licensing negotiations is either costly to overcome or leads to costly conflict. This is consistent with the idea that information asymmetry causes breakdowns in the licensing of ideas (Gans, Hsu, & Stern, 2008). All else equal, the expectation of costly conflict deters firms from bringing products to market.

Based on this reasoning, firms with large status differences from their neighbors above and below in the status hierarchy should increase their product line sizes more than do firms with small status differences from their neighbors. Essentially, firms in nearby positions cannot correctly discriminately in their dealings with one another as much as firms in distant positions can. Firms with less status difference from their neighbors incur an added cost to bringing products to market, whether due to unrecovered investments in quality or the cost of fighting to recover these investments.

Hypothesis 3: A firm's propensity to expand its product line increases with status differentiation.

The attenuating effect of inconsistency

In the preceding section, I discussed the potential for status inconsistency due to the existence of different subdivisions in the population of peers from which a firm could receive endorsements. Status inconsistency weakens the constraints imposed by firms' positions in the status hierarchy. As a result, it increases the magnitude of status differentiation necessary for firms to avoid conflict. The source of inconsistency discussed before is obviously most relevant in the context of firms competing across many different markets. However, the potential for status inconsistency also arises within the context of a single market.

Status inconsistency within a single market arises because it is possible for endorsements to flow from a higher-status actor to a lower-status actor at the same time that the higher-status actor expects deference from the lower-status actor. A higher-status actor may endorse the quality of particular lower-status actors but still be higher in status because it is more highly-endorsed in general. In this case, the higher-status actor would be ‘in the right’ to expect deference from these lower-status actors. However, the fact that the direction of endorsement runs opposite to the direction of expected deference makes the deference rule more contestable in these particular instances.

The reason for contestability is that for the deference rule to appear sensible to both sides, both must consider not just the endorsement between the two of them but also the endorsements they receive from others. In other words, they must consider the broader network of endorsements that make up the status hierarchy in order to reach the same conclusion about where they sit relative to each other. Up to this point, I have treated this as being unproblematic.

In practice, however, firms are boundedly-rational and may not necessarily consider the entire network of endorsements underlying the status hierarchy. A firm may instead stop at a localized picture of the network. This is especially tempting when it implies a favorable deference rule for the firm. Consider the case of a higher-status firm that endorses the quality of a lower-status firm. The high-status firm, based on the endorsement it receives from others, will expect deference from this particular lower-status firm in a dispute. The lower-status firm will reach the same conclusion only if it also considers the higher-status firm’s broader network of endorsements relative to its own. If, however, the lower-status firm considers only the endorsement it receives from the higher-status firm, it will conclude that it ought to receive rather than give deference. In practice, it is likely that a variety of factors influences how broadly or narrowly firms view one another's networks. The main point here is that certain networks are invariant to these factors while other networks are particularly sensitive. I characterize this as the degree of consistency in the status difference between the two firms.

The presence of inconsistency weakens the power of status difference to help resolve disputes. To see why, consider why it would not be sensible for firms to correct one another's beliefs about status and deference in a dispute. In principle, one firm could share its broader knowledge of the network with the other firm. But the resulting deference rule would lose its taken-for-granted quality. A deference rule can help in a dispute because it allows both sides to avoid the more costly scrutiny of technical and legal details. If both sides arrive at the negotiating table with mutually agreeable initial offers and expectations, then this reduces both sides' incentives to demand costly scrutiny of details. The implied deference rule has worked.

If, however, the two sides do not share the same deference expectations to begin with and negotiations begin on the wrong foot, then explicitly invoking a deference rule at this point may actually hurt negotiations. As discussed before, status differences accurately reflect quality differences between higher-status firms and lower-status firms in general but not necessarily in the case of every particular dispute. A firm that invokes status explicitly would essentially be proposing to settle the dispute using a standard that favors itself and has no grounding in the particular circumstances of the dispute. If anything, this is likely to generate more indignation and mistrust than simply debating technical and legal details. In other words, for a deference rule to help in a dispute, it must remain implicit. The status difference behind the rule must confront both firms as an immutable social reality and, therefore, be taken for granted rather than become another point to be argued.

If status differentiation encourages product line expansion by reducing the risk of conflict, then the presence of inconsistency weakens this effect. The presence of inconsistency makes deference rules contestable and, therefore, less useful substitutes for costly (but technically justifiable) bases for resolving disputes. The same consequences of weak status differentiation apply. Firms in nearby positions cannot deal as discriminately with one another and, as a result, incur an added cost to expanding their product lines, whether due to unrecovered investments in quality or the cost of fighting to recover these investments. For product line expansion to be worthwhile, a firm's position must not only be differentiated in status from nearby positions but differentiated in a consistent way.

Hypothesis 4: The positive effect of status differentiation on product line expansion is weaker in the presence of inconsistency.

Chapter 5: Patent litigation in the semiconductor industry

Introduction

In this chapter, I examine the consequences of status ambiguity in the context of patent litigation in the worldwide semiconductor industry. I find support for the argument that status ambiguity increases the likelihood that disputes escalate. I use an original data set in which I have hand-matched competitors sued by a patent owner for infringement to other competitors making the same product that were not sued. This yields a sample of dyadic observations in which some instances of infringement resulted in litigation while other instances did not. Analysis of this sample suggests that, conditional on infringement, two firms which are different in status are much less likely to litigate than firms which are similar in status. However, the effect of status difference is attenuated by the degree to which these status differences are consistent across different subpopulations in the industry that are responsible for conferring status. These results provide support for the argument that clear status differences reduce the likelihood that disputes escalate.

Empirical Strategy

The focus of this chapter is to examine empirically why some disputes escalate to litigation while others do not. To do this, I need a sample that allows me to compare litigated to non-litigated disputes. The semiconductor industry provides a useful context for the purposes of this study because features of the technology and products make it possible to construct such a sample. First, most products in the industry use patented technology, and so instances of infringement are common. Research has argued that prolific patenting in semiconductor technology has created a ‘patent thicket’ (Shapiro, 2001; Ziedonis, 2004). Dense, overlapping patent claims make it virtually impossible for products not to incorporate patented technology. Second, for every litigated instance of infringement, there are likely to be many more instances of infringement that were not litigated. Compared to other industries with similar rates of patenting, overall rates of patent litigation in the semiconductor industry are relatively low (Lanjouw & Schankerman, 2004). In addition, anecdotal evidence suggests that firms frequently enter into licensing agreements to settle suspected infringement, thereby reducing the frequency with which these incidences of infringement precipitate litigation (Cohen et al., 2000; Grindley & Teece, 1997).

While we know that non-litigated instances of infringement occur, we have much less systematic information about these than we do about instances that are litigated[1]. The challenge is to somehow ‘recover’ information about the population of non-litigated instances of infringement based on the subpopulation of litigated ones. To do this, I rely on the fact that for any given device made in the semiconductor industry, there are likely to be several firms offering substitutes with similar if not identical attributes—‘pin-for-pin replacements’. A major reason for this is that switching costs often make customers reluctant to adopt a particular device unless they can be assured of consistent supply at reasonable prices. To encourage adoption, firms commonly license competitors to enter the market as ‘official’ second sources for their products (Shepard, 1987) and also allow entry by ‘unofficial’ second sources—competitors making substitutes without prior

[1] A common challenge in studying legal disputes in general is that while information on the population of litigated disputes is often available from court records, there is usually no comparable information on the population of disputes that do not enter the legal system (Cooter & Rubinfeld, 1989).

permission. In the latter case, firms usually try to persuade unauthorized competitors to take out a license (Tilton, 1971).

This generates two mutually exclusive conditions that become useful for testing the predictions in this study. Either two firms make the same product and litigation occurs, or two firms make the same product and no litigation occurs. Stated in another way, for every competitor sued for making an infringing product, there is likely to be one or more other competitors making the same product that were not sued. These competitors commit the same 'offense', so to speak, as the ones who were sued—use patented technology and compete directly with the patent owner for market returns accruing to that technology. However, the division of market returns occurs in an arguably more orderly fashion in instances when litigation does not occur than in instances when litigation occurs. Examining why litigation occurs in some instances but not in others provides a way to test the main predictions in this study.

Following this logic, I construct a sample of dyad-level patent 'disputes'. On one side of the dyad are plaintiffs in patent infringement suits. On the other side of the dyad are direct competitors of these plaintiffs that I have identified as producing an infringing product. This group of competitors consists of the set of actual defendants in each suit along with a matched control set of *potential* defendants—firms that produced the infringing product involved a given suit but were not actually sued. Hence, each case-control group consists of two or more dyad-level observations, at least one of which is an instance of infringement that resulted in litigation and at least one of which is an instance of infringement that did not. Using this data structure, I test my predictions about the likelihood that disputes escalate by comparing attributes of dyads in which litigation occurred to attributes of dyads in which litigation did not.

This data structure has several advantages for testing the predictions in this study. First, the sample as a whole is constructed based on litigated cases. This allows the results of the analysis to be interpreted as being conditional on the fact that underlying patents and products are sufficiently valuable to merit enforcement. This is important in the context of the semiconductor industry because many semiconductor patents are thought to have little or no economic value (Moore, 2005). Second, the results are also conditioned on instances of infringement against patent owners that have a propensity to enforce their patents. This is important in the semiconductor industry because there is likely to be high variation in firms' propensities to enforce their patents (Lanjouw & Schankerman, 2004). Finally, the data structure allows me to compare litigated and non-litigated disputes within the same case. This should, in principle, allow me to 'condition out' much of the between-case differences in patents, products, and plaintiffs, so that predictions can be tested primarily based on within-case variation in dyad-level attributes.

Dependent variable: likelihood of litigation

Constructing my sample required several steps. I began by identifying the population of semiconductor firms that were candidates, either as plaintiffs or defendants, for litigation. The firms in my sample are taken from *IC Master*, an annually published handbook of integrated circuit devices. In the current paper, I construct my sample from the 237 firms that appear in the 1997 edition of *IC Master*[2]. The *IC Master* handbook is meant to be a comprehensive listing of products and so should contain all firms that manufacture at least one integrated circuit product for sale on the open market[3]. As a comparison, the Compustat database lists 166 firms under SIC 3674 (Semiconductors and Related Devices) between 1996 and 1997. Table 5.1 gives descriptive statistics of firms in this sample. The distribution covers a wide range of firm sizes. The smallest firms make one product, while the largest firms offer thousands[4]. The distribution is also highly skewed (see Figure 5.1). A quarter of the firms make fewer than 33 products, and half make fewer than 123. Their products belong to 90 different market segments, covering five major categories of integrated circuit devices—memory, linear, digital, interface, and microprocessors (see Table 5.2)[5].

Having identified the relevant population, I then identified patent infringement litigation among these firms. I relied on two sources: Derwent's *LitAlert* database and press reports from Factiva. The *LitAlert* database consists of patent suits reported to the US Patent and Trademark office (USPTO). This is the primary data source used in previous studies of patent litigation (e.g. Bessen & Meurer, 2005; Lanjouw &

---

[2] The current sample of litigation by firms in the 1997 *IC Master* edition is a preliminary one. The full sample will consist of litigation involving firms in the industry from 1984 through 2001.

[3] This excludes captive manufacturers, manufacturers of discrete devices only, and foundries. Alternative samples of semiconductor firms based on industry classifications such as SIC and NAICS will vary in some of these respects. For instance, International Rectifier, a manufacturer of discrete devices, and Taiwan Semiconductor, a foundry, are classified as semiconductor manufacturers according to their primary SIC (3674), but neither appears in my sample. Additionally, my sample includes electronics firms like Hitachi and Motorola that are not primarily classified as semiconductor manufacturers but are major players in the semiconductor industry.

[4] The three largest firms are National (6,726 products), Texas Instruments (6,060 products), and Motorola (5,058 products).

[5] I do not yet have data on firms making programmable logic devices, though this is an important segment of the semiconductor industry, especially from the 1990s onwards.

Schankerman, 2001, 2004). The *LitAlert* database is incomplete, however, since courts where patent suits are filed may neglect to report all filings to the USPTO. Lanjouw and Schankerman (2004) matched cases in *LitAlert* to the comprehensive Federal Judicial Center (FJC) Database of all civil suits in the US and found underreporting of about 15 percent[6]. Hence, I also relied on industry and general business press reports in the *Factiva* database. Using these two sources, I identified 147 patent infringement suits in which both the plaintiff and defendant are firms in my sample[7]. Consistent with Lanjouw and Schankerman's (2004) figures, 25 of these suits, about 13 percent, are not reported in *LitAlert*.

Because the sample was constructed based on all firms that make integrated circuits rather than only firms with a primary SIC in semiconductor manufacturing, it includes diversified electronics firms such as Motorola and Philips. Suits between these firms may involve downstream products beyond the scope of the market for integrated circuits. I searched press reports on the 147 suits and identified 28 which covered products besides integrated circuits. For instance, one suit covered pagers made by Motorola. Fifty-seven of the suits had no press coverage at all, so I was not able to

---

[6] The FJC database consists of court filings stripped of identifying information about the parties involved. Lanjouw and Schankerman (2004) were able to match *LitAlert* entries to FJC entries by docket number. However, it is generally not possible to recover any additional information from suits that appear only in the FJC database. Hence, the FJC database is generally unsuitable for studies involving firm or dyad-level attributes.

[7] The records in *LitAlert* are for all events related to patent litigation, which are not necessarily distinct lawsuits. There are occasionally duplicate records for a given lawsuit for different motions, different patents, different defendants, and differently coded docket numbers for the same suit. I excluded records involving the same plaintiff and defendant and on the same date as an earlier filing. In addition, the database does not distinguish between suits over patent infringement and suits over patent validity. I excluded records in which the defendant and assignee for the litigated patent are the same, which implies that the suit was filed to invalidate a patent. Validity suits are almost always filed as a response to an infringement suit, so including these as additional observations carries the risk of artificially strengthening the effects of variables that predict the onset of infringement litigation.

establish what product these suits covered. After excluding these, I was left with 62 suits that I could establish as covering integrated circuit devices.

The sample of 62 suits is confined to suits filed between 1984 and 2000. The reason is that in order to construct my case-control sample and to compute the variables needed to test my predictions requires complete data on the product lines of all competitors in the industry. I have been able to collect these data for the years 1984, 1988, 1992, and 1997 from the corresponding editions of *IC Master*[8]. If I treat the product market data as being valid for four years, then this allows me to test predictions involving suits occurring between 1984 and 2000.

Case-control sampling and estimation

In order to construct a matched control sample of non-litigated disputes, I had to identify the particular product or products in a defendant's product line that were responsible for infringement and then use this information to identify other competitors that made these same products. To do this, I relied on accounts of new product releases and patent suits in industry journals such as *Electronic Engineering Times* and *Electronic Buyers' News*. I used product data sheets to match descriptions given in press reports to products listed in the *IC Master* handbook. In some cases, press reports give the name of the product family, for instance, Analog Devices' AD-800 series of operational amplifiers. In other cases, only a functional description was available. In a suit against Advanced Micro Devices, for instance, the allegedly infringing products were AMD's

---

[8] I was not able to obtain the 1996 edition of *IC Master*, so there is a four rather than three year gap between the last and second to last years of data.

error detection and correction devices. I found three devices listed in *IC Master* that belong to the same family of AMD error detection and correction devices. In some cases, suits cover broad classes of related products. In a suit between Cypress and Quality Semiconductor, for instance, Cypress alleged that Quality's entire line of FCT logic circuits infringed a Cypress patent on FCT technology. This covered several subcategories of CMOS digital logic circuits, ranging from arithmetic circuits to counters. Similarly, suits involving memory products usually cover entire classes of memory devices, for instance, DRAM. In these cases, I view the breadth of the suit as reflecting the true breadth of the patents given the commodity nature of products such as memory. In contrast, I excluded those suits that had only broad product descriptions due to sparse press coverage but were not actually commodity products.

Using this approach, I was able to collect reliable product information for 51 lawsuit filings. For each of these suit filings, I identified all competitors that the made the infringing products in question, both those that were sued by the patent owner and those that were not. I created a dyad-level observation for each of these competitors, with the patent owner on one side and the competitor making the infringing product on the other. This yielded a sample consisting of 814 dyad-level observations representing instances of plausible infringement[9]. Of these 814 observations, 65 resulted in litigation, and 749 did not. The reason that there are 65 instances of litigation for 51 lawsuit filings is that in

---

[9] Strictly speaking, I cannot be absolutely certain using data at this level that all products performing the same function, even those that are pin-for-pin replacements, do indeed infringe on the same patents. To truly establish infringement would require much closer scrutiny of the products, for instance, comparing the circuit layout between the patent owner's and competitor's devices. This is, in fact, what is often done in litigation. However, given that this is costly for patent owners to do for all competitors, it seems reasonable for the purposes of this study to treat competitors making substitute products as plausible candidates for infringement litigation.

several lawsuits, more than one competitor is named as a defendant. I treat each separate defendant as a separate litigation event. This is necessary because the two explanatory variables of interest, status difference and inconsistency, vary across defendants within lawsuits. The alternative is to aggregate variables across defendants within the same lawsuit, for instance, by taking averages. However, in the context of a case-control sample, this approach discards potentially meaningful information that could provide contrary evidence against my predictions. For instance, if a predicted risk factor is extremely high for one case but low for all other cases and controls, then averaging across the cases could generate misleading support for the effect of the risk factor. Treating each defendant as a separate litigation event makes for an arguably more discriminating test of the prediction. Table 5.3 shows the distribution of cases and controls across the five categories of products represented in these suits. The distribution of cases and controls suggests that there may be variation across categories in the average probability of litigation. The lowest implied probability of litigation is in the memory category. For every competitor sued for making a given product, there are on average 25 competitors making the same product that were not sued. The highest implied probability of litigation is in the digital category. For every competitor sued for making a give product, there are on average 4 competitors making the same product that were not sued. For the entire sample, there are on average 12 controls (no litigation) for every case (litigation).

The structure of the data presents particular statistical issues that must be taken into account in the analysis. Most importantly, the data represent a highly non-random sample from the population of all possible competitor dyads, constructed based on the criteria outlined above. It does not contain the full population of patent owners that could have filed infringement suits, the population of suits actually filed, or the population of instances of infringement that could have resulted in litigation. Instead, it consists of a select sample of cases for which reliable controls could be constructed.

To take this into account, I use conditional logistic regression to estimate the probability of litigation between two firms. Conditional, or stratified, logistic regression is often used in biostatistics research to analyze matched case-control data (Collett, 2003). It allows matched observations to be grouped and computes likelihood relative to each group. Hence, the effects of variables of interest are estimated by comparing cases to their controls, i.e. using within-group variation only. Computationally, it is equivalent to fixed-effects logit regression in econometrics. In my analysis, I group observations by suit. This should, in principle, 'condition out' between-suit differences in patents, products, and plaintiffs, so that my predictions are tested based on within-suit variation in dyad-level attributes, i.e. attributes of the relationship between the patent owner and its competitors. The model is specified as follows:

$$\ln\left[\frac{\Pr(Y_{ij}=1 \mid X_{ij})}{1-\Pr(Y_{ij}=1 \mid X_{ij})}\right] = \alpha_j + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{1ij} x_{2ij} + Z'\gamma_{ij} + \varepsilon_{ij}$$

where Y_{ij} is a binary variable that takes the value of "1" if the litigation occurred in dyad *i* and a value of "0" otherwise. The subscript *j* indexes the suits in the sample on which the observations are grouped. The vector *Z* represents right-hand side variables that enter as controls. The independent variable x_{1ij} represents status difference between two firms in a dyad, and x_{2ij} represents inconsistency in status difference. *Hypothesis 1* predicts that $\beta_1 <0$. The larger the status difference between two firms, the less likely they are to litigate. *Hypothesis 2* predicts that $\beta_3 > 0$. The greater the inconsistency in status difference between two firms, the weaker is the negative effect of status difference on the likelihood of litigation.

Independent variables

Status difference. Status represents the level of endorsement that firm receives from its peers in an industry. Since I am primarily concerned with status as a signal of the quality of a firm's innovations, I follow previous work in measuring status based on the number of patent citations that a firm receives from its peers (Podolny et al., 1996; Stuart, 1998). Since status is meant to convey relative standing within a population of peers, I compute a firm's status as its share of all citations sent by peers in the industry in the preceding four years[10]. I define the population of peers in the industry as the set of firms that produce at least one integrated circuit device in a given year. As mentioned before, I was able to collect product data for all firms that operated in the years 1984, 1988, 1992, and 1997[11]. I matched the names of these firms to assignees in the NBER patent data set to identify citations between peers in the industry. I compute status difference between two firms as the absolute value of the difference between their status scores. The estimated coefficient for status difference provides a test for *Hypothesis 1*, which predicts a negative sign.

[10] I update the status measure in the same four-year windows as the years of product data I collected. This becomes important for computing the inconsistency measure, which is based on both the patent and product data.

[11] Firms that entered and exited in between these years do not appear in my data and, hence, do not contribute to the status measure. These are likely to be extremely small firms that send few if any patent citations. Omitting patent citations by these firms should have a small impact on the status scores of firms in the sample.

While I use patent citations to measure endorsement, patent citations are also commonly used as empirical measures of the actual quality of a firm's innovations. The assumption behind using patent citations as measures of quality is that other firms in the industry are in a position to judge the quality the focal firm's innovations. Firms cite a patent because they judge the underlying innovation to be high in quality. They judge the innovation to be high in quality presumably because it is actually high in quality. Podolny and Stuart (1995) and Podolny, Stuart, and Hannan (1996) suggest that, in practice, the technical merits of an innovation are difficult to assess, and so citations are more accurately interpreted as indicators of firms' judgments of an innovation rather than indicators of objective, technical information they possess about the innovation. In other words, they justify the measure by assuming that true quality is unobserved by the actors involved.

A potential problem with using patent citations as measures of status is that we may not be able to empirically distinguish whether firms are acting based on their objective knowledge of true quality or whether they are acting based on signals of quality provided by peer endorsements. I suggest that the problem lies primarily with the choice of dependent variable. Many studies of status are concerned with estimating the relationship between *level* of status and various material advantages. This is problematic not for lack of a good independent measure of status but because theory would predict that quality does not vary independently of status. In other words, the identification strategy of examining the relationship between level of status and material advantages inherently cannot distinguish between the effects of status and the effects of quality differences that are observable to actors but not to researchers.

In my design, the problem is different, primarily because my independent variables, dependent variable, and hence my identification strategy are different. My design focuses on the signaling effect of status differences rather than the material advantages of status *level*. My dependent variable of litigation, which I have argued as a measure of bargaining failure, should vary with the degree of information asymmetry but not with the degree of quality difference *observable* to the disputing parties. The intuition is that under perfect information, the settlement of a dispute where one party is only slightly more at fault than the other is not more difficult than the settlement of a dispute where one party is much more at fault than the other. Hence, if my measure of status difference is picking up quality difference observable to the parties involved, it should not have any effect on likelihood of bargaining failure.

In addition, discussion of the licensing process in previous chapters suggests that there is substantial information asymmetry in licensing negotiations when it comes to firms' contributions to innovation. One reason is that patents do not provide a perfect window on a firm's R&D activities. I have made this argument at various points in the preceding chapters. It is widely acknowledged that the content of patents does not always map well to the reality of R&D investments and R&D outcomes. Moreover, there is sharp variation. Some patents reflect true innovative contribution while others reflect legalistic patent harvesting strategies. The implication is that in order to establish the true R&D activity reflected in firms' patents requires substantial and costly investigation. I have illustrated this in Chapter 3. This costly investigation is conducted after a dispute goes to litigation. Licensing allows firms to avoid this costly verification but requires them to take each other's claims on faith.

Perhaps a more important reason is that even if citations measure past quality which is observable at the time of negotiations, licensing agreements are forward-looking. The unobservable 'quality' is a firm's future investments in innovation. This varies independently of a firm's present patent portfolio and depends on the firm's own choices, which are difficult to monitor and enforce. Access to a firm's patent portfolio is only as good as the R&D behind it. It is possible for a cross-licensing partner to 'overpay' if it agrees to a pay a high royalty rate but the other partner does no further R&D. Technology obsolesces much faster than patents expire. A firm is more likely to agree to pay a high royalty rate if it believes the other party is likely to make valuable ongoing R&D contributions. It will be reluctant to accept a high royalty rate if it believes the other party is trying cheat--lock in a high royalty now and then cut back on R&D investment in the future. Chapter 2 provided anecdotal evidence that this is a real concern to firms in the semiconductor industry. Cohen et al. (2000) have argued that in semiconductor industry, because knowledge flows between competitors are prevalent and difficult to trace, licensing agreements are used to divide broad market returns accruing to broad pools of knowledge rather than specific sales associated with specific patents.

The bottom line is that for a firm to maintain high-quality output into the future is not effortless but requires ongoing investment. This is why, even with information about a firm's past quality, licensing partners require credible signals of ongoing future commitment. In light of these information asymmetries, backward-looking measures based on patent citations should only have an impact on likelihood of bargaining failure if they provide credible signals to the negotiating parties.

Status difference inconsistency. The idea behind measuring inconsistency in the status difference between two firms is that the peers in the industry responsible for conferring status may be subdivided into distinct subpopulations that do not necessarily agree with one another. In other words, a given firm may have high status to one subpopulation but low status to another. As a consequence, the relative status of two firms might vary dramatically or even reverse across different subpopulations. *Hypothesis 2* suggests that the degree of inconsistency of this kind will attenuate the negative effect of status difference on the likelihood of litigation.

To compute a measure of status inconsistency, I need to first identify cohesive groupings of firms into subpopulations and then determine each firm's status among each of these subpopulations. The criteria I use for identifying subpopulations are that firms within the same subpopulation should be maximally similar to one in terms of the market segments in which they compete and maximally different from firms in different subpopulations. The logic is that firms competing in similar market segments have greater social influence with one another when it comes to evaluating technologies than do firms in different market segments. This logic is broadly consistent with well-established findings that firms are most susceptible to influence by reference groups consisting of similar peers (Greve, 1996, 1998, 2000). It is also consistent with empirical findings that firms in structurally-equivalent market positions influence one another's decisions to adopt innovations; Bothner's (2003) findings in the diffusion of microprocessors are particularly relevant for this study.

The semiconductor industry is highly segmented, and firms occupy a diverse array of market positions across these segments. Table 5.2 shows the three largest segments in each of the major categories and the three largest producers in each segment. This table reveals two interesting features of the market positions of semiconductor firms. First, few firms appear in this table more than once. While the most diversified firms such as Texas Instruments and National Semiconductor, unsurprisingly, have strong positions in several segments, in general, the biggest players in one segment tend not to be the biggest players in another segment. Second, the biggest players in the industry are not equally specialized in the segments in which they dominate. At one extreme, ESC Electronics specializes entirely in delay lines and is the largest producer in this segment. At the other extreme, National is the largest producer of voltage regulators, but this segment amounts to only 5 percent of its total product line. There is also variation within segments. In 8-bit processors, Philips is the largest producer, but the segment amounts to only 8 percent of its total product line. In contrast, Zilog, the second largest producer of 8-bit processors, has 93 percent of its products in this segment. The implication is that the competitors Zilog encounters most frequently in the market are unlikely to be the same as the competitors Philips encounters most frequently.

To identify subpopulations of semiconductor firms on the basis of market position, I use the *k*-means clustering procedure. The *k*-means clustering procedure groups items in such a way that items within the same cluster are similar to one another in terms of some attribute vector while being different from items in other clusters. Specifically, it assigns items to clusters so as to minimize angular separation between each item's attribute vector and the vector of attribute means for all other items in the cluster. To group firms by their market positions, I create a vector for each firm consisting of the proportion of the firm's product line devoted to each market segment in the industry. The angular separation between two firms, i and j at time t is given by:

$$c_{ijt} = \frac{S_i' S_j}{\sqrt{S_i' S_i}\sqrt{S_j' S_j}} \qquad j \neq i$$

where S_i is a vector whose typical element s_{ik} is the proportion of firm i's products in segment k. The measure is similar to Burt and Talmud's (1993) measure of structural equivalence in market niche. Applied in this way, the *k*-means procedure should group firms so members of the same group have high market overlap with one another, have overlap with one another's competitors, and their competitors have overlap with one another's competitors, and so on. If similar market position reflects social influence, then this cluster procedure should identify subpopulations of firms such that social influence is strong within subpopulations and weak between subpopulations.

Implementing the *k*-means procedure requires me to specify an initial grouping of firms. The procedure then iteratively reassigns firms to reduce centroid distance until no further reassignments can be made. The choice of starting conditions is necessarily arbitrary. For the sake of reproducibility of results, I begin by grouping firms by product line breadth, measured as simply the number of market segments in which a firm competes[12]. In addition to initial grouping, the procedure requires me to choose the number of clusters. This choice, also, is necessarily arbitrary. I experimented with between 5 and 40 clusters. The amount of variation in firm product lines explained by cluster membership (based on Wilk's lambda) is highest at 25. In other words, choosing

---

[12] I experimented with other starting conditions, such as grouping firms by dispersion of their product lines across segments (Herfindahl) and grouping firms by size (number of products) with similar results.

25 clusters yielded groupings that were maximally distinct from one another and homogeneous within[13]. Table 5.4 gives a description of typical firms in each cluster. For simplicity, I have characterized each cluster in terms of the market segment to which members tend to devote most of their products. In reality, a cluster is characterized the position of its center in the multidimensional space of market segments. Figure 5.2 gives a spatial representation, scaled to two dimensions, of how clusters are positioned relative to one another in market space.

After identifying the different subpopulations in the industry, I compute each firm's status in each subpopulation. A firm's status in a subpopulation is simply its share of all citations sent by members of that subpopulation[14]. To compute a measure of inconsistency in the status difference between two firms, I identify which of the two has higher status in each subpopulation where both receive at least one citation. I then multiply the proportion of subpopulations in which one firm has higher status by the proportion of subpopulations in which the other has higher status. Figure 5.3 illustrates how this measure is computed in an example with Motorola and Hitachi. Motorola has higher status in 9 of the 17 subpopulations from which both firms receive citations. Hitachi has higher status in the remaining 8. This measure gives an inconsistency score of 0.249. Obviously, the inconsistency measure takes on its maximum value (0.25) when

[13] The 'best-fitting' grouping of firms is not the same as the grouping of firms that yields the most significant hypothesis test results. In other words, I can generate stronger support for my predictions by choosing a different number of clusters or perturbing the starting conditions. However, I consider my current results the most sensible.

[14] That a firm could receive patent citations originating from different subpopulations as defined by market position suggests that technologies do not map in a one-to-one way to product markets. This is a commonly-observed property of semiconductor technology. Figure 4 shows the pattern of patent citations between subpopulations. If we interpret a patent citation as identifying a technological linkage between an innovation and the previous innovations on which it builds, then the citation pattern between subpopulations appears to confirm the many-to-many mapping from technologies to markets.

neither firm has higher status in more than half of the subpopulations they have in common. When two firms do not receive citations from at least two common subpopulations, then the inconsistency measure takes on a value of zero.

Controls variables

To study the escalation of disputes using litigation as a dependent variable, I have to be able to interpret litigation events as bargaining failures. This requires that I estimate the effects of my main variables of interest relative to a meaningful baseline reflecting factors that may systematically influence the likelihood of litigation for reasons besides bargaining failure. To construct this baseline, I include several control variables.

First, I include a dummy variable for observations appearing in the years of 1992 through 2000, the second half of the time period covered in my sample. This is meant to capture any unobserved changes over time in the underlying propensity for firms to litigate.

Second, I include a count of the number of infringing products made by a given competitor. Many of the suits in my sample covered more than one product. In constructing the case-control sample, I included competitors that made at least one infringing product. It is reasonable to expect that competitors making many infringing products are more likely to be targets of litigation than competitors making few. This is important to control for since the extent of infringement is likely to influence the incentive for the patent owner to sue a given competitor.

Third, I include a count of the number of products made by the competitor that directly compete with products made by the patent owner. I define directly competing products as integrated circuits performing the identical function. In most cases in my data, these are pin-for-pin replacements, i.e. perfect substitutes. It is important to control for the extent of direct product competition, since this is likely to influence the incentive for the patent owner to sue a given competitor.

Fourth, I include two controls for size difference. I include a control for difference in the number of products made by the patent owner and competitor, and I include a control for difference in the number of patents received by the patent owner and competitor in the previous four years. A large size difference between the patent owner and competitor may generate asymmetric incentives for settlement. For instance, recent research suggests that patent owners may use litigation as a deliberate signal of 'toughness' when it comes to enforcing its patents (Agarwal, Ganco, & Ziedonis, 2008). If this is a case, then a lawsuit that is unprofitable in the short run, i.e. one in which the legal fees exceed the expected settlement, may in fact serve the purpose of providing a credible signal of toughness and pay off in the long run. Hence, it should not be interpreted as a bargaining failure. If two firms differ sharply in size, then the value of the signal to the larger firm (which can spread the benefits of the signal over more patents or products) may exceed the ability of the smaller firm to offer a settlement to prevent litigation.

Fifth, I include a dummy variable indicating that the patent owner and competitor belong to the same subpopulation in the industry. Recalling that subpopulations are identified based on groups of firms in similar market positions, this variable controls for any effects that similar market position might have on the baseline probability of litigation between two firms.

Finally, I include a control for the relative status of the patent owner compared to a given competitor. It is computed as the patent owner's status minus the competitor's status. Because it is directional, this control is different from the main variable of interest, status difference, which is an absolute value. Research on predatory behavior suggests that high-status entrants are less likely to be targets of predatory behavior by incumbents than low-status entrants (Podolny & Scott-Morton, 1999). The logic is that status reflects an entrant's propensity to behave cooperatively in a cartel. If this is the case, then we might expect high-status patent owners to be more likely to sue a low-status competitor than a high-status competitor.

Table 5 gives descriptive statistics of the dependent, independent, and control variables used in my analysis. Note that for all variables the require product data to compute, I am limited to the years 1984, 1988, 1992, and 1997. This means that for patents suits occurring in the intervening years, the 'true' values of these variables may differ from the actual values on the date of the suit. In my analysis, I used the most recent year of data before the date of a suit[15].

Results

[15] In a few cases, a firm (either case or control) does not appear in the most recent year of data before the suit. For instance, if a suit occurs in 1991, some of the producers of the infringing product may have been founded after 1988. In these cases, I use the earliest year of data in which a firm appears (but no earlier than one year before the suit). For instance, for a firm founded after 1988 and involved in a suit occurring in 1991, I use data from 1992. However, if a suit occurred any more than one year before the earliest year of complete product data for the firms involved, I exclude the suit from the analysis.

Table 6 reports results from conditional logit estimates of the likelihood of litigation between patent owners and competitors in the case-control sample. *Model 1* is the baseline model with control variables only. The results imply that several factors have strong effects on the baseline probability of litigation between firms. The probability of litigation increases with the number of infringing products made by the competitor. The probability of litigation is also higher when the patent owner and competitor belong to the same subpopulation than when they belong to different subpopulations.

Model 2 adds the independent variable, status difference, as a test of *Hypothesis 1*. As predicted, status difference has a negative effect on the likelihood of litigation between the patent owner and competitor. However, the effect is significant only at the $p<0.10$ level. This provides weak support for *Hypothesis 1*. *Models 3* and *4* add the independent variables for inconsistency in status difference and the interaction between status difference and inconsistency. In *Model 4*, the interaction between status difference and inconsistency has a positive and significant coefficient ($p<0.01$). This is consistent with *Hypothesis 2*. The positive and significant sign of the interaction term suggests that the effect of status difference on the likelihood of litigation between two firms becomes weaker to the extent that the status difference is not consistent across all subpopulations in the industry. In addition, accounting for this attenuating effect of inconsistency causes the main effect of status difference to become significant ($p<0.05$). The results of *Model 4* provide support for *Hypothesis 2* and stronger support for *Hypothesis 1*.

In nonlinear models such as logistic regression, coefficient estimates and significance tests for interaction effects cannot be interpreted in the same way as in linear models (Ai & Norton, 2003). To appropriately represent and assess the significance of the interaction effect in a logistic regression model, I follow Zelner's (2009) suggestion to look at the change in predicted probability of an event due to a discrete change in the first variable in the interaction, computed at high and low levels of the second variable in the interaction. I use the simulation-based approach suggested by King, Tomz, and Wittenberg (2000) and implemented in Zelner (2009) to compute appropriate confidence intervals for the predicted probabilities. Because existing software for implementing these methods do not yet accommodate conditional logistic regression, I estimate my full model using logistic regression with unconditional fixed-effects for each case-control group. The results are reported in column 5 of Table 5.6. Compared to the conditional logit model, the logit model with unconditional fixed-effects produces very similar results both in terms of estimated coefficients and standard errors.

I use estimates from the logit model with unconditional fixed-effects to simulate predicted probabilities and confidence intervals. The effects are represented in Figures 5.5 and 5.6. Figure 5.5 plots the relationship between predicted probability of litigation and status difference. At low levels of status inconsistency, status difference has the predicted effect of reducing the probability of litigation. At high levels of inconsistency, status difference does not reduce the probability of litigation and in fact slightly increases the probability of litigation. Figure 5.6 plots the change in the effect of status difference between high and low levels of inconsistency. For most of the observed range of status difference in this sample (status difference > 0.04), the effect is significantly differently between high and low levels of inconsistency. These results provide support for *Hypotheses 1 and 2*.

Robustness checks

A common statistical issue in the analysis of dyad-level data is non-independence of observations. In my data, non-independence arises from the fact that firms can appear in multiple dyads in the sample, and unobserved attributes of these firms that influence the probability of litigation do not vary across the dyads in which they appear. For instance, there may be unobserved attributes that influence a patent owner's underlying propensity to litigate or a competitor's underlying susceptibility to being a target of litigation. Failing to account for non-independence of observations can result in systematic underestimation of standard errors (Lincoln, 1984; Stuart, 1998). A solution to the non-independence problem, proposed by Lincoln (1984) is to include a right-hand side variable which, for an observation containing actors *i* and *j*, is the mean value of the dependent variable across all other dyads containing either actor *i* or actor *j*.

To account for potential autocorrelation due to non-independence, I compute a variable which, for a focal patent owner and competitor, is the mean number of suits either sent by the focal patent owner or received by the focal competitor. Table 5.7 reports model estimates after introducing this control. The results are highly consistent with the results in Table 5.6. In *Model 4*, status difference has a negative and significant effect on the likelihood of litigation between two firms ($p<0.05$). The interaction between status difference and inconsistency has a positive and negative sign ($p<0.01$). This implies that that inconsistency in status difference between two firms attenuates the effect of status difference in reducing the likelihood of litigation. After adjusting for potential autocorrelation, the results become slightly stronger. The magnitudes of the coefficients for status difference and the interaction with inconsistency become larger. This may be due to the fact that the autocorrelation control provides an additional way to account for unobserved heterogeneity across observations (Stuart, 1998). These results provide additional support for *Hypotheses 1* and *2*.

In addition to adjusting for potential autocorrelation, I checked the robustness of my main results to other empirical choices. First, I checked that my results are robust to alternative measures of inconsistency in status difference. In the present analysis, the inconsistency variable has a non-zero value only for those dyads of firms that receive citations from at least two common subpopulations. Moreover, subpopulations that send citations to one firm but not the other do not contribute at all to the inconsistency score. This potentially biases the measure to have higher values for pairs of large, diversified firms that are more likely to receive citations from the same subpopulations. As an alternative, I computed an inconsistency score which treats a firm as having higher status in a subpopulation if the other firm in the dyad receives no citations from the subpopulation. This generates substantially more reversals of status difference across subpopulations. Another alternative to the present inconsistency measure is to look at the standard deviation in each firm's status scores across subpopulations. These standard deviations serve as an additional measure of the degree of noisiness of status as a signal of quality. In analyses not reported here, I re-estimated all models with these alternative measures and found consistent results. In all cases, the main effects of interest were in the predicted direction and statistically significant.

Second, I checked that my results are robust to a more inclusive sample of observations. As an alternative to the present sample of case-control observations matched on infringing products, I constructed a broader sample that includes all dyads of direct competitors in the semiconductor industry for the years 1984, 1988, 1992, and 1997. I define direct competitors as firms making at least one integrated circuit that performs the identical function, in other words, at least one perfect substitute for each other's products. This sample is much less restrictive than the case-control sample. However, it also presents a less precise counterfactual of instances in which litigation did not occur against which to compare the instances in which litigation did occur. Specifically, I do not know whether either of the two competitors actually owns patents covering the products they have in common and whether either has a propensity to enforce the patents. As a partial solution I restricted attention to those dyads in which at least one of the two competitors filed at least one patent infringement suit in the given year[16]. In analyses not reported here, I estimated all models using this broader sample and found consistent results. The main effects of interest were in the predicted direction and statistically significant.

Third, I checked that my results are robust to alternative estimation approaches. The conditional logit model is arguably the most appropriate given the matched case-control structure of the data. However, I also estimated all models using probit regression with random effects as an alternative way to account for unobserved heterogeneity across suits. The results were consistent. The main effects of interest were in the predicted direction and statistically significant.

Fourth, I checked that my results are robust to the inclusion of controls for underlying propensity to litigate. Research suggests that some firms may be more

[16] For this broader sample, I was not able to collect additional information from press reports about the products involved.

litigious than others, for instance, as part of a deliberate strategy to build a reputation for toughness (Agarwal et al, 2008). Failing to account for this unobserved variation in firms' baseline likelihoods of litigation could produce biased results. My current findings partially account for this by including Lincoln's (1984) suggested autocorrelation control. Since this counts the total number of other suits involving either the plaintiff or defendant in a focal suit, it provides a control for underlying variation in firms' propensities to litigate, and it is computed in essentially the same way as Agarwal, Ganco, and Ziedonis's (2008) measure of toughness. The main difference is that it controls for contemporaneous variation, whereas Agarwal, Ganco, and Ziedonis (2008) use a lagged measure. To incorporate lagged variation, I include two additional controls, one for the number of lawsuits filed by the focal plaintiff against the focal defendant in the preceding four-year period and one for the number of lawsuits filed against the focal defendant by any plaintiff in the preceding four-year period. These two variables attempt to account for further unobserved variation at the dyad-level and defendant-level, respectively, which is reflected in firms' past rather than contemporaneous activities and which could lead to different baseline likelihoods of litigation. I do not include an analogous control for number of past lawsuits filed by the plaintiff because it would not vary across observations within a case-control group and so would not enter into the conditional logit model. By conditioning on the case-control group, the model is essentially accounting for unobserved plaintiff-level variation. The results are not reported here because neither variable has a significant effect on likelihood of litigation, and including them has virtually no impact on the estimates of the other variables in the model.

Discussion

My aim in this study is to understand the role of status hierarchies on the conduct of disputes. I take as a starting point Podolny's (1993) explanation of how status processes restrict quality claims that an actor can make given its position and argue that these same forces also restrict the claims that actors are willing to make and consider from one another in a dispute. This helps to generate predictability, align expectations, and prevent disputes from escalating. I find empirical support for this argument based on analysis of an original data set that matches instances of infringement that resulted in litigation to comparable instances of infringement that did not. Analysis of this sample suggests that, conditional on infringement, two firms which are different in status are much less likely to litigate than firms which are similar in status. However, the effect of status difference is attenuated by the degree to which these status differences are consistent across different subpopulations in the industry that are responsible for conferring status. These results provide support for the argument that clear status differences reduce the likelihood that disputes escalate and conversely that status ambiguity increases the likelihood that disputes escalate.

There are several ways in which the theoretical arguments and empirical findings in this study could be further developed in future research. First, future research can examine whether absolute position in the status hierarchy influences the likelihood of engaging in litigation. This study focuses only on status difference, i.e. relative position in the status hierarchy, and does not differentiate among high-, middle-, and low-status actors. In doing so, it assumes that status differences are equally meaningful regardless of the absolute position of the actors. However, research suggests that absolute position may be important. Specifically, research suggests that high- and low-status actors are more likely than middle-status actors to engage in activities for which there is a stigma (Phillips & Zuckerman, 2001). In the context of this study, anecdotes in chapter 2 suggested the practice of making aggressive licensing demands is viewed with a certain stigma, specifically, as being antithetical to the strategy of profiting through superior innovation. If this is the case, then high- and low-status firms may be more willing than middle-status firms to engage in the kinds of aggressive licensing practices that put them at risk of litigation. This idea could have bearing on how to interpret the empirical findings of this study. Dyads of high- and low-status firms will have larger status differences than dyads of high- and middle- or low- and middle-status firms. The findings in this study suggest that litigation is most likely in disputes between middle-status firms and either high- or low-status firms. Litigation is least likely between high- and low-status firms. Further theoretical and empirical work is needed to examine whether this pattern is consistent with the theoretical mechanisms behind the middle-status conformity prediction (Phillips & Zuckerman, 2001).

Second, future research can extend the empirical analysis in this study by developing more nuanced conceptions of inconsistency. The analysis in this study relied on a simple measure which treats all subpopulations in an industry as contributing equally to a firm's status consistency. In practice, firms are unlikely to view all subpopulations as equally important for conferring status. It is possible that some technological fields or markets are considered to be more prestigious than others and so carry a greater weight in influencing how firms view one another's status. Future research can explore this idea by examining whether alternative weighting schemes do a better job of predicting how firms interact with one another. This may also provide a direction for future research on the middle-status conformity idea. If some technological fields carry more weight than others, then having high status in a prestigious field may free a firm to engage in activities that normally detract from status such as participating in less prestigious fields and using aggressive licensing strategies. This may provide an alternative explanation for the findings of this study that is consistent with the middle-status conformity argument.

Third, future research can examine how status influences licensing terms. This study focused on the influence of status on how successfully firms are able to reach licensing solutions but did not theorize explicitly or examine empirically the influence of status on the nature of licensing solutions. The theory assumes that higher-status firms demand and receive better licensing terms. Future research is needed to establish whether this is in fact the case. There is a significant practical obstacle in that licensing agreements that are not part of a court-ordered settlement are usually kept secret. However, future research may go beyond the context of patent disputes and identify contexts where licensing terms are publicly available.

Appendix

Table 5.1. Manufacturers of Integrated Circuit Devices, 1997

	Products
Min	1
25th percentile	33
50th percentile	123
75th percentile	485
Max	6,726
Mean	508.8
Std. Dev.	1003.7

Figure 5.1. Size Distribution of Sample Firms

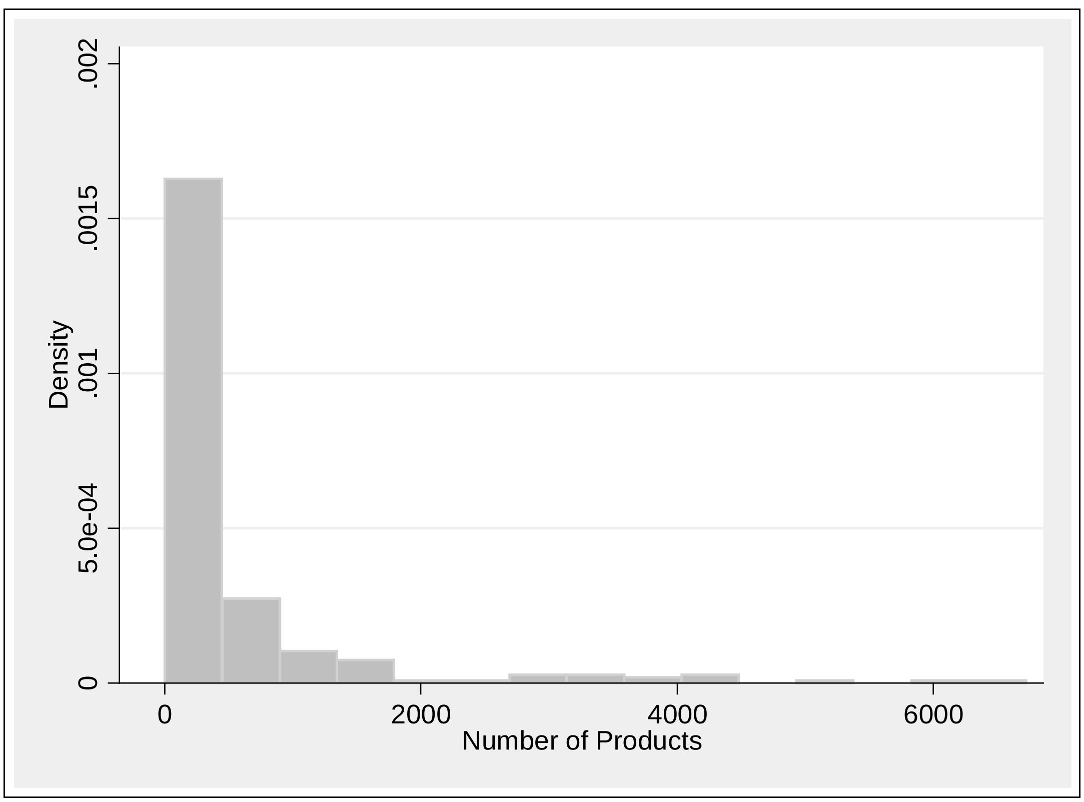


Table 5.2. Market Segments in the Semiconductor Industry, 1997[1]

Category	Segment	Firm	Firm Size (All Products)	Products in Segment	Specialization in Segment
Memory	SRAM	DensePac	4,115	2,416	0.59
(43,434)	(18,148)	IDT	4,220	2,411	0.57
		Cypress	2,835	1,432	0.51
	DRAM	Mitsubishi	3,705	1,392	0.38
	(10,382)	Samsung	3,222	1,370	0.43
		MicronTech	2,728	1,054	0.39
	PROM	ATMEL	1,661	604	0.36
	(4,595)	Cypress	2,835	488	0.17
		AMD	1,329	461	0.35
Linear	Operational Amplifiers	AD	3,268	676	0.21
(30,423)	(4,667)	TI	6,060	585	0.10
		National	6,726	447	0.07
	Telecommunications Circuits	LucentTech	406	191	0.47
	(3,674)	Mitel	149	144	0.97
		National	6,726	142	0.02
	Voltage Regulators	National	6,726	363	0.05
	(3,667)	SGSThomson	3,622	322	0.09
		Alpha	475	300	0.63
Digital	Bus-oriented Circuits	TI	6,060	701	0.12
(29,083)	(4,660)	IDT	4,220	508	0.12
		Philips	4,075	442	0.11
	Gates	National	6,726	471	0.07
	(3,482)	TI	6,060	409	0.07
		Motorola	5,058	326	0.06
	Delay Lines	ESC	925	925	1.00
	(2,885)	DataDelay	1,236	904	0.73
		BelFuse	777	714	0.92

(Table 5.2 continued on next page)

[1] The table lists the three largest producers in the three largest segments in each category.

Table 5.2 (continued). Market Segments in the Semiconductor Industry, 1997[1]

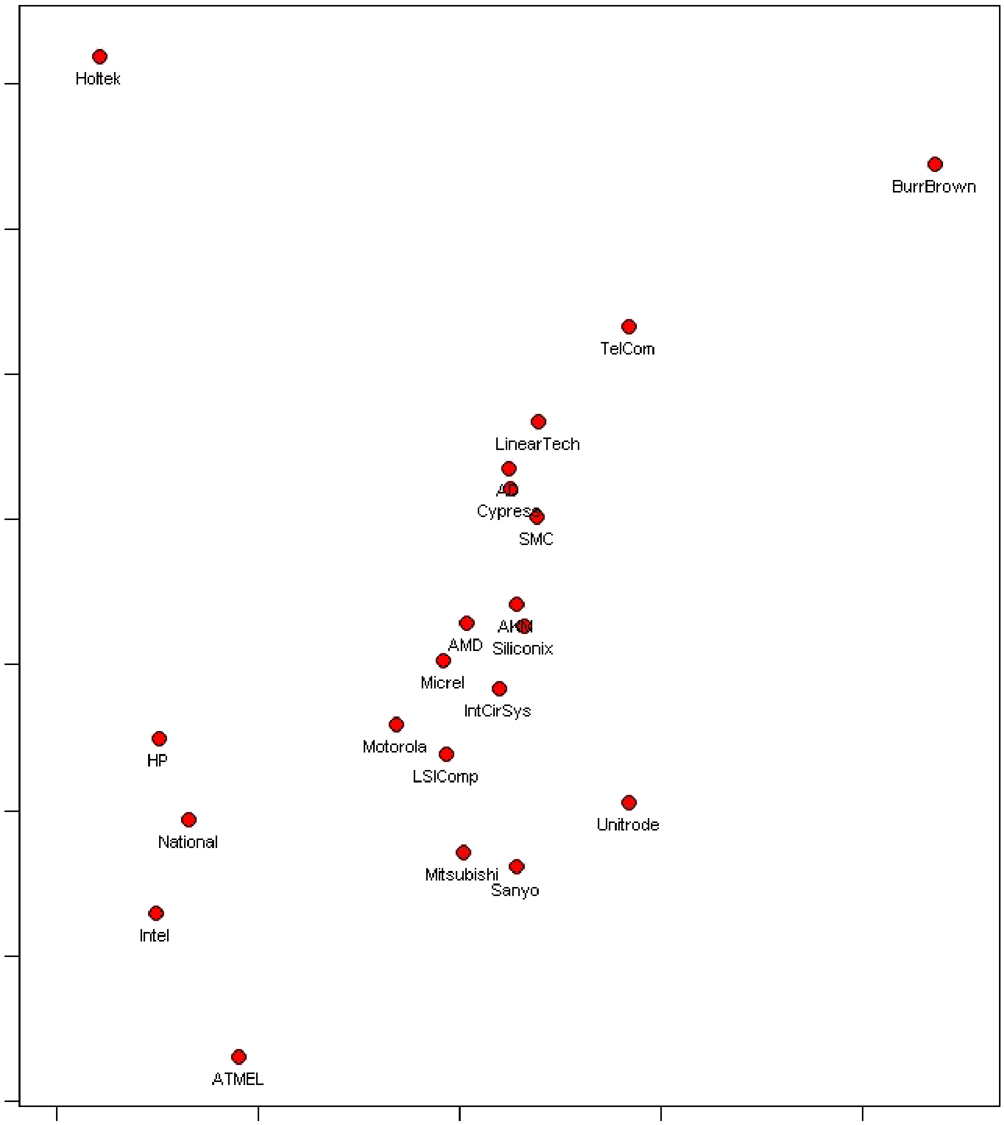
Category	Segment	Firm	Firm Size (All Products)	Products in Segment	Specialization in Segment
Interface	Analog-digital Converters	AD	3,268	609	0.19
(13,533)	(3,363)	Maxim	2,365	329	0.14
		MicroNet	396	251	0.63
	Memory and Peripherals Drivers	SiliconSys	360	332	0.92
	(2,795)	Allegro	597	276	0.46
		National	6,726	221	0.03
	Digital-analog Converters	AD	3,268	795	0.24
	(2,724)	Maxim	2,365	371	0.16
		BurrBrown	1,430	200	0.14
Processors	8-bit	Philips	4,075	316	0.08
(4,114)	(2,504)	Zilog	307	285	0.93
		Microchip	521	209	0.40
	4-bit	NEC	1,883	164	0.09
	(544)	Toshiba	3,006	95	0.03
		Panasonic	1,548	69	0.04
	16-bit	Intel	1,163	124	0.11
	(531)	Mitsubishi	3,705	117	0.03
		AMD	1,329	47	0.04

[1] The table lists the three largest producers in the three largest segments in each category.

Table 5.3. Distribution of cases and controls across infringement suits

Category	Number of suits	Average cases per suit (competitors sued)	Average controls per suit (competitors not sued)	Ratio (controls per case)
Digital	2	1.0	4.0	4
Interface	12	1.1	8.7	8
Linear	10	1.3	10.6	8.2
Memory	17	1.6	25.1	15.8
Processor	10	1.0	10.4	10.4
Total	51 suits	65 cases	749 controls	11.5

Figure 5.2. Subpopulations in the Semiconductor Industry: Relative Positions in Market Space, 1997[1]



[1] This figure shows firm positions in market space, scaled to two dimensions. For readability, the largest firm in each subpopulation is shown.

Table 5.4. Subpopulations in the Semiconductor Industry, 1997

Subpopulation	Representative Segment	Category	Average Specialization in Segment	Average Products in Segment	Subpopulation Breadth	Firm (Largest, Smallest)	Size
1	32-bit Processor	Processor	0.74	15	8	Temic 8x8	207 1
2	Line Circuits	Interface	0.50	6	3	NJR Optotek	667 2
3	NOVSRAM	Memory	0.64	79	9	Dallas Greenwich	629 19
4	Telecommunication Circuits	Linear	0.64	43	38	Holtek Sensory	600 2
5	DC-DC Converters	Linear	0.75	204	17	BurrBrown PCA	1430 31
6	Delay Lines	Digital	0.87	674	3	DataDelay MikroChips	1236 6
7	8-bit Processor	Processor	0.52	146	2	Intel WDC	1163 36
8	Voltage Regulators	Linear	0.41	129	38	Micrel Vortex	694 59
9	EEPROM	Memory	0.66	142	12	ATMEL Lattice	1661 2
10	ROM	Memory	0.44	68	14	AKM Hughes	314 77
11	PROM	Memory	0.78	151	5	AMD Altera	1329 7
12	Optoelectronic Devices	Linear	0.72	79	4	HP Dalsa	141 44
13	Clock Circuits	Linear	0.57	40	4	IntCirSys Focus	148 1
14	Transmitters-receivers	Interface	0.60	19	8	SMC Eurom	50 6
15	Analog-digital Converters	Interface	0.33	148	42	AD Analogic	3268 86

(Table 5.4 continued on next page)

Table 5.4 (continued). Subpopulations in the Semiconductor Industry, 1997

Subpopulation	Representative Segment	Category	Average Specialization in Segment	Average Products in Segment	Subpopulation Breadth	Firm (Largest, Smallest)	Size
16	SRAM	Memory	0.59	529	51	Motorola ArrayMicro	5058 20
17	Operational Amplifiers	Linear	0.45	83	28	LinearTech TLSI	1265 4
18	Analog Switches	Interface	0.46	116	5	Siliconix Supertex	267 200
19	Radio Circuits	Linear	0.59	4	7	LSIComp RFMicro	92 2
20	Pulse-width Modulators	Linear	0.46	116	5	Unitrode ACCMicro	591 2
21	SRAM	Memory	0.36	291	24	Cypress Music	2835 19
22	Audio Circuits	Linear	0.48	47	25	Sanyo Seponix	466 1
23	Memory and Peripheral Drivers	Interface	0.60	83	39	TelCom Adaptec	797 2
24	DRAM	Memory	0.58	438	68	Mitsubishi EMS	3705 49
25	Gates	Digital	0.14	196	75	National Universal	6726 42

Figure 5.3. Inconsistency in Status Difference: Example

Subpopulation	Representative Segment	Higher status firm
1	Microprocessors, 32-bit	Hitachi
4	Telecommunication circuits	Motorola
5	DC-DC converters	Hitachi
7	Microprocessors, 8-bit	Motorola
9	EEPROM	Hitachi
10	ROM	Motorola
11	PROM	Hitachi
14	Transmitters-receivers	Motorola
15	Analog-digital converters	Motorola
16	SRAM	Hitachi
17	Operational amplifiers	Motorola
18	Analog switches	Motorola
19	Radio circuits	Motorola
22	Audio circuits	Hitachi
23	Memory drivers	Hitachi
24	DRAM	Motorola
25	Standard logic gates	Hitachi

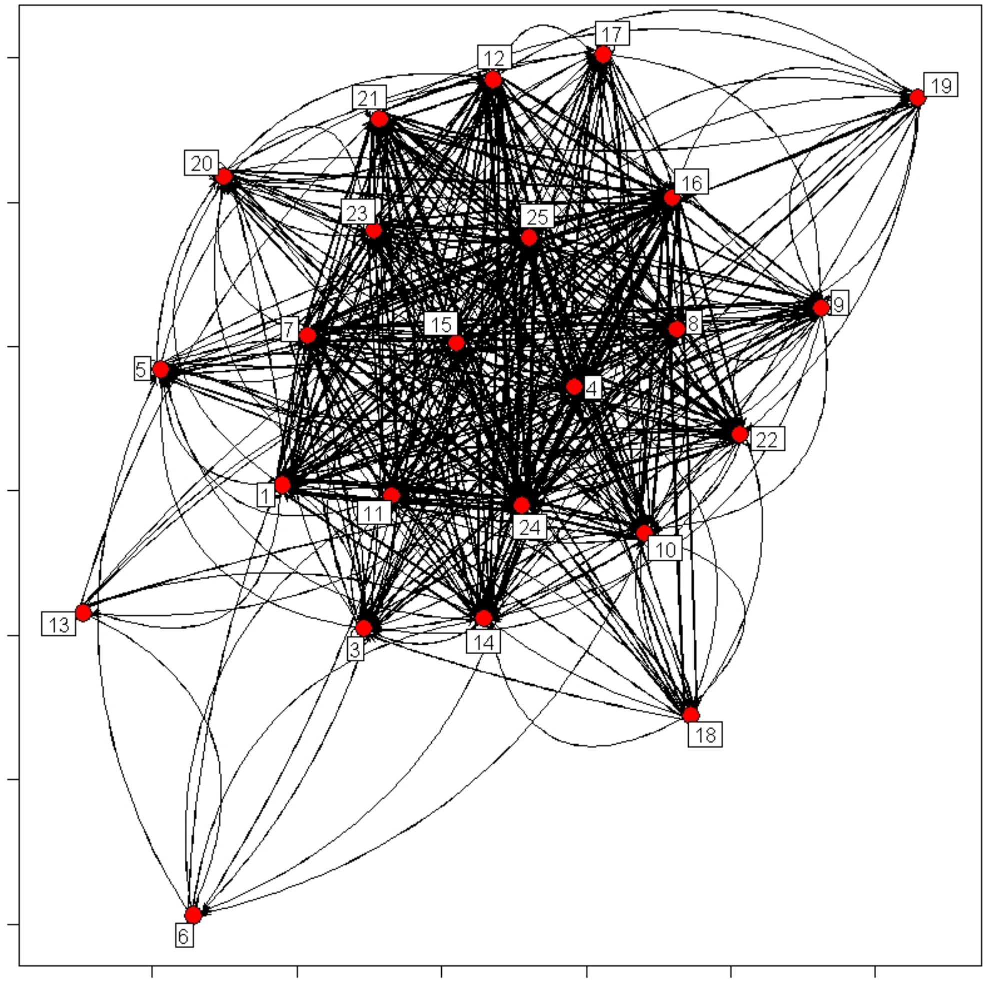
Inconsistency (Hitachi, Motorola) =

(Proportion in which **Hitachi** has higher status) (Proportion in which **Motorola** has higher status)

(0.47) (0.53)

= 0.249

Figure 5.4. Patent Citations between Subpopulations, 1997 [1] [2]



[1] The numbers in this figure correspond to the subpopulation numbers in Table 5.4.
[2] Line width corresponds to the total number of citations from one subpopulation to another in the preceding four years.

Table 5.5. Descriptive Statistics

	N=814	Mean	Std. Dev.	Min	Max
1	Suit filed	0.1	0.3	0	1
2	Year is 1992-2000	0.6	0.5	0	1
3	Infringing products made by competitor	3.5	6.7	1	55
4	Directly competing products	38.2	76.1	1	416
5	Relative products (thousands)	1.3	1.1	0	6.70
6	Relative patents (thousands)	1.1	1.5	0	8.20
7	Same subpopulation	0.2	0.4	0	1
8	Relative status of patent owner	0.0	0.1	-0.28	0.28
9	Status difference	0.0	0.1	0	0.28
10	Inconsistency in status difference	0.0	0.1	0	0.25

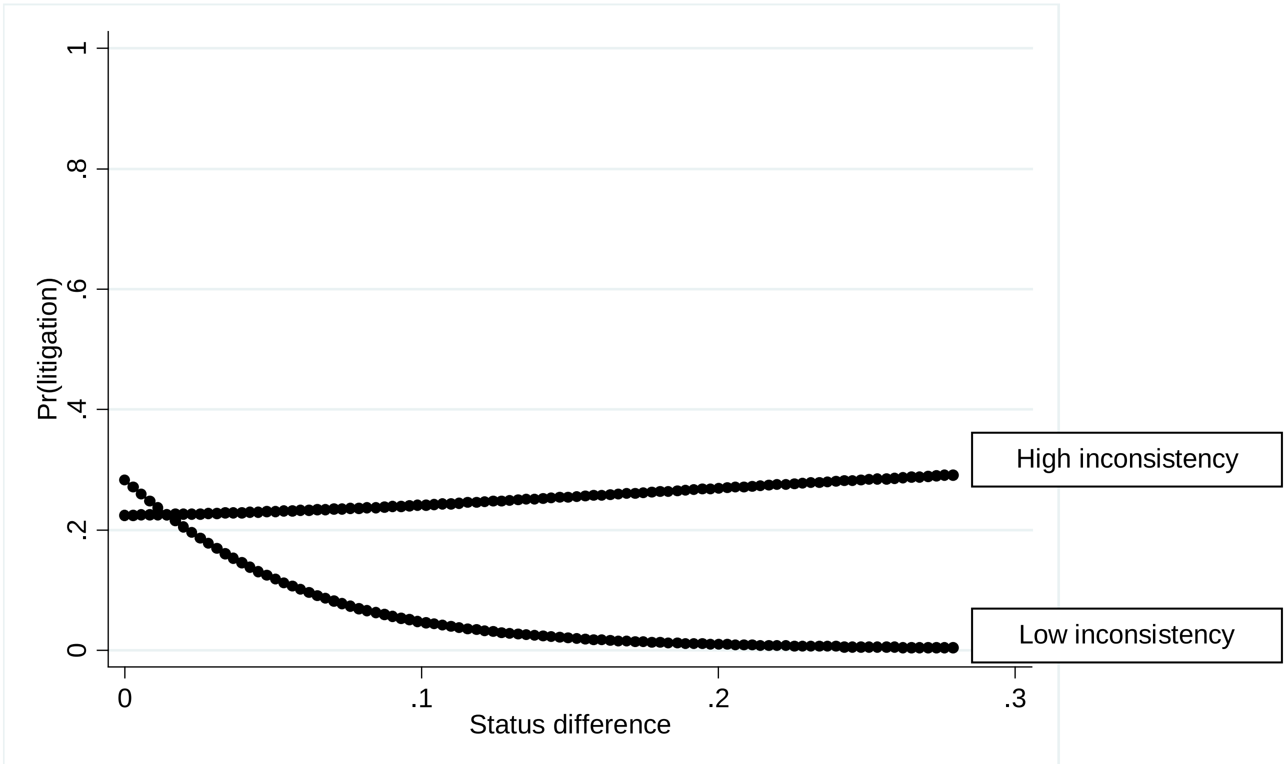
	N=814	1	2	3	4	5	6	7	8	9
1	Suit filed									
2	Year is 1992-2000	-0.05								
3	Infringing products made by competitor	0.06	0.26							
4	Directly competing products	0.02	-0.17	-0.03						
5	Relative products (thousands)	0.05	0.02	0.08	0.03					
6	Relative patents (thousands)	-0.08	0.32	0.24	0.07	0.03				
7	Same subpopulation	0.09	-0.05	0.02	0.49	-0.11	0.06			
8	Relative status of patent owner	-0.03	0.10	-0.12	-0.03	-0.09	0.19	-0.08		
9	Status difference	-0.05	0.06	-0.01	0.14	-0.08	0.69	0.03	0.51	
10	Inconsistency in status difference	0.08	0.03	-0.07	0.12	-0.03	-0.12	0.03	-0.07	-0.06

Table 5.6. Main results

Conditional logit estimates of likelihood of litigation between patent owner and competitor					Logit with unconditional fixed-effects
	(1)	(2)	(3)	(4)	(5)
Year is 1992-2000	-0.133	-0.185	-0.187	-0.202	-0.252
	(0.907)	(0.893)	(0.893)	(0.895)	(1.016)
Infringing products made by competitor	0.945**	0.905**	0.904**	1.010**	1.216***
	(0.192)	(0.195)	(0.196)	(0.206)	(0.229)
Directly competing products	-0.005+	-0.006*	-0.006*	-0.008*	-0.009*
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Relative products (thousands)	-0.128	-0.148	-0.148	-0.199	-0.279
	(0.167)	(0.169)	(0.170)	(0.183)	(0.211)
Relative patents (thousands)	-0.390+	-0.225	-0.233	0.222	0.267
	(0.225)	(0.228)	(0.235)	(0.282)	(0.313)
Same subpopulation	1.299**	1.264**	1.266**	1.246*	1.494**
	(0.482)	(0.481)	(0.481)	(0.489)	(0.526)
Relative status of patent owner	-5.417	-11.934*	-12.391+	4.676	5.777
	(4.100)	(5.672)	(6.427)	(7.867)	(8.474)
Status difference		-13.563+	-13.908+	-22.658*	-27.685*
		(7.761)	(8.095)	(11.038)	(11.501)
Inconsistency in status difference			-0.330	-2.835	-3.429
			(2.194)	(2.476)	(2.690)
Status difference * Inconsistency				181.271**	225.408**
				(64.222)	(68.245)
Observations	814	814	814	814	814

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

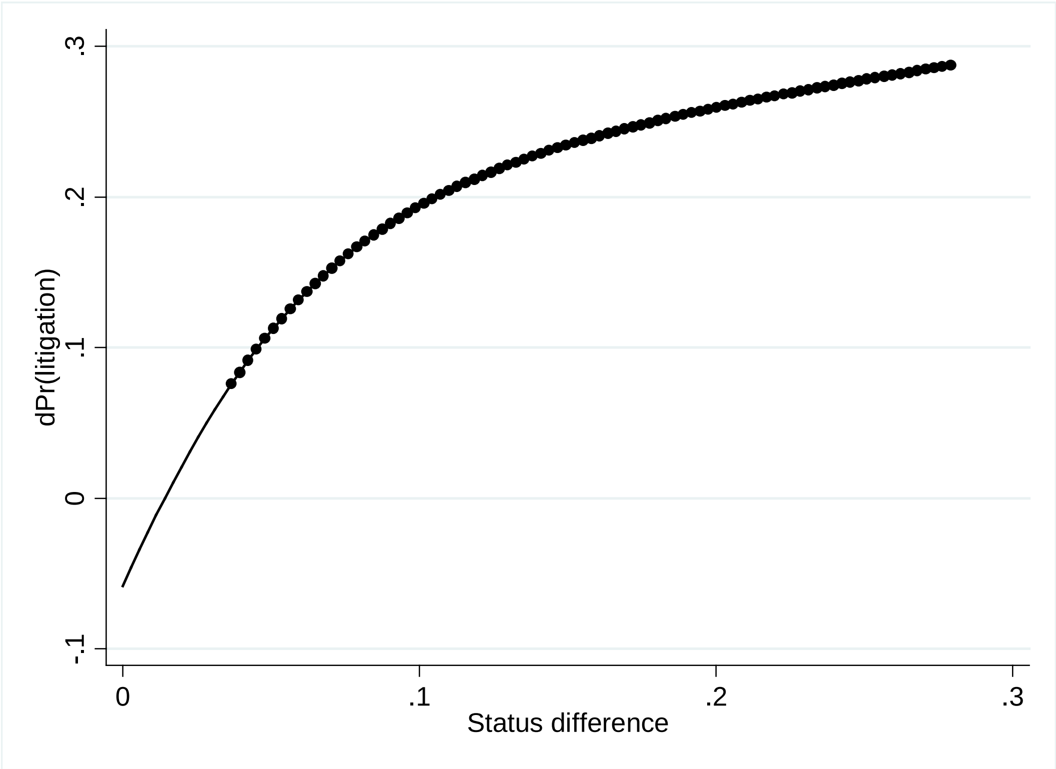
Figure 5.5. Effect of status difference on predicted probability of litigation[1][2][3]



---

[1] Dots indicate predicted values that are significantly different from zero (95% confidence interval does not include zero).
[2] High = sample mean(inconsistency) + 1 std.dev.; low = sample mean(inconsistency) – 1 std.dev.
[3] Estimates based on logit model with unconditional fixed-effects

Figure 5.6. Change in effect of status difference between high and low inconsistency[1][2][3]



[1] Dots indicate predicted values that are significantly different from zero (95% confidence interval does not include zero).
[2] High = sample mean(inconsistency) + 1 std.dev.; low = sample mean(inconsistency) – 1 std.dev.
[3] Estimates based on logit model with unconditional fixed-effects

Table 5.7. Main results adjusted for autocorrelation

Conditional logit estimates of likelihood of litigation between patent owner and competitor				
	(1)	(2)	(3)	(4)
Year is 1992-2000	-0.406	-0.456	-0.458	-0.423
	(0.933)	(0.917)	(0.917)	(0.899)
Infringing products made by competitor	0.980**	0.944**	0.944**	1.042**
	(0.198)	(0.202)	(0.202)	(0.212)
Directly competing products	-0.005+	-0.006+	-0.006+	-0.008+
	(0.003)	(0.003)	(0.003)	(0.004)
Relative products (thousands)	-0.159	-0.200	-0.199	-0.230
	(0.187)	(0.195)	(0.195)	(0.214)
Relative patents (thousands)	-0.514*	-0.348	-0.356	0.217
	(0.229)	(0.230)	(0.239)	(0.301)
Same subpopulation	1.227*	1.215*	1.213*	1.139*
	(0.536)	(0.535)	(0.536)	(0.549)
Relative status of patent owner	-6.696+	-13.175*	-13.586*	4.270
	(4.061)	(5.688)	(6.450)	(7.738)
Status difference		-13.432+	-13.729+	-26.781*
		(7.898)	(8.204)	(12.316)
Inconsistency in status difference			-0.323	-3.203
			(2.390)	(2.744)
Status difference * Inconsistency				206.570**
				(70.956)
Mean number of suits in other dyads including the focal patent owner or competitor	-9.475**	-9.495**	-9.489**	-9.619**
	(2.952)	(2.981)	(2.982)	(2.938)
Observations	814	814	814	814

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Chapter 6: Product line expansion in the EEPROM market

Introduction

In this chapter, I examine the impact of status differentiation and consistency in status differentiation on firms' product line sizes. Firms should be more likely to increase their product line sizes when they occupy a differentiated place in the status hierarchy. Moreover, the effect of status differentiation should depend on the consistency of the differentiation. I test these predictions in the context of the electrically-erasable programmable read-only memory (EEPROM) market, using data on expansion and contraction in firms' product lines from 1977 through 2001.

The logic behind studying firms' product line sizes is that bringing products to market both allows firms to profit from innovation and also exposes them to litigation risk. The number of different products offered by a firm reflects its coverage of the various pockets of consumer preferences that make up the market. A firm can spread the cost of innovation over a larger market if it introduces more product variants embodying the same innovations (Cohen & Klepper, 1992). However, offering more products also brings firms into competition with one another. In the semiconductor industry, where many firms own patents covering complementary or cumulative innovations, they either reach licensing agreements to give one another freedom to introduce products or enter into costly disputes that deter one another from introducing products. Hence, studying firms' product line sizes gives a window on how firms divide returns to innovation.

Context

I test hypotheses 3 and 4 in the context of the worldwide semiconductor industry, focusing on the market for EEPROM (Electrically-Erasable Programmable Read Only Memory) devices from 1977 to 2001. The market emerged following the invention of the first true electrically-erasable PROM device in 1978 by George Perlegos at Intel (Rostky, 2002). Previously, EPROM (erasable PROM) technology allowed information stored in memory chips to be programmed and erased, but it required a lengthy and cumbersome exposure to ultraviolet radiation. In addition, the key drawback was that information stored on EPROM chips could only be erased in bulk. Perlegos's invention was the first nonvolatile memory device that could be rapidly erased and programmed one byte at a time without ultraviolet exposure. In 1981, Perlegos left Intel along with two other employees to found Seeq Technlogy, the first de novo entrant into the new EEPROM market. By 1985, a variety of firms had entered the market, including Advanced Micro Devices, Motorola, and National. The EEPROM market offers the advantage of a long observation period while allowing me to observe the evolution of a new market from virtually its inception.

There are several reasons I focus on one market in this study. The first is that it allows me to better isolate the effects of status inconsistency as developed in hypothesis 4. In hypothesis 4, I suggested that endorsements of lower-status actors by higher-status actors are contradictory to the deference rules normally implied by the status difference between higher and lower-status actors. As a result, the presence of these inconsistencies weakens the effect of status differences in helping firms to manage disputes. Studying these inconsistencies across markets is problematic because firms may have different status in different markets. An endorsement of a lower-status actor in one market by a higher-status actor in another market does not have the same interpretation.

The second reason I focus on one market is that it allows me to collect more detailed information on firms' products and patents. Within a single market, I am able to collect complete data on a firm's product line for every year over the life of the firm and market. This allows me to study within-firm changes in product line size. In addition, focusing on one market allows me to better isolate patents specific to the market. Semiconductor patents are normally difficult to link to meaningful, specific product categories. If the network of endorsements among firms in a market (which comprises the status hierarchy) provides them with a mechanism for dividing returns to innovation, then it is important to identify the network of endorsements associated with the innovations and products in that market. Focusing on a single market allows me to better identify the patents specific to the market and therefore construct the appropriate network of endorsements thought to influence opportunities and constraints in the market.

Data Sources

Data on EEPROM producers and product lines come from *IC Master*, a directory of integrated circuit devices published annually since 1977. The directory provides a comprehensive listing of products in all major categories including memory, digital, linear, interface, and microprocessor. From *IC Master*, I was able to collect data on 13,709 EEPROM devices produced by 68 firms between the years 1977 and 2001[1].

---

[1] I was not able to obtain the 1979, 1981, 1996, 1998, and 2000 editions of *IC Master*. Observations for these years are treated as missing.

I use data on patent citation patterns to compute measures of status differentiation and status consistency. I collect patent data from two sources. First, I used Google's full-text patent search to identify patents on EEPROM technology. To identify these patents, I used the terms *erasable* + *programmable* + (*read-only memory* or *ROM*), *EEPROM*, excluding *random access*, *RAM*, and *flash*. This process yielded 3,056 patents issued between 1976 and 2001. Second, I matched these patents to citation data in the NBER patent database (Hall, Jaffe, & Trajtenberg, 2001).

Dependent variable

Product line size. The aim of this study is to examine expansion and contraction in a firm's position in the market. I measure this using the size of a firm's product line as the dependent variable. Each product listed in *IC Master* is identified by a unique device code. For each firm in each year, I counted the number of distinct products to obtain product line size. When modeled using firm fixed-effects, product line size as a dependent variable effectively captures expansion and contraction over time.

Independent variables

Status differentiation. Following previous studies on status (e.g. Podolny, Stuart, & Hannan, 1996; Podolny, 2001; Jensen, 2003), I measure status using Bonacich's (1987) centrality score. Conceptually, this measure is derived from the notion that an actor's status is a function of deference ties sent to the actor, weighted by the status of the senders. Since each sender contributes to the status scores of all other actors in a network, which in turn directly and indirectly contributes back to status of the sender, the set of status scores of all actors in a network is determined recursively.

Following Podolny, Stuart, and Hannan's (1996) study of status in the semiconductor industry, I use citations between firms' patent portfolios as measures of deference ties. I construct a matrix whose typical element is the number of citations that a firm *i* receives from a firm *j* in a five-year window up to but not including a given year *t*. Since my focus is on status within the context of the EEPROM market, I use only citations between patents which I have identified as covering EEPROM technology (both citing and cited patents are EEPROM).

Bonacich's (1987) approach identifies a vector of status scores (one for each firm) such that each firm's status is proportional to the sum of status scores of other firms from which it receives deference (as given by the citation matrix). Each of these other firms' status scores must be proportional to the sum of status scores of other firms from which they receive deference, and so on. The status scores for a network of *n* firms comes out as the solution to a system of *n* equations with *n* unknowns.

In my analyses, I use the 'raw' status scores in two ways. First, I order the status scores for all firms in a given year and take each firm's rank as its vertical position in the market status hierarchy[1]. I include this as a control variable in all models as *status rank*. Second, I compute the difference between each firm's status score and the scores of the firms immediately above and immediately below the focal firm in the status hierarchy. I take the absolute value of these differences to obtain continuous measures of differentiation in the focal firm's position. These measures enter into the analyses in two

[1] It is possible for two firms to have the same status rank if they have precisely the same status score.

ways: separately, as *upward differentiation* and *downward differentiation*, and summed together as *status differentiation*.

Status inconsistency. As developed in the theory section, status inconsistency reflects the tendency for the dyadic relationship between two actors in a network to appear to contradict their overall rankings when the dyadic relationship is considered in isolation. In the context of intellectual property and the flow of technological innovations, this is reflected in the tendency for 'downstream' firms to be in positions of technological leadership over the 'upstream' firms on whose technologies they build. In other words, status inconsistency reflects the tendency for deference to flow from higher-ranked actors to lower-ranked actors. To measure this I create two indicator variables. The first variable, *downward inconsistent*, takes a value of '1' if a focal firm cites patents of the firm ranked immediately below it in the status hierarchy in a given year and a value of '0' otherwise. The second variable, *upward inconsistent*, takes a value of '1' if a focal firm's patents are cited by the firm ranked immediately above it in the status hierarchy in a given year and a value of '0' otherwise.

Model

The predictions in this study concern growth in firms' market positions. I model this using fixed-effects Poisson regression with firm product line size, a count measure, as the dependent variable. The unit of observation is the firm-year. The use of a model with firm fixed-effects is important, since the inclusion of a time-invariant, firm-specific component allows within-firm, across time variation in product line size to be interpreted as expansions and contractions.

An alternative model is the more commonly used fixed-effects negative binomial model (Hausman, Hall, and Griliches, 1984). This model has an advantage over Poisson models in that it allows for overdispersion through the inclusion of an additional variance parameter. However, it has the disadvantage of producing inconsistent maximum likelihood estimates if the underlying distribution is misspecified, i.e. the dependent variable is not negative binomial.

Wooldridge (1997) shows that the fixed-effects Poisson model produces consistent quasi-maximum likelihood estimates under more general conditions and, hence, is more robust to misspecification. In addition, it is robust in the presence of arbitrary dependence between observed independent variables and the unobserved component. In all models, I compute the robust standard errors recommended by Wooldridge (1997) using Tim Simcoe's (2007) xtpqml command for Stata.

Control variables

Applications. I include a control for number of EEPROM patent applications filed by a firm in a given year (that were ultimately granted) to account for possible unobserved changes in firm innovative activity that may be driving changes in product line size.

Patents. I include a control for patent portfolio size, computed as the number of patents received by a firm in the four-year window up to but not including the focal year.

Product line breadth. I include a control for number of market segments spanned by a firm's product line. I define a market segment based on memory capacity, i.e. 4K segment.

Status rank. Since I am primarily interested in differentiation around a firm's position, it is important to account for the effect of simply being higher or lower in the status hierarchy. Hence, I include a control for rank in the status hierarchy. As described above, I compute this by ordering the Bonacich (1987) centrality scores for all firms in a given year and give the firm with the highest centrality score a status rank of '1', the next highest a rank of '2', and so on.

Table 6.1 gives descriptive statistics of my sample.

Results

Table 6.2 gives results from fixed-effects Poisson models of product line size. Model 1 includes only the control variables in order to establish baseline levels of firms' product line sizes. All controls have significant positive effects on product line size. The results suggest that firms expand their product lines in years when they increase their number of patent applications, the size their patent portfolios, and the breadth of their product lines. The positive effect of status rank is somewhat surprising. It implies that firms expand (contract) their product lines in years when they fall (rise) in rank.

Model 2 adds the variable for status differentiation. As predicted, status differentiation has a positive and significant effect on product line size ($p<0.05$). The result implies that firms expand their product lines when the status difference between them and their competitors increases. Estimates for all controls remain stable after including the variable for status differentiation.

Model 3 splits the overall effect of status differentiation into its two parts: upward differentiation from a firm's competitor immediately above it in the status hierarchy and downward differentiation from a firm's competitor immediately below it in the status hierarchy. Both variables have positive effects. Downward differentiation has a stronger effect and is significant at the 5% level, whereas upward differentiation has a weaker effect and is significant only at the 10% level.

Tables 6.3 and 6.4 report model estimates after adding the interaction effect between status differentiation and inconsistency. Table 6.3 shows the moderating effect of inconsistency on upward status differentiation. Model 1 gives the main effect of upward status differentiation. Model 2 adds the interaction effect between upward status differentiation and inconsistency. Because inconsistency is measured as a dummy variable, the interaction effect is easier to interpret if the effect of status differentiation is broken out into two parts: when inconsistency is present and when inconsistency is not present. Consistent with predictions, the positive effect of status differentiation appears to be weaker when inconsistencies are present and stronger when inconsistencies are not present. When inconsistencies are not present, status differentiation has the predicted positive effect on product line size ($p<0.05$). When inconsistencies are present, however, the positive effect of status differentiation is no longer significant.

Table 6.4 shows the analogous moderating effect of inconsistency on downward status differentiation. Model 1 gives the main effect of downward status differentiation. Model 2 adds the interaction effect between upward status differentiation and inconsistency. Again for easier interpretation, the interaction effect is incorporated by breaking out the effect status differentiation into two parts: when inconsistency is present and when inconsistency is not present. Consistent with predictions, the positive effect of status differentiation is stronger when inconsistencies are not present and weaker when inconsistencies are present. When inconsistencies are not present, status differentiation has the predicted positive effect on product line size ($p<0.05$). However, when inconsistencies are present, status differentiation does not have the predicted positive effect and in fact has a significant negative effect on product line size ($p<0.01$). On the one hand, the fact that status differentiation only has a positive effect when inconsistencies are not present provides support for predicted moderating effect of inconsistency. On the other hand, the significant negative effect, while not clearly contrary to theory, is not predicted by the theory in this study either. The results should be interpreted with caution as support for the predicted moderating effect.

Table 6.5 shows estimates from the full model, which includes the moderating effects of inconsistency for both upward and downward differentiation. Models 1 and 2 reproduce the results from the previous two tables when the moderating effects were modeled separately for upward and downward differentiation. Model 3 gives the full model. When the moderating effects of inconsistency for both upward and downward differentiation are added together, the estimates remain consistent with previous models and consistent with predictions. When inconsistencies are not present, status differentiation has the predicted positive effect on product line size for both upward and downward differentiation. When inconsistencies are present, the predicted positive effect goes away. Again, though, downward differentiation has an unpredicted negative effect in the presence of inconsistencies. Overall, the results provide support for the predicted moderating effect, but the unpredicted negative effect suggests caution in interpreting the results.

Additional analyses

Table 6.6 reports results of additional analyses to further explore the main findings. Models 1 and 2 report the results of splitting the sample into years of product line expansion and years of product line contraction to examine whether status differentiation has symmetric effects for growth and decline. In Model 1, status differentiation has a significant positive effect on product line expansion. In model 2, status differentiation does not have a significant negative effect on product line contraction.

Model 3 replaces product line size with probability of exit from the market as a dependent variable. If decreased status differentiation does not predict product line contraction, then it may be that it increases likelihood of exit. Model 3 reports results of a logit model with firm fixed-effects. Status differentiation does not appear to have a significant effect on probability of exit.

Models 4 and 5 in Table 6.6 examine whether changes in firm product lines may be due to unobserved changes in firm innovative activity. I split the dependent variable of product line size into 'new technology' and 'old technology' products. For each year, I identified each firm's technological frontier as the fastest speed of its highest capacity memory chip. I measure speed as nanoseconds of access time and capacity as number of bytes of memory. I define a product as 'new technology' if it is beyond the firm's technological frontier for the previous year, i.e. it must improve on at least one dimension while being at least equivalent on the other. Otherwise, I define a product as 'old technology'.

Model 4 uses number of new technology products as the dependent variable. Downward status differentiation has a slight positive effect, while upward differentiation has no significant effect. Model 5 uses number of old technology products as the dependent variable. Consistent with the main results, both downward and upward status differentiation have positive and significant effects on product line size. Also consistent with main results, downward differentiation has a stronger effect than upward differentiation. These results provide some evidence that increases in product line size due to status differentiation reflect expansion in firms' market positions rather than unobserved technological innovation.

Discussion

My aim in this chapter is to understand the influence of status ambiguity on the viability of positions in a status hierarchy. Status differences between firms provide signals of quality differences. To the extent that firms can clearly differentiate themselves from one another in quality and, more importantly, to the extent that firms recognize the distinctiveness of one another's positions, then these are viable positions in the role structure of the market (White, 1981).

Studying the effect of status ambiguity on firms' product line sizes provides a window on the viability of firms' positions. In previous chapters, I have argued and found evidence that status difference reduces the likelihood of conflict. I have also found evidence that the effect of status difference is weakened by the presence of inconsistency. In this chapter, I find that firms increase their product line sizes more when they have high status differentiation. I also find that the effect of status differentiation is weakened by the presence of inconsistency. These results suggest that the same factors leading to costly conflict also deter firms from bringing products to market.

In the context of innovation, these results are consistent with the idea that bringing products to market both allows firms to profit from innovation and also exposes them to litigation risk. In the semiconductor industry, where many firms own patents covering complementary or cumulative innovations, firms either reach licensing agreements to give one another freedom to introduce products or enter into costly disputes that deter one another from introducing products. Complementing the findings in chapter 5, the findings in this chapter about how status ambiguity influences product line size give an additional window on how firms divide returns to innovation.

Appendix

Table 6.1. Descriptive statistics

		Mean	Std. Dev.	Min	Max	(1)	(2)	(3)	(4)	(5)
(1)	Products	25.15	59.13	0	636					
(2)	Densities	3.35	3.18	0	14	0.669				
(3)	Applications	1.85	3.65	0	29	-0.033	0.005			
(4)	Patents	4.44	9.31	0	65	0.022	0.083	0.661		
(5)	Status rank	10.94	7.04	1	26	0.147	0.173	-0.127	-0.151	
(6)	Status differentiation	0.13	0.30	0	1.79	-0.016	-0.054	0.295	0.401	-0.352

Table 6.2. Fixed-effects Poisson regression of product line size

	(1)	(2)	(3)
Controls			
Applications	0.046**	0.046**	0.047**
	(0.016)	(0.015)	(0.015)
Patents	0.019*	0.018*	0.018*
	(0.008)	(0.008)	(0.008)
Product breadth	0.213**	0.211**	0.212**
	(0.020)	(0.020)	(0.020)
Status rank	0.029+	0.035*	0.035*
	(0.015)	(0.014)	(0.014)
Status differentiation			
Status differentiation		0.447*	
		(0.176)	
Upward differentiation			0.323+
			(0.180)
Downward differentiation			0.731*
			(0.327)
Firm-year observations	546	529	529
Firms	59	59	59

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 6.3. Moderating effect of inconsistency on upward status differentiation

	(1)	(2)
Controls		
Applications	0.047**	0.042**
	(0.015)	(0.014)
Patents	0.018*	0.014+
	(0.008)	(0.007)
Product breadth	0.212**	0.209**
	(0.020)	(0.019)
Status rank	0.035*	0.031*
	(0.014)	(0.013)
Downward status differentiation	0.731*	0.510**
	(0.327)	(0.186)
Presence of inconsistency (1/0)		0.191**
		(0.044)
Status Differentiation		
Upward status differentiation	0.323+	
	(0.180)	
Status Differentiation × ***Inconsistency***		
Upward differentiation when inconsistency = 1		-0.054
		(0.213)
Upward differentiation when inconsistency = 0		0.467*
		(0.185)
Firm-year observations	529	546
Firms	59	59

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 6.4. Moderating effect of inconsistency on downward status differentiation

	(1)	(2)
Controls		
Applications	0.047**	0.045**
	(0.015)	(0.014)
Patents	0.018*	0.019*
	(0.008)	(0.009)
Product breadth	0.212**	0.210**
	(0.020)	(0.019)
Status rank	0.035*	0.034*
	(0.014)	(0.014)
Upward differentiation	0.323+	0.268
	(0.180)	(0.177)
Presence of inconsistency (1/0)		0.088*
		(0.041)
Status Differentiation		
Downward status differentiation	0.731*	
	(0.327)	
Status Differentiation × ***Inconsistency***		
Downward differentiation when inconsistency = 1		-2.177**
		(0.791)
Downward differentiation when inconsistency = 0		0.958**
		(0.305)
Firm-year observations	529	529
Firms	59	59

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 6.5. Moderating effect of inconsistency on status differentiation: full model

	(1)	(2)	(3)
Controls			
Applications	0.042**	0.045**	0.040**
	(0.014)	(0.014)	(0.013)
Patents	0.014+	0.019*	0.017*
	(0.007)	(0.009)	(0.008)
Product breadth	0.209**	0.210**	0.208**
	(0.019)	(0.019)	(0.019)
Status rank	0.031*	0.034*	0.031*
	(0.013)	(0.014)	(0.012)
Upward differentiation		0.268	
		(0.177)	
Downward differentiation	0.510**		
	(0.186)		
Upward inconsistent	0.191**		0.193**
	(0.044)		(0.052)
Downward inconsistent		0.088*	0.005
		(0.041)	(0.048)
Status Differentiation* × *Inconsistency			
Upward differentiation when inconsistency = 1	-0.054		-0.012
	(0.213)		(0.226)
Upward differentiation when inconsistency = 0	0.467*		0.420*
	(0.185)		(0.209)
Downward differentiation when inconsistency = 1		-2.177**	-2.037**
		(0.791)	(0.781)
Downward differentiation when inconsistency = 0		0.958**	0.558*
		(0.305)	(0.225)
Firm-year observations	546	529	546
Firms	59	59	59

Standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 6.6. Additional analyses

	(1) Product line expansion	(2) Product line contraction	(3) Probability of Exit	(4) New technology products	(5) Old technology products
Breadth	0.185**	0.329**	0.003	0.170	0.206**
	(0.018)	(0.081)	(0.203)	(0.147)	(0.021)
Applications	0.043**	-0.003	0.102	0.029	0.048**
	(0.013)	(0.038)	(0.202)	(0.043)	(0.016)
Patents	0.013	0.033	0.118	-0.036+	0.018*
	(0.009)	(0.021)	(0.102)	(0.021)	(0.009)
Status	0.046**	-0.005	0.388**	-0.036	0.037**
	(0.010)	(0.041)	(0.098)	(0.041)	(0.014)
Status differentiation	0.603**	-0.162	2.078		
	(0.222)	(0.229)	(1.467)		
Downward differentiation				2.086+	0.807*
				(1.154)	(0.327)
Upward differentiation				-2.339	0.363*
				(1.789)	(0.185)
Firm-year observations	205	299	208	246	423
Number of firms	40	48	28	19	54

Standard errors in parentheses

+ significant at 10%; * significant at 5%; ** significant at 1%

Chapter 7: Conclusion

Introduction

In the preceding chapters, I have presented a framework for understanding status ambiguity and conflict. In this concluding chapter, I discuss how the framework I have presented relates to existing status theory and other relevant bodies of research.

Significance of conflict for status hierarchies

Much of the literature on status in market settings focuses on the role of status hierarchies in generating inequality. For instance, Podolny (1993) has examined price differentials across investment banks at different positions in the status hierarchy. Podolny and Stuart (1995) have examined differential influence that technological inventions of different status have on subsequent inventions. Podolny, Stuart, and Hannan (1996) have examined differential rates of growth among semiconductor firms of different status. Similarly, other studies have followed this approach and explored the way status hierarchies generate inequalities in a variety of outcomes (e.g. Benjamin and Podolny; Stuart, 1998; Jensen, 2003; 2006). While inequality is certainly an important (and often the most salient) consequence of status hierarchies, status hierarchies involve a rich set of underlying mechanisms, some of which are arguably more fundamental to status as a distinct theoretical concept. Revisiting these more distinctive features of status theory holds promise for extending the status literature, both theoretically and empirically.

Conflict, or more precisely the mutual misjudgment between actors that generate conflict, is important for understanding a status hierarchy's source of stability as a form of role structure. A role structure is a socially-constructed system of distinct profiles of duties and obligations. It is socially-constructed in the sense that it only has substance insofar as the roles comprising it are mutually-recognized by occupants and to the extent that the occupants are prevented from occupying the social space in between roles (White, 1981). If actors were able to claim occupancy of a common role but differ persistently in their conduct, then the role would become meaningless. A role acquires meaning from the common set of expectations placed on occupants of the role. Moreover, a role has meaning because the expectations associated with a role apply categorically. That is, they are invariant to individual-level heterogeneity. For stability to obtain, some internal mechanism must operate to penalize deviance. Status ambiguity represents an attempt to occupy the social spaces forbidden by the role structure. Conflict provides a form of penalty.

Status hierarchy as a source of order without law

As sociologists and legal scholars have argued, it is possible to have order in market settings without the law (Granovetter, 1985; Macaulay, 1963). Sociological research on this issue has largely focused on the role of cohesive, multiplex ties between transaction partners in overcoming problems of hidden information and opportunism (Uzzi, 1997; 1999).

The framework I have presented suggests signaling as an alternative mechanism to trust through which social structure can facilitate order. In the context of settling business disputes, the key problem is that actors must be able to appropriately discriminate in the terms they offer to and accept from one another. The framework I have presented provides a way to bring together insights from the status literature and insights from the law and economics literature. The law and economics literature, inspired by Coase (1960), views disputes as occasions for economic transactions. To the extent that ‘damages’ in a dispute can be priced and exchanged like any other economic good, disputes can be resolved within the market. Market resolution of disputes (as with any market transaction in general) hinges critically on the ability to discriminate between exchange partners with different preferred terms of exchange. The inability to discriminate generates failure to achieve gains from trade.

In addition to serving as an alternative, the framework I have presented may also present a challenge to the ‘embeddedness’ perspective. Contrary to intuition, Gould (2003) observes that violent conflict occurs more often between acquaintances, friends, and relatives than between complete strangers. Similarly, instances of patent litigation often involve firms with prior histories of business dealings with one another. The incidence of conflict between actors with prior ties becomes less counterintuitive when we recognize that actors with extensive prior ties are also more likely to be similar in status and, hence, face status ambiguity. It may be that when compared with complete strangers, actors of similar status are more likely consummate business transactions but are simultaneously more vulnerable to conflict when disputes arise. Current theory provides little guidance for reconciling the potentially countervailing effects of embeddedness and status ambiguity. This provides an avenue for future research.

Market competition versus market conflict

At the broadest level, the aim of this study is to better understand the sources of order in a market. The business strategy literature frequently evokes the imagery of conflict to characterize particularly intense episodes of competition among producers in a market. For instance, research has analyzed producers' market strategies in terms of attack and retaliation moves against competitors (Chen & Miller, 1994; Smith, Grimm, Wally, & Young, 1997). In addition, an extensive literature in the industrial organization tradition has viewed intense competition as a failure in and antithesis to cooperation among producers (McGahan, 1995; Scherer & Ross, 1990). Business strategy research in this tradition has analyzed industry substructure (Caves & Porter, 1977, 1978; McGee & Thomas, 1986) and multimarket contact (Baum & Korn, 1996, 1999; Gimeno, 1999) as determinants of competitive intensity. While evoking the imagery of conflict, the business strategy literature focuses largely on forms of contestation that, however intense, still fall under the umbrella of market competition, in other words, that occur through market mechanisms such as price, quantity, and quality. This emphasis is natural given the assumption in much of the business strategy and related economics literatures that underlying market mechanisms are stable and consistent.

In contrast, research on the sociology of markets suggests that more fundamental forms of conflict are possible when market institutions are contested. Economic exchange is characterized by a diversity of actors, strategies, goods, and circumstances. As a precursor to exchange in the sense of neoclassical economics,—anonymous buyers and sellers interacting through market mechanisms—this diversity must be 'evened out' in a way that seems sensible to all parties (Crawford & Knoer, 1981; Rosen, 1974; Stigler & Sherwin, 1985). This is by no means a straightforward process. Consistent pricing requires buyers and sellers to evaluate goods through common interpretive frames (Zuckerman, 2004). Stable patterns of rivalry require both dominant and challenger firms to share common 'conceptions of control' (Fligstein, 2001). Even the most fundamental of market institutions, such as currency and contracts, involve social construction of shared understandings before they can operate in a consistent way (Carruthers & Babb, 1996; Suchman, 2003). Once in place, these social institutions operate in the background of more visible market institutions to impose order on otherwise ambiguous terms of exchange.

Once we recognize the social processes operating behind market mechanisms, it becomes important to draw a theoretical distinction between orderly, albeit intense, competition occurring within market mechanisms and deeper forms of conflict whose resolution makes orderly competition possible. Recent work in economic sociology has studied contestation over ‘conceptions of control’ (Fligstein, 2001), institutional logics (Haveman & Rao, 1997), and categorical boundaries (Rao, Monin, & Durand, 2005; Rosa, Porac, Runser-Spanjol, & Saxon, 1999). These sorts of conflict center on the presence of some fundamental ambiguity in setting terms of exchange in the market. This leads to divergent beliefs about which producers can rightfully charge some given price, which producers can rightfully offer some given product, or which producers can rightfully serve some class of customer. In each case, resolution of conflict means producers reaching consensus about the answers to the preceding questions, which in turn yields a stable and collectively understood market structure (White, 1981).

References

Agarwal, R., Ganco, M., & Ziedonis, R. H. 2008. Reputations for Toughness in Patent Enforcement: Implications for Knowledge Spillovers via Inventor Mobility. Strategic Management Journal, Forthcoming.

Alster, N. 1988. New profits from patents, Fortune.

Aspray, W. 1997. The Intel 4004 microprocessor: what constituted invention? . IEEE Annals of the History of Computing, 19(3): 4-15.

Baum, J. A. C. & Korn, H. J. 1996. Competitive dynamics of interfirm rivalry. Academy of Management Journal, 39(2): 255-291.

Baum, J. A. C. & Korn, H. J. 1999. Dynamics of dyadic competitive interaction. Strategic Management Journal, 20(3): 251-278.

Bebchuk, L. A. 1984. Litigation and Settlement under Imperfect Information. Rand Journal of Economics, 15(3): 404-415.

Benjamin, B. A. & Podolny, J. M. 1999. Status, quality, and social order in the California wine industry. Administrative Science Quarterly, 44(3): 563-589.

Bessen, J. & Meurer, M. J. 2005. The patent litigation explosion. Boston University School of Law Working Paper.

Bonacich, P. 1987. Power and centrality - a family of measures. American Journal of Sociology, 92(5): 1170-1182.

Bothner, M. S. 2003. Competition and social influence: The diffusion of the sixth-generation processor in the global computer industry. American Journal of Sociology, 108(6): 1175-1210.

Burt, R. S. & Talmud, I. 1993. Market Niche. Social Networks, 15(2): 133-149.

Business Wire. 1992. Advanced Micro Devices, Cypress Semiconductor agree to cross-license patents.

Business Wire. 1993. Brooktree and Advanced Micro Devices settle patent litigation.

Calabresi, G. & Melamed, A. D. 1972. Property rules, liability rules, and inalienability - one view of the cathedral. Harvard Law Review, 85(6): 1089-1182.

Carroll, G. & Hannan, M. T. 1999. The demography of corporations and industries. Princeton, N.J.: Princeton University Press.

Carruthers, B. G. & Babb, S. 1996. The color of money and the nature of value: Greenbacks and gold in postbellum America. American Journal of Sociology, 101(6): 1556-1591.

Caves, R. E. & Porter, M. E. 1977. Entry barriers to mobility barriers - conjectural decisions and contrived deterrence to new competition. Quarterly Journal of Economics, 91(2): 241-261.

Caves, R. E. & Porter, M. E. 1978. Market structure, oligopoly, and stability of market shares. Journal of Industrial Economics, 26(4): 289-313.

Chen, M. J. & Miller, D. 1994. Competitive attack, retaliation and performance - an expectancy-valence framework. Strategic Management Journal, 15(2): 85-102.

Cohen, W. M. & Klepper, S. 1992. The Anatomy of Industry Research-and-Development Intensity Distributions. American Economic Review, 82(4): 773-779.

Cohen, W. M. & Levin, R. C. 1989. Empirical studies of innovation and market structure. In R. Schmalensee & R. D. Willig (Eds.), Handbook of Industrial Organization: Elsevier.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. 2000. Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). NBER Working Paper.

Cooter, R. D. & Rubinfeld, D. L. 1989. Economic analysis of legal disputes and their resolution. Journal of Economic Literature, 27(3): 1067-1097.

Crawford, V. P. & Knoer, E. M. 1981. Job matching with heterogeneous firms and workers. Econometrica, 49(2): 437-450.

Duke, P. J. 1987. Texas Instruments settles patent suits it brought against two Japanese firms, Wall Street Journal.

Dunn, D. 1990. ITC to investigate complaint, Electronic Buyers News.

Dunn, D. 1996. TI wins IC patent reversal, Electronic Buyers News.

Electronic Buyers News. 1990. The patent-suit binge—an era of benign neglect has come to an end.

Electronic Engineering Times. 1988. TI aims DRAM guns at Micron.

Electronic Engineering Times. 1991. AMD-Brooktree spat escalates.

Electronics Times. 1986. US chip firms beginning to rely on patents to ward off competition,

Faggin, F. 1992. The birth of the microprocessor, BYTE.

Fearon, J. D. & Laitin, D. D. 2003. Ethnicity, insurgency, and civil war. American Political Science Review, 97(1): 75-90.

Fisher, L. M. 1990. Chip Makers Combine To Fight Suits, New York Times.

Fligstein, N. 2001. The architecture of markets : an economic sociology of twenty-first-century capitalist societies. Princeton: Princeton University Press.

Galasso, A. & Schankerman, M. 2008. Patent thickets and the market for innovation: evidence from settlement of patent disputes. Working paper.

Gans, J. S., Hsu, D. H., & Stern, S. 2008. The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. Management Science, 54(5): 982-997.

Gardner, W. D. 1990. Microprocessor pioneers never heard of Hyatt: patent grant opens old wounds, Electronic Engineering Times.

Gardner, W. D. 1996. At the same time... Electronic Engineering Times.

Gimeno, J. 1999. Reciprocal threats in multimarket rivalry: Staking out 'spheres of influence' in the US airline industry. Strategic Management Journal, 20(2): 101-128.

Gould, R. V. 1999. Collective violence and group solidarity: Evidence from a feuding society. American Sociological Review, 64(3): 356-380.

Gould, R. V. 2002. The origins of status hierarchies: A formal theory and empirical test. American Journal of Sociology, 107(5): 1143-1178.

Gould, R. V. 2003. Collision of Wills: How Ambiguity about Social Rank Breeds Conflict. Chicago: University of Chicago Press.

Granovetter, M. 1985. Economic-action and social-structure - the problem of embeddedness. American Journal of Sociology, 91(3): 481-510.

Green, J. R. & Scotchmer, S. 1995. On the division of profit in sequential innovation. Rand Journal of Economics, 26(1): 20-33.

Greve, H. R. 1996. Patterns of competition: The diffusion of a market position in radio broadcasting. Administrative Science Quarterly, 41(1): 29-60.

Greve, H. R. 1998. Managerial cognition and the mimetic adoption of market positions: What you see is what you do. Strategic Management Journal, 19(10): 967-988.

Greve, H. R. 2000. Market niche entry decisions: Competition, learning, and strategy in Tokyo banking, 1894-1936. Academy of Management Journal, 43(5): 816-836.

Grindley, P. C. & Teece, D. J. 1997. Managing intellectual capital: Licensing and cross-licensing in semiconductors and electronics. California Management Review, 39(2): 8-&.

Hall, B. H. & Ziedonis, R. H. 2001. The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995. Rand Journal of Economics, 32(1): 101-128.

Hall, B. H., Jaffe, A. B., & Tratjenberg, M. 2001. The NBER patent citation data file: lessons, insights, and methodological tools. NBER Working Paper 8498.

Harbert, T. 1990. Patent portfolios emerge as corporate money makers, Electronic Business Buyer.

Hausman, J., Hall, B. H., & Griliches, Z. 1984. Econometric-Models for Count Data with an Application to the Patents R and D Relationship. Econometrica, 52(4): 909-938.

Haveman, H. A. & Rao, H. 1997. Structuring a theory of moral sentiments: Institutional and organizational coevolution in the early thrift industry. American Journal of Sociology, 102(6): 1606-1651.

Heller, M. A. & Eisenberg, R. S. 1998. Can patents deter innovation? The anticommons in biomedical research. Science, 280(5364): 698-701.

Henry, M. D. & Turner, J. L. 2006. The court of appeals for the federal circuit's impact on patent litigation. Journal of Legal Studies, 35(1): 85-117.

Hsu, G. 2006. Jacks of all trades and masters of none: Audiences' reactions to spanning genres in feature film production. Administrative Science Quarterly, 51(3): 420-450.

Jaffe, A. B. & Lerner, J. 2004. Innovation and its discontents : how our broken patent system is endangering innovation and progress, and what to do about it. Princeton, N.J.: Princeton University Press.

Jensen, M. 2003. The role of network resources in market entry: Commercial banks' entry investment banking, 1991-1997. Administrative Science Quarterly, 48(3): 466-497.

Jensen, M. 2006. Should we stay or should we go? Status accountability anxiety and client defections. Administrative Science Quarterly, 51(1): 97-128.

Katz, M. L. & Shapiro, C. 1985. Network Externalities, Competition, and Compatibility. American Economic Review, 75(3): 424-440.

Kehoe, L. 1984. Japan challenges US micros - leading Japanese semiconductor companies aim to breach American stronghold, Financial Times.

Kehoe, L. 1986. US-Japan chip trade friction grows—semiconductor makers prepare for more legal battles, Financial Times.

Kehoe, L. 1990. Divided when the chips are down: US semiconductor makers are at odds over the industrial policy issue, Financial Times.

Kehoe, L. 1991. Sea of litigation threatening to drown U.S. chip industry Financial Post.

King, G., Tomz, M., & Wittenberg, J. 2000. Making the most of statistical analyses: Improving interpretation and presentation. American Journal of Political Science, 44(2): 347-361.

Kortum, S. & Lerner, J. 1999. What is behind the recent surge in patenting? Research Policy, 28(1): 1-22.

Lanjouw, J. O. & Schankerman, M. 2004. Protecting intellectual property rights: Are small firms handicapped? Journal of Law & Economics, 47(1): 45-74.

Lemley, M. A. & Shapiro, C. 2005. Probabilistic patents. Journal of Economic Perspectives, 19(2): 75-98.

Lemley, M. A. & Shapiro, C. 2007. Patent holdup and royalty stacking. Texas Law Review, 85(7): 1991-2049.

Lincoln, J. R. 1984. Analyzing relations in dyads - problems, models, and an application to interorganizational research. Sociological Methods & Research, 13(1): 45-76.

Macaulay, S. 1963. Non-Contractual Relations in Business - a Preliminary-Study. American Sociological Review, 28(1): 55-67.

Macneil, I. R. 1978. Contracts - Adjustment of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law. Northwestern University Law Review, 72(6): 854-905.

Manners, D. 1996. Lust for life - chip pioneer Dr. Federico Faggin, Electronics Weekly.

McGahan, A. M. 1991. The emergence of the national brewing oligopoly: competition in the American market, 1933-1958. Business History Review, 65(2): 229-284.

McGee, J. & Thomas, H. 1986. Strategic groups - theory, research and taxonomy. Strategic Management Journal, 7(2): 141-160.

Menache, J. 1991. Reasonable royalties, Electronic Buyers News.

Merges, R. P. & Nelson, R. R. 1990. On the Complex Economics of Patent Scope. Columbia Law Review, 90(4): 839-916.

Merges, R. P. 1994. Of property rules, Coase, and intellectual property. Columbia Law Review, 94(8): 2655-2673.

Meurer, M. J. 1989. The settlement of patent litigation. Rand Journal of Economics, 20(1): 77-91.

Moore, G. 1965. Cramming more components onto integrated circuits, Electronics.

Moore, K. A. 2005. Worthless patents. Berkeley Technology Law Journal, 20(4): 1521-1552.

New York Times. 1988. Chip makers in patent suit.

Newsbytes News Network. 1993. National Semi files to bar Mitsubishi devices from US.

Norman, D. 1989. Motorola vs. Hitachi: MPU rights fight—patent suit follows stormy relations, Electronic Buyers News.

Norman, D. 1990. Judge lowers boom on Moto, Hitachi—enjoins 68030, microcontroller, Electronic Buyers News.

Olzak, S. 1992. The Dynamics of Ethnic Competition and Conflict. Stanford: Stanford University Press.

Phillips, D. J. & Zuckerman, E. W. 2001. Middle-status conformity: theoretical restatement and empirical demonstration in two markets. American Journal of Sociology, 107(2): 379-429.

Podolny, J. M. & Morton, F. M. S. 1999. Social status, entry and predation: The case of British shipping cartels 1879-1929. Journal of Industrial Economics, 47(1): 41-67.

Podolny, J. M. & Stuart, T. E. 1995. A role-based ecology of technological change. American Journal of Sociology, 100: 1224-1260.

Podolny, J. M. 1993. A status-based model of market competition. American Journal of Sociology, 98(4): 829-872.

Podolny, J. M. 1994. Market uncertainty and the social character of economic exchange. Administrative Science Quarterly, 39(3): 458-483.

Podolny, J. M. 2001. Networks as the pipes and prisms of the market. American Journal of Sociology, 107(1): 33-60.

Podolny, J. M., Stuart, T. E., & Hannan, M. T. 1996. Networks, knowledge, and niches: Competition in the worldwide semiconductor industry, 1984-1991. American Journal of Sociology, 102(3): 659-689.

Pollack, A. 1984. A move into microprocessors, New York Times.

Pollack, A. 1988. The new high-tech battleground, New York Times.

PR Newswire. 1988. TI files patent infringement suit against Micron Technology.

Rao, H., Monin, P., & Durand, R. 2005. Border crossing: Bricolage and the erosion of categorical boundaries in French gastronomy. American Sociological Review, 70(6): 968-991.

Reinganum, J. F. & Wilde, L. L. 1986. Settlement, litigation, and the allocation of litigation costs. Rand Journal of Economics, 17(4): 557-566.

Resnick, R. 1991. What will the patent war cost your company? , Electronic Business Buyer.

Richards, E. 1990. Firms cash in on patent cases—competitive pressures raise stakes for electronics giants, Washington Post.

Rosa, J. A., Porac, J. F., Runser-Spanjol, J., & Saxon, M. S. 1999. Sociocognitive Dynamics in a Product Market. The Journal of Marketing, 63(Fundamental Issues and Directions for Marketing): 64-77.

Rosen, S. 1974. Hedonic prices and implicit markets - product differentiation in pure competition. Journal of Political Economy, 82(1): 34-55.

Rostky, G. 2002. Remembering the PROM knights of Intel, Electronic Engineering Times: 85-86.

Saul, M. 1996. Samsung, TI trade lawsuits Dispute endangers royalty revenue Dallas Morning News.

Scherer, F. M. & Ross, D. 1990. Industrial market structure and economic performance. Boston: Houghton Mifflin.

Schweizer, U. 1989. Litigation and Settlement under 2-Sided Incomplete Information. Review of Economic Studies, 56(2): 163-178.

Shannon, K. 1988. Texas Instruments files patent suit against fellow Sematech member, Associated Press.

Shapiro, C. 2001. Navigating the patent thicket: cross licenses, patent pools, and standard setting. In A. B. Jaffe & J. Lerner & S. Stern (Eds.), Innovation Policy and the Economy, Vol. 1. Cambridge, MA: MIT Press.

Shepard, A. 1987. Licensing to Enhance Demand for New Technologies. Rand Journal of Economics, 18(3): 360-368.

Sherry, E. F. & Teece, D. J. 2004. Royalties, evolving patent rights, and the value of innovation. Research Policy, 33(2): 179-191.

Simcoe, T. 2007. XTPQML: Stata module to estimate fixed-effects Poisson (Quasi-ML) regression with robust standard errors.

Smith, K. G., Grimm, C. M., Wally, S., & Young, G. 1997. Strategic groups and rivalrous firm behavior: Towards a reconciliation. Strategic Management Journal, 18(2): 149-157.

Spence, A. M. 1974. Market signaling: informational transfer in hiring and related screening processes. Cambridge: Harvard University Press.

Spence, M. 1984. Cost Reduction, Competition, and Industry Performance. Econometrica, 52(1): 101-121.

Steinert-Threlkeld, T. 1989. Steak and chips—quiet dinner kicks off potentially explosive alliance of former competitors, Dallas Morning News.

Steinert-Threlkeld, T. 1990. TI trying to reap royalties without getting black eye Dallas Morning News.

Stigler, G. J. & Sherwin, R. A. 1985. The extent of the market. Journal of Law & Economics, 28(3): 555-585.

Stuart, T. E. 1998. Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry. Administrative Science Quarterly, 43(3): 668-698.

Suchman, M. C. 2003. The contract as social artifact. Law & Society Review, 37(1): 91-142.

Thompson, J. 1988. Changes to patent rules increase industry revenue, Computer System News.

Tilton, J. E. 1971. International Diffusion of Technology: The Case of Semiconductors. Washington: Brookings Institution.

Trajtenberg, M. 1990. A penny for your quotes: patent citations and the value of innovation. Rand Journal of Economics, 21: 172-187.

Uzzi, B. 1997. Social structure and competition in interfirm networks: The paradox of embeddedness. Administrative Science Quarterly, 42(1): 35-67.

Uzzi, B. 1999. Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. American Sociological Review, 64(4): 481-505.

Wade, J. 1995. Dynamics of Organizational Communities and Technological Bandwagons: An Empirical Investigation of Community Evolution in the Microprocessor Market. Strategic Management Journal, 16(Special Issue: Technological Transformation and the New Competitive Landscape): 111-133.

Wade, J. 1996. A community-level analysis of sources and rates of technological variation in the microprocessor market. Academy of Management Journal, 39(5): 1218-1244.

Wade, W. 1997. Samsung to Fujitsu: see you in court, Electronic News.

Wall Street Journal. 1987. Texas Instruments settles patent suit against Hitachi Ltd.

Weber, S. 1991. The patent battle takes an ugly turn, Electronics.

White, H. C. 1981. Where do markets come from? American Journal of Sociology, 87: 517-547.

Whiting, R. 1992. Protecting the power of the idea, Electronic Business Buyer.

Wiegner, K. K. 1987. When the going gets tough, the tough go to court, Forbes.

Willett, H. G. 1990. They're battle weary—but patent suits hurt small firms, Electronic Buyers News.

Wirbel, L. 1989. LTC, National end eight-year battle, Electronic Engineering Times.

Wood, L. 2008. Forgotten PC history: The true origins of the personal computer, Computerworld.

Wooldridge, J. M. 1997. Quasi-likelihood methods for count data. In M. H. Pesaran & P. Schmidt (Eds.), Handbook of Applied Econometrics, Vol. 2: 352-406. Oxford: Blackwell.

Yoder, S. K. 1990. U.S. judge stays injunction of sale of Motorola chips, Wall Street Journal.

Zelner, B. A. 2009. Using simulation to interpret results from logit, probit, and other nonlinear models. Strategic Management Journal, Forthcoming.

Ziedonis, R. H. 2004. Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. Management Science, 50(6): 804-820.

Zuckerman, E. W. 1999. The categorical imperative: Securities analysts and the illegitimacy discount. American Journal of Sociology, 104(5): 1398-1438.

Zuckerman, E. W. 2004. Structural incoherence and stock market activity. American Sociological Review, 69(3): 405-432.

Zuckerman, E. W., Kim, T. Y., Ukanwa, K., & von Rittmann, J. 2003. Robust identities or nonentities? Typecasting in the feature-film labor market. American Journal of Sociology, 108(5): 1018-1074.