ABSTRACT

Title of dissertation: ESSAYS IN CORPORATE FINANCE

Donald E. Bowen III, Doctor of Philosophy, 2017

Dissertation directed by: Professor Michael Faulkender

Department of Finance
Robert H. Smith School of Business

This dissertation is comprised of three essays about investment, technology transfer, and corporate governance mandates.

The first essay, "Patent Acquisition, Investment, and Contracting", examines the transfer of intellectual property via the secondary market for patents and asks how patent acquisitions interact with firm investment policy. I find that patent acquirers subsequently invest in more R&D, increase internal patenting, and eventually make new investments in CAPX. Firms with more technological expertise and investment opportunities acquire more patents. Patent sales are the dominant type of contract and maximize investment incentives; patent licenses frequently contain royalties, which induce underinvestment problems. Nevertheless, licensing can be explained in part by financial and strategic considerations. Licensing is more likely when buyers become financially constrained, when revenue can be shifted to low tax sellers, and when the buyer is a competitor acquiring rights to a valuable patent. Overall, these results suggest patent acquisitions are motivated by the pursuit of in-

vestment synergies, rather than innovation substitution, commercialization motives, or legal threats.

The second essay, “What’s your Identification Strategy? Innovation in Corporate Finance Research”, co-authored with Laurent Frésard and Jérôme P. Taillard, studies the diffusion of techniques designed to identify causal relationships in corporate finance research. We estimate the diffusion started in the mid-nineties, lags twenty years compared to economics, and is now used in the majority of corporate finance articles. Consistent with recent theories of technology diffusion, the adoption varies across researchers based on individuals’ expected net benefits of adoption. Younger scholars, holders of PhDs in economics, and those working at top institutions adopt faster. Adoption is accelerated through networks of colleagues and alumnis and is also facilitated by straddlers who cross-over from economics to finance. Our findings highlight new forces that explain the diffusion of innovation and shape the norms of academic research.

The third essay, “Were non-independent boards really captured before SOX?”, exploits the legal implementation of rules used by the major US stock exchanges following Sarbanes-Oxley (SOX) to study the pre-SOX optimality of board structure. The rules allowed firms to change the legal independence of their board without changing personnel by reclassifying a director from non-independent to independent. Many firms required to change their board structure used reclassification in order to minimize the alterations they made to their pre-SOX board structure, and I call these “placebo firms”. This observation makes feasible a DDD test that identifies the effect of the mandate by comparing treatment firms to placebo firms. Consistent with the view that boards are chosen optimally, real outcomes (profitability) are better for placebo firms than treatment firms. The magnitude of the difference, 4.9 percentage points, is economically meaningful, implying that the constraint is a significant impediment to the conduct of firms targeted by the regulations. Increased profitability is accounted for by increased revenue and typically flat expenses, including investment levels.

ESSAYS IN CORPORATE FINANCE

by

Donald E. Bowen III

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Professor Michael Faulkender, Chair/Advisor
Professor Laurent Frésard
Professor Richmond Mathews
Professor Gerard Hoberg
Professor Erkut Ozbay

© Copyright by
Donald E. Bowen III
2017

Dedication

To my wife, Jordan, for constant support and love.

Acknowledgments

I owe a vast debt to a vast collection of people that have been a part of my life and the journey towards this dissertation.

First and foremost, I'd like to thank my advisor, Professor Michael Faulkender. He was a simply fantastic mentor and friend throughout. His knowledge and perspective on the entire process of research—development, execution, writing, and presenting—tremendously impacted my work. His advice will be embedded in my work forever. I thank him greatly for his guidance and support.

I am also deeply indebted to Professor Laurent Frésard. He was a constant force driving the development of my career with support, guidance, and humorous coffee breaks. This brief acknowledgment section can not appropriately express the scale of the support he gave me or my gratitude for that support.

The other three members of my thesis committee have been excellent to me and deserve special thanks. Professor Gerard Hoberg is exceptionally giving, thoughtful, and encouraging. His support has opened new opportunities for funding, research, and networking. His faith in me has been constant and generous. Professor Richmond Mathews has always offered his time and given me thoughtful and intelligent feedback on issues large and small. Professor Erkut Ozbay is due thanks for generously offering his invaluable time to serve on my committee.

The broader ecosystem around the finance department and Ph.D. program at Maryland were amazing. Professors Mark Loewenstein, Vojislav Maksimovic, and Albert Kyle contributed excellent insights at important and difficult points in the research process. My classmates provided important friendship when it mattered. Justina Blanco is a rock star, impossibly helpful, and a phenomenal person. She is universally acknowledged as central to the functioning and success of the the Ph.D. program at large—but she is still undervalued! My roommates in the early years of the program enriched my life and work.

I owe thanks to some people I met before coming to Maryland. Jeffrey Coles supported my goal of heading to graduate school when I was just persistently knocking on his door. Cliff Smith and Toni Whited shaped my early view of academia and gave me crucial knowledge about the field. Robert Parham and Mihail Velikov are the best people.

Finally, my family deserve the deepest thanks. My parents—all four of them—had ceaseless faith in me and frequently provided support at a moment's notice. Thanking them is an impossible task. My in-laws deserve thanks for being loving and welcoming to me and more importantly, for raising an incredible woman—Jordan, my wife. I'm blessed and dumbstruck that I have her by my side for life. She's fiercely supportive, amazingly sweet, and incomparably caring. This document does not exist without her.

Table of Contents

List of Tables

List of Figures

Chapter 1: Patent Acquisition, Investment, and Contracting

1.1 Introduction

The secondary market for patents—buying and licensing patents—involves over 20% of all patents granted since 1980 and is 5.6 times larger than the number of patents involved in mergers. Despite its transaction volume, the buying and selling of patents has been understudied. This lack of attention is surprising given the voluminous literature analyzing the interactions between financing and intellectual property and the acquisitions of firms with patents. Thus, this paper focuses on the fundamental question of whether firms are in secondary patent markets to substitute or complement for their own investments in research.

For constant guidance and support, I am grateful and indebted to feedback and guidance my chair, Michael Faulkender, and my committee: Laurent Frésard, Jerry Hoberg, and Rich Mathews. For helpful comments, I thank Kenneth Ahern, Cecilia Bustamante, Tom Chang, Francesco D'Acunto, Matt Gustafson, Authur Korteweg, Vojislav Maksimovic, William Mullins, Nagpurnanand Prabhala, Shrihari Santosh, Jerome Taillard, Mihail Velikov, Fernando Zapatero, and seminar participants at the University of Southern California and the University of Maryland.

Existing literature provides ambiguous guidance on whether patent acquirers will subsequently invest more or less in R&D. R&D might decrease for three reasons. First, patent acquisitions could simply substitute for internal innovation. Second, patent acquirers driven by commercialization motives could shift investment away from R&D (Phillips and Zhdanov [2013]). Third, post-acquisition innovation may be stifled by the acquiring firm (Seru [2014] and Gompers, Lerner, and Scharfstein [2005]). Alternatively, R&D might increase if firms acquire patents that have synergies with firm capabilities (Bena and Li [2014] and Hoberg and Phillips [2010]), patents that blocked innovative efforts (Williams [2013]), or patents that allow firms to pursue investment opportunities (Rhodes-Kropf and Robinson [2008]).

Thus, the relationship between patent acquisitions and future R&D is an empirical question. To examine it in a systematic way, I construct a new patent-transaction level dataset that links the names of 1.1 million patent buyers and sellers to Compustat. Transaction data comes from the USPTO's Patent Assignment Dataset (PAD), which covers all US patent assignments for the 1980-2014 period. Matching these patent transactions to firm characteristics overcomes the central challenge that has stunted research on the patent market.

The final firm-linked patent transaction level dataset contains 1.4 million patent sales and licenses and reveals that participation of public firms in the secondary patent market is widespread. In fact, about 64% of all Compustat firms and 38% of firm-years participate in patent transactions. Additionally, innovation among public firms is almost entirely accounted for by years in which firms are active in the secondary patent market: These firm-years account for 92% of all patent grants to public firms and 88% of all Compustat R&D expenditures during the sample period.

Figure 1.1 shows that the patent market is playing an increasingly large role in the economy. The annual fraction of public firms buying or selling patent rights nearly tripled between 1980 and 2010. By 2010, the majority of firms used the patent market each year. The implication of the large, growing, and widely used patent market is that the gap in knowledge regarding the role of patent transactions in firm investment decisions is increasingly important. This paper, by linking market transactions to firm data, represents an initial step towards bridging that gap.

[Insert Figure 1.1 about here]

To evaluate whether patent acquisitions are followed by increases or decreases in R&D, I construct a firm-year panel with variables that capture a firm's yearly activity in the secondary patent market. Specifically, I measure the citation weighted number of patents in which a firm acquires rights via purchase or license, and likewise for patents in which a firm divests rights. I add these measures to a standard investment regression which controls for determinants of investment policy and unobserved heterogeneity at the firm and industry-year level. I find that doubling patent acquisitions in a year from 1.5 to 3 is associated with one year ahead R&D that is 7.9 percentage points higher.[1] This relationship is large: roughly 72% of the

[1] Among firm-years with positive acquisitions, increasing acquisitions from 1.5 to 3 is equivalent to moving from the 30th to the 45th percentile of acquisitions. Among all observations, including those without acquisitions, this move is equivalent to moving from the 75th to the 80th percentile

sample mean R&D rate of 10.7%.

This finding indicates that, on average, patent acquisitions complement future R&D. To support this interpretation, I examine the broader pattern of firm innovation and investment. I find that acquiring firms receive more patents in the following years. These patents build on a wider array of technology fields and earn more citations in the long run. This is inconsistent with the notion that innovation is stifled following acquisitions in the patent market. This is consistent, however, with patent acquisitions and internal R&D being complementary and suggests that acquired patents increase productivity—a possibility I explore in later tests.

To add credence to the main interpretation, I examine physical investment (CAPX) and advertising, which are investments plausibly related to commercialization. Tests show that patent acquisitions do not predict CAPX or advertising spending in the short run. This finding runs counter to the commercialization motive at the heart of a substitution prediction. The lack of comovement among all types of investment also indicates that higher post acquisition R&D is not due to an idiosyncratic shock to all investment opportunities, but is due instead to the patent acquisition itself. Investigating long run relationships, I find that physical investment is significantly higher three and four years after an increase in patent acquisitions. This delayed increase, which occurs simultaneously to a detectable increase in citations, is consistent with R&D projects maturing into operations with a delay.

The overall pattern of innovation, R&D, and physical investment strongly

---

of the distribution.

support the case of a complementary relationship with R&D and conversely, do not support a view of patent acquisitions driven strictly by innovation substitution, commercialization, or legal hold up.[2] Instead, the evidence suggests that patent acquisitions facilitate R&D.

In a second set of tests, I offer evidence that firms with more plausible research and investment synergies acquire more patents, which supports the notion that patent acquisitions facilitate R&D. To do so, I pose several explanations for higher R&D following acquisitions and test whether they determine participation in the patent market. I find that a one standard deviation increase in technological expertise measured by intangible capital stock and patent stock are associated with increases in patent acquisitions of 59% and 25%, respectively. Moreover, a one standard deviation increase in investment opportunities measured by Total Q and sales growth are associated with 4% and 3.3% increases in acquisition activity. These results lend credence to the notion that research complementarities arise from technological synergies, as found by Bena and Li [2014] in the context of mergers.[3] Conversely, I find no evidence that patent acquisitions are related to competitive threats or the amount of patents that peer firms are litigating. The lack of a relationship between patent acquisitions and litigation might be due to seller factors; sellers in litigious industries substantially reduce patent sales. I find that patent sell-

[2] Survey evidence in Feldman and Lemley [2015], concentrating on patent licenses, suggests that patent transactions are a sideshow. The implication of a "sideshow" view of transactions is that patent acquisitions are unrelated to future R&D and CAPX.

[3] See also Hoberg and Phillips [2010] and Rhodes-Kropf and Robinson [2008].

ers have more intellectual capital, as expected, but deteriorating performance in the product and technology markets. This suggests that patent transactions reallocate patent rights to firms more able to exploit them.

In the final set of tests, I ask whether the patent transaction itself is structured to improve investment incentives. I focus on whether patent rights are acquired by purchase or license. The distinction between these two contract types is important, as the choice essentially allocates the ownership and cash flow rights of the patent going forward. In a license, ownership of the patent is retained by the seller but used by the buyer. In a sale, the patent transfers to the new owner following a one-time exchange of cash. Thus, the contract sets firm boundaries, and this setting offers a novel way to understand the dynamics that contribute to organizational form. The primary difference between purchase and license agreements, besides ownership, is that patent licenses often contain royalties calculated as a share of revenue or income. Unlike with leases of physical capital, royalties effectively transfer equity (variable rents) in products based on the patent from the buyer to the seller. Thus, licenses decrease the buyer's incentives to invest and induce underinvestment. Consistent with that notion, I find that patent sales are the dominant type of contract, comprising nearly 98% of transactions.[4]

While sales comprise the bulk of transactions in the data, analysis of the contractual choice reveals that financial considerations justify the use of licensing. I

[4] All contract level analysis is based on the subset of transactions with only one buy-side firm and one sell-side firm. This rules out cross-licenses and bundled sales, which have substantively different contracting dynamics.

specify an empirical contracting model that estimates the likelihood of a transaction being a license agreement and find that licensing is more likely when the buyer is financially constrained, which extends the Eisfeldt and Rampini [2009] finding regarding physical capital leases to patent licensing. I also find that licensing is more likely when the seller is the low tax party. This result can be explained by the use of royalties in order to reduce the combined taxes incurred by the buyer and seller. This allocation tendency—ownership by the low tax party—runs counter to intuition developed in studies of capital leases, wherein the high tax party tends to own equipment because of the use of depreciation tax shields.

Lastly, I show that strategic considerations also justify licensing. Classic contracting models ignore strategic competition and predict that as the productivity of a transacted patent increases, underinvestment following a license agreement is exacerbated. This, in turn, causes the transaction surplus of licensing to decline relative to purchasing. Thus, a key prediction of these models is that licensing is decreasing in patent productivity. Yet, I find the opposite correlation in the data. I measure preexisting citations between acquirers and the acquired patents. Preexisting acquirer citations show that acquirers have existing research capacity directly related to the patent and thus reveal a type of complementarity. This complementarity captures a dimension of buyer-specific value or productivity. I find that as acquirer-patent citation links increase, licensing agreements become more probable.[5]

[5] Moreover, other measures of ex-ante patent value are positively and significantly related to licensing propensity. This relationship holds whether patent value is measured using the stock market impact at grant date (following Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming])

I hypothesize that product market considerations strongly mediate this "puzzling" result. The idea is that, all else equal, sellers prefer to maintain control over valuable patents, but especially when dealing with product market rivals. Indeed, I find that acquirer-patent complementarities are positively related to licensing only among transactions between firms in the same industry. In fact, when a transaction is between firms in different industries, transactions with more acquirer-patent citations are more likely to be sales, as predicted by models that focus on investment optimization, such as Aghion and Tirole [1994].

Altogether, this paper makes three main contributions. First, this paper highlights the importance of the secondary patent market and its relevance for corporate finance. In linking patent market transactions to public firms, I extend the literature on patent transactions by directly examining corporate investment policy. In doing so, I add to the growing innovation-to-investment literature by showing that research investment is higher after acquiring external innovation. As a whole, the findings do not support the view that patent acquisitions are driven strictly by commercialization or innovation substitution, nor do they support the view that patent transactions are a legal sideshow resulting in few products. Rather, patent acquisitions seem to facilitate research efforts. This conclusion extends recent studies in corporate finance showing firms increase innovation after accessing external intellectual property via mergers (Bena and Li [2014]) or corporate venture capital (Ma [2015]) to the largest market for intellectual capital—the secondary patent market.

Second, in examining how firm boundaries are set by the contracting decision,

or pre-transaction citations.

I extend the finance literature on the buy or lease decision for tangible capital to intangible capital.[6] Like tangible capital leases, licensing allows sellers to extend a type of financing. Unlike tangible capital leases, however, the contracting parties to a patent transaction allocate the patents to the low tax party. This allocation rule reduces depreciation tax shields but generates tax savings on the portion of income subject to royalty payments. Thus, the key lesson of the buy or lease literature does not apply directly to patent transactions. This contracting analysis also extends the vertical integration literature by highlighting the interaction between underinvestment costs and strategic competition.[7] Finally, the contractual analysis highlights the strong dominance of contracts which maximize incentives to build on patents and innovation and thus relates to the financial literature on optimal innovation contracting.[8]

Third, a complementary relationship between patent acquisitions and R&D suggests that for the acquirer, the presence of the acquired patent was hindering follow-on research before the transaction. This finding sits at the heart of a debate about the role of patents in facilitating research. Whether patents reduce the rate of cumulative innovation is central to patent policy design and recent research on the topic has been cited in Supreme Court decisions. This literature focuses on how citations (either in patents or scientific publications) respond to the removal

[6] On tangible capital leasing, see, e.g., Smith and Wakeman [1985], Eisfeldt and Rampini [2009], Sharpe and Nguyen [1995], and Rauh and Sufi [2012].

[7] E.g., Aghion and Tirole [1994], Grossman and Hart [1986], and Hart and Moore [1990].

[8] E.g., Acharya, Baghai, and Subramanian [2014], Hackbarth, Mathews, and Robinson [2014], Manso [2011], and Robinson [2008].

of intellectual property rights or surprise patent approvals. I complement these studies by directly examining research spending following patent transfers. This amounts to testing the impact of the removal of patents from competitors. My findings align with recent studies that find evidence that decreased patent litigation risk increases R&D (Mezzanotti [2015]) and increased intellectual property rights depress citations by about 30% (see Murray, Aghion, Dewatripont, Kolev, and Stern [2009] and Williams [2013]). Because these studies examine non-patent forms of protections, Sampat and Williams [2016] instead look at a quasi-exogenous variation in whether patents are granted for DNA sequences and find no evidence that a grant depresses subsequent citations to the same DNA sequence. Their finding can be reconciled with the evidence here if patents depress research by the most likely alternative user (leading to my results) but not firms with lower valuations of the patent (thus diluting the economy-wide impact which is the focus of Sampat and Williams [2016]).

The paper proceeds as follows. Section 1.2 presents the main results on investment following patent acquisitions. Section 1.3 examines the determinants of participation in the patent market. Section 1.4 analyzes the contracting decision. Section 1.5 concludes.

1.2 Patent market activity and investments

This section establishes the complementary relationship between patent acquisitions and R&D. Theories and related evidence on R&D investment following patent acquisitions are discussed in Section 1.2.1. I introduce the main variables based on a new, US-economy wide patent transaction level dataset in Section 1.2.2. Then, I specify and discuss the central test equation in Section 1.2.3 and present the main results in Section 1.2.4. Finally, I discuss the interpretation of the findings in Section 1.2.5.

1.2.1 Theoretical tension and discussion

Efficient markets for tangible goods reallocate resources to their first best use, facilitating investment and growth. While research on the role of markets for physical goods is well developed, research on the market for ideas—especially in corporate finance—is still up and coming.[9] This gap in research is crucial given the role of patents in firm dynamics (such as financing investments and firm growth) and the role of R&D in theories of endogenous growth.[10]

Thus, this paper focuses on whether patent acquisitions complement or substitute for future R&D on average. This question, which focuses on market transactions and not the market itself, matters because it sits at the center of four important discussions. One, it relates to the large debate about the role of patents—either the firm's own patents or those of its competitors—in research and investment decisions.[11] Two, it relates to the discussion about whether the patent market eases

[9] Notable exceptions include Bena and Li [2014], Ma [2015], Serrano [2010], Seru [2014].

[10] On financing, see, for example, Cockburn and MacGarvie [2009], Hochberg, Serrano, and Ziedonis [2014], Hsu and Ziedonis [2008], and Mann [2015]. On endogenous growth, see the classic article Romer [1990].

[11] See Hausman, Hall, and Griliches [1984] and Hall, Griliches, and Hausman [1986].

R&D investment frictions.[12] Three, it relates to studies of vertical integration and particularly to studies focusing on the role of patents and technological know-how in merger decisions.[13] Four, it also relates to broader discussions about the nature of transactions and contracts in this large and growing market.[14] Despite the importance of each of these debates and the large bodies of research that accompany them, there is no study that investigates the role of patent acquisitions in R&D decisions. In lieu of direct evidence, we can look to related literatures for theoretical and empirical guidance about economic mechanisms that would lead to higher or lower R&D following patent acquisitions.

A reasonable starting point is the literature on cumulative innovation, which focuses in particular on how patents impact the R&D decisions others make. Because innovation is often cumulative, a patent held by an inventor might reduce or eliminate follow on research by others if any market friction prevents the transfer of rights between the inventor and others wishing to use the patent. In the absence of such a transaction, litigation risks (e.g. costs due to injunction and infringement) can maintain the patent holder's monopoly use over the patent. If patents do hinder research by others, then the implication in the setting of this paper—assuming firms buy patents that were blocking R&D—would be that patent acquisitions precede R&D increases. Yet, the rich theoretical literature in the area has produced ambiguous predictions (Sampat and Williams [2016]), and the most recent (and well

[12] See Ziedonis [2004].

[13] See Aghion and Tirole [1994], Hitt, Hoskisson, Ireland, and Harrison [1991], and Phillips and Zhdanov [2013].

[14] See Arora, Fosfuri, and Gambardella [2004] and Ahuja and Katila [2001].

identified) empirical efforts disagree about the impact of intellectual property on follow on research.[15] In fact, if patents have positive technological spillovers (a la Bloom, Schankerman, and Van Reenen [2013]), then future patent acquirers might have follow on research in place before the acquisition with the expectation of a patent deal. When its follow on research matures, the project shifts from R&D to commercialization and continued infringement becomes more risky. This prompts the patent acquisition as R&D becomes secondary to commercialization.[16] Thus, this literature does not offer clear theoretical or empirical guidelines for the test I conduct.

Another related and growing literature that involves external patent acquisitions examines the role of patents in mergers. Mergers are more complicated transactions than patent acquisitions, but acquisitions of small technology firms (whose patents constitute a larger share of firm value) might approximate patent purchases. Exploring purchases of small firms, Phillips and Zhdanov [2013] outline a commercialization motive, wherein larger firms avoid an R&D race with smaller firms because they can acquire the smaller firm (and its R&D) and commercialize the target's innovation. If this finding applied to patent acquisitions, a commercialization motive would predict a substitute finding. Another explanation for decreased ex-post innovation comes from Seru [2014]. He finds that acquirers stifle innovation

[15] See Galasso and Schankerman [2014], Murray, Aghion, Dewatripont, Kolev, and Stern [2009], Sampat and Williams [2016], and Williams [2013].

[16] On this point, see Arora [1995], Arora, Fosfuri, and Gambardella [2004], Kieff [2000], Gans, Hsu, and Stern [2002], Gans, Hsu, and Stern [2008].

of target firms and attributes this to the organizational structure of diversified acquirers. However, other papers in this literature offer contrasting implications. A leading example of this is Bena and Li [2014], which finds that mergers are more likely when the target's technology overlaps with the acquirer's. They argue that this represents synergies in technological expertise, and find that such synergies lead to more produced patents. Importantly, they do not directly examine post acquisition R&D choices.

A growing literature on patent transactions questions the nature of patent transactions and the role of non-practicing entities (colliqually, "patent trolls") in the patent market. In a survey of practitioners, Feldman and Lemley [2015] report that patent licenses lead to few new products. In effect, they find that firms paying patent licenses typically continue operations unaltered, even if the license demand comes from a non-troll. The implication of this hypothesis would be no change in R&D post acquisition.

Overall, the ambiguity of preexisting evidence, both theoretical and empirical, highlights the need for this research. The remainder of this section presents the main results. After determining whether the primary relationship is one of complementarity or substitution, this paper will confront the challenge of assessing which channels are likely driving the relationship.

1.2.2 Data and sample

The focus of this paper requires data on secondary patent market transactions. Thus, I begin by constructing a comprehensive set of patent sales and licenses. The USPTO Patent Assignment Dataset (PAD) contains records of 6.3 million patent assignments involving over 10.1 million patents between 1980 and 2014. I overcome two major hurdles to use this dataset. First, there are many reasons for recording a patent assignment.[17] I follow Serrano [2010] and Ma [2015] to define which assignments are sales and which are licenses. This step results in a dataset containing about 350,000 transactions with 1.4 million patents and 1.1 million names of buyers and sellers.[18] Second, I match the listed names of each assignee (buy-side of the transaction) and assignor (sell-side) to US public firm identifiers (GVKEY and PERMNO). This match allows systematic study of how patent transactions are related to firm characteristics.

The appendix describes the construction of the sample in detail, but the most important and novel step is the matching process to link the 1.1 million names of buyers and sellers to Compustat. I standardize the names and then create three auxiliary datasets to link the standardized PAD names to the standardized names of public firms. The first auxiliary dataset exploits the fact that the first assignment

[17] According to Marco, Myers, Graham, DAgostino, and Apple [2015], which documents PAD, reasons for assignment include initial assignments from the inventor-employee to their employing firm, sales, corrections to the record, and security interests.

[18] Transactions can contain multiple patents and some transactions are cross-licensing deals or bundled sales involving multiple buyers and sellers.

of a patent is often not a sale or license, but the assignment from the inventor to the employer firm. In fact, Marco, Myers, Graham, DAgostino, and Apple [2015] reports that these assignments account for over 82% of PAD assignments. For these 5.2 million assignment records, I merge in the firm identifier for the firm that was granted the patent (as listed in the NBER patent data project and Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming]). This produces an exact map from the name of the assignee (the employer) in the PAD data to a firm identifier while bypassing the need to do string matching. Applying this PAD-name-to-firm-identifier map to the 1.1 million buyer and seller names accounts for about 50% of the matches I obtain. The second auxiliary dataset is a comprehensive list of firm name variations compiled by the SEC with which I accept exact name matches. This second step accounts for about 10% of successful matches. Because these steps will fail to match a large number of firm name variations, I build a third auxiliary dataset containing pairs of standardized PAD names and names from the SEC list, along with similarity scores from several fuzzy match functions. Using a random sampling procedure, I pick an algorithm that designates accepted matches with a low false positive rate based on the different similarity scores. This step accounts for the remaining 40% of name to firm identifier matches. Overall, 66% of assignees and 60% of assignors are public firms.

At this stage, the dataset now contains patent sales and licenses along with firm identifiers for buyers and sellers. From this, I construct a firm-year panel with variables that capture activity in the secondary patent market. $PatAcq(CW)$ is the citation weighted sum of patents a firm acquires in a year and $PatDiv(CW)$

is computed similarly for patents a firm divests.[19] I merge these variables with Compustat to create a firm-year sample. All variables are defined in Appendix A.1. Utilities, government, and financial firms (SIC 49, 6, and 9) are excluded from the sample. Since the focus is on R&D expenditures, I restrict the main analysis to firms with active R&D programs as indicated by the Compustat variable XRD.[20]

Table 1.1 shows that participation in secondary patent markets is widespread. Over half of all firms (whether or not the firm has active R&D) acquire a patent at some point in their lifetime and a majority sell patent rights as well. Overall, nearly two-thirds of firms participate in the patent market. Most innovative activity and research in the US economy are done while firms are actively using the market: While firms only buy a patent in 28% of firm-years, these firm-years account for nearly 86% of patents granted to public firms and 81% of inflation adjusted R&D expenditures by public firms during the sample period. Sellers are innovative firms as well, as the 24% of firm-years that include a patent sale account for over 82% of patent grants and 77% of R&D expenses accured by public firms. Only 8% of patents are granted to firms not active in the patent market.[21]

[19] Citations are measured pre-transaction for both variables.

[20] This impact of this restriction is assessed in Table B.4 and discussed in Section 1.2.4.

[21] The overlap between firm-years with transactions and grant reception is apparent at the industry level. Table B.1 reports the fraction of patent acquisitions, divestitures, and patent grants by industry at the SIC2 and SIC3 levels. The broad pattern is industries that receive the most patent grants also buy and sell patent rights most often. High R&D industries are well represented, as most acquisitions are in electronic components, computer equipments, drugs, scientific instruments, and computing services.

[Insert Table 1.1 about here]

Figure 1.1 shows that the patent market is playing an increasingly large role in the economy. Over the entire sample, firms in 38% of their years buy or sell a patent and during that time, participation has nearly tripled. In the early 1980s, roughly a quarter of public firms were active participants in a market where less than 8,000 patents were bought by public firms. By the late 2000s, the annual fraction of firms in the market exceeded 50% and public firms acquired over 25,000 patents annually. Additionally, the patent market has become more involved in the economic life of patents. Patent grants in 2010 are three times more likely to be sold within four years than patents granted in 1980.

Research is increasingly highlighting the importance of patents for firm dynamics.[22] Within that context, the implication of the large, growing, and widely used technology market illustrated by Table 1.1 and Figure 1.1 is that the gap in knowledge regarding the role of patent transactions for firm investment (and other outcomes such as competition, financing policies, and productivity growth) is widening. This paper, by linking market transactions to firm data, represents an initial step towards bridging that gap.[23]

[22] See Hall and Lerner [2010] and Kerr and Nanda [2014] for recent surveys of the large literature on the interaction between patents and financing.

[23] While researchers have used patent sales and other assignments in questions focusing on patent level outcomes or technology areas, this paper represents one of the first efforts to examine firm outcomes. Early contributions based on PAD data include Chesbrough [2006] and Serrano [2010]. More recently, Fischer and Henkel [2012] examines the role of patent trolls in the patent

To ensure that correlation with R&D attributed to patent market purchases isn't confounded with the concurrent economy-wide shift towards obtaining patent grants or firms specific changes, I measure and control for internal innovation. Internal innovation—patents a firm receives directly from the USPTO, rather than acquired in purchases or mergers—is measured in terms of quantity (grants received), quality (citations), and originality (the span of technology fields a patent uses as reflected in the patents it cites).[24] I obtain data on patent grants through 2010 with firm identifiers from Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming][25] and augment this with citations data current through 2014 scraped from Google Patents. In addition to playing a key role as a control variable, how internal innovation changes following patent acquisitions will shed light on the nature and productivity of the research firms undertake.

[Insert Table 1.2 about here]

Table 1.2 describes the main regression sample, which differs from the universe

market, Galasso, Schankerman, and Serrano [2013] estimates gains from trade in the market and shows patent sales reduce litigation risk, and Mann [2015] uses the dataset to explore how patents are used as collateral.

[24] I capture this by defining "Originality" as in Hall, Jaffe, and Trajtenberg [2001], where originality is assumed to capture patents synthesizing a wider variety of technologies into its patents. Specifically, for a patent, originality is one minus the HHI of technology fields the patent cites. Then, for a firm-year, originality is the average of this patent level measure across patent grants received that year.

[25] The dataset can be downloaded at https://kelley.iu.edu/nstoffma/.

of Compustat firms due to the sample requirement that firms have active R&D programs. Relative to all firms, firms in the sample do more R&D (11% to 4%),[26] less physical capital investment (6% to 7%), and have higher averages for all innovation measures. I examine the role of sample selection in Table B.4. I find that, as expected if non-R&D firms are unlikely to pay large fixed costs to begin R&D programs, results become attenuated as the sample includes more firms for which R&D is irrelevant. Nevertheless, the conclusions are robust to different methods of treating missing R&D observations and various sample restrictions based on R&D.

1.2.3 Methodology and Predictions

The economic question is whether patent acquisitions are complementary to or substitute for future R&D. Thus, the main specification is designed to directly estimate this relationship while addressing several complications. Specifically, on the firm-year panel, I estimate

$$\begin{aligned}\text{R\&D}_{i,j,t} = {} & \alpha_1 \times \text{Acquired Patent (Flow)}_{i,t-1} + \alpha_2 \times \text{Divested Patent (Flow)}_{i,t-1} \\ & + \alpha_3 \times \text{Internal Innovation (Flow)}_{i,t-1} + \beta X_{i,t-1} + \eta_i + \eta_{j,t} + u_{i,j,t} \qquad (1.1)\end{aligned}$$

where firm i is in industry j in year t. R&D is normalized by lagged assets. The main variable of interest, Acquired Patent $(\text{Flow})_{i,t-1}$ is defined as $Log(1 + PatAcq(CW))_{i,t-1}$. Divested Patent $(\text{Flow})_{i,t-1}$ is similarly defined as $Log(1+PatDiv(CW))_{i,t-1}$. To ensure that the new measures aren't simply picking up firm specific shifts to-

[26] Compustat averages are not reported in the table. For this statistic, I assume missing R&D equals zero whenever missing.

wards patenting, I control for the flow of internal innovation, defined as $Log(1 + NewPatStock)_{i,t-1}$, where $NewPatStock$ is the sum of new internal grants and citations received by internally generated patents. $X_{i,t-1}$ includes a measure of Q (specifically, *Total Q*) and cash flow as is standard in the investment literature to control for variation in invetment opportunities and cash flow constraints.[27] To prevent firm size from driving coefficient estimates due to the denominator in the dependent variable, $X_{i,t-1}$ also includes $\text{Log(Assets)}_{i,t-1}$. I standardize the controls in $X_{i,t-1}$ (Q, cash flow, and assets) to ease economic interpretation. Standard errors are clustered by firm.

Although the aforementioned theories make strong predictions regarding associations, it is relevant to also explore the challenges that get in the way of causal inference, and the steps I take to mitigate the most direct concerns. First, I tackle two forms of unobserved heterogeneity. To address unobserved heterogeneity due to economic conditions, I include industry-year fixed effects. These controls imply that the coefficients can not be explained by shifts in economic conditions at the economy or industry level, even if the shifts vary by industry over time. For example, a technology shock might increase growth opportunities for all firms in an industry, but this would be absorbed by the industry-year fixed effects. To address unobserved

[27] See Fazzari, Hubbard, Petersen, Blinder, and Poterba [1988] and Erickson and Whited [2000]. *Total Q* is from Peters and Taylor [Forthcoming] and is downloaded from WRDS. Peters and Taylor [Forthcoming] developed *Total Q* to explicitly account for intangible capital and show in a neoclassical framework that it proxies for intangible investment opportunities. Cash flow is measured as in Chang, Dasgupta, Wong, and Yao [2014].

heterogeneity at the firm level, I include firm fixed effects. These controls imply that the coefficients are identified largely by within firm variation. For example, an estimated positive relationship between patent acquisitions and R&D can not be due to stable firm specific factors such as founder CEOs capable of synthesizing and building on new ideas. The combined set of fixed effects therefore accounts for time-invariant unobserved firm heterogeneity and time-varying industry heterogeneity. Second, I explicitly test for pre-trends and find no evidence of that patent acquirers systematically increase R&D prior to transactions. Specifically, firms that are already increasing R&D do not acquire more patents. Third, I relax the one period ahead model to allow for more flexible dynamic relationships. These tests, presented in Tables 1.5 and B.5, show that the conditional increase in R&D the year after acquisitions is robust to dynamic consideration.

While these tests reduce likelihood of bias in the coefficient, I take additional steps to pin down the interpretation of complementarity. The preferred methodology, if feasible, is to exploit exogenous variation in patent acquisitions. This route, however, is challenging. Valid instruments must be associated with acquisitions (relevant) but conditionally unrelated to investment decisions (the "exclusion" condition). Relevant buyer specific factors and other demand factors are highly likely to impact R&D and investment decisions, and thus fail the exclusion condition. Seller factors that increase the supply of patents for purchase must not only be unrelated to the economic conditions that drive buyer investment decisions but also be related to purchasing decisions. Factors that satisfy both conditions are hard to identify.

Instead of an identification approach, I look to rule out alternative stories by examining the broader pattern of investment and innovation following acquisitions. In particular, I examine the evolution of investment in physical assets and advertising and innovation outputs (patent grants, patent citations, patent grant breadth). To do this, I replace the dependent variable R&D with, respectively, CAPX divided by lagged assets, advertising divided by lagged assets, Log(1+Grants)$_{i,j,t}$, Log(1+Cites)$_{i,j,t}$, and Originality$_{i,j,t}$. Each of Grants, Cites, and Originality are based only on internal innovation of a firm. In addition to short run responses, I estimate a dynamic specification to estimate long run relationships. The predictions I discuss and test below are summarized in Table 1.3.

[Insert Table 1.3 about here]

In either case—complement or substitute—I expect the short run response of R&D to persist in the long run. Turning to CAPX and advertising, the commercialization hypothesis would predict immediate and long run increases following patent acquisitions. Alternatively, a complementary relationship would predict either no or negative response in CAPX and advertising in the short run, but a long run increase in CAPX as R&D projects begin to mature. A delay in the onset of increased CAPX and advertising supports the complementarity story, wherein the patent acquisition is an important event, for another reason: If the acquisition and increased R&D are simultaneously driven by an idiosyncratic shocks to investment opportunities not captured by Q, investment of all types is likely to comove.

Changes to output of internal innovation following patent acquisitions also shed light on the nature of R&D following these acquisitions. If acquisitions are complementary to research, perhaps because the patents have synergies with existing but untapped firm resources, changes in R&D policy (both the amount and project selection of R&D) would likely result in more patents and more patent originality, with or without a delay. If, instead, the relationship is substitutive, patent grants and citations might decline as resources are diverted away from research both in the short run and the long run. The diversion of research resources from R&D would likely result in subsequent firm patents leaning on a smaller array of technologies. This prediction is stronger if firms concentrate on incremental patents to protect the product.[28]

To examine the dynamic predictions about firm spending and innovation in the years following a patent acquisition, I specify a distributed lag model

$$\mathrm{y}_{i,j,t} = \sum_{k=1,2,3,4} \Big(\alpha_{1,k} \times \text{Acquired Patent (Flow)}_{i,t-k} + \alpha_{2,k} \times \text{Divested Patent (Flow)}_{i,t-k} \Big)$$
$$+ \alpha_3 \times \text{Internal Innovation (Flow)}_{i,t-1} + \beta X_{i,t-1} + \eta_i + \eta_{j,t} + u_{i,j,t} \tag{1.2}$$

where three additional lags for *Acquired* and *Divested Patents* are included.[29] Coefficients on the first lag of acquired patents corresponds with short run predictions, while the additional lags correspond to the long run predictions. The summary in Table 1.3 includes predictions for coefficients on the additional lags in Equation 1.2.

[28] This notion is in line with Bhide [2000], Gompers, Lerner, and Scharfstein [2005], and Seru [2014], who note that large firms and acquirers can suppress entrepreneurial spirit and ideas.

[29] As in the static model, I standardize Q, cash flow, and assets (included in $X_{i,t-1}$) to facilitate interpretation, and standard errors are clustered by firm.

1.2.4 Main results

Table 1.4 presents the estimation of Equation 1.1. Controlling for determinants of investment policy and unobserved heterogeneity at the firm and industry-year level, patent acquisitions are positively related to next period R&D in a statistically and economically significant way. Specifically, doubling acquisitions (a 100% increase) is associated with 7.9 percentage point increase in R&D investment (which is 72% of the sample mean R&D rate of 10.7%).[30] This sensitivity is large – roughly half of that of the 13.4 percentage point relationship with doubling successful internal innovation ($NewPatStock$). Further consistent with a complementary relationship, firms that acquire patents receive more patents in the following year (column 4) and these patents build on a wider array of technology fields (column 6). Moreover, firms do not spend more on commercializing activities, as there is no significant relationship with CAPX or advertising. None of these findings fit the predictions of the substitution case. Finally, a 100% increase in patent divestitures is related to statistically significant decreases in R&D and CAPX of 10.5 and 8.5 percentage points, respectively, providing additional support for the conclusion.

[Insert Table 1.4 about here]

[30] A 100% increase in acquisitions is a reasonable amount of variation to examine. The mean of the main (logged) acquisition variable is 0.85. Exponentiating it and subtracting one, this corresponds to acquiring 1.5 citation weighted patents in a year. Doubling this from 1.5 to 3, represents a move from the 30th to the 45th percentile of firm-years with positive acquisitions. Among all firm-years, i.e. including years with zero acquisitions, doubling from 1.5 to 3 represents a move from the 75th to the 80th percentile of the distribution.

The estimates of the (standardized) control variables are sensible. Q is significantly and positively related to each type of investment (e.g. Peters and Taylor [Forthcoming]). Cash flow is positively related with CAPX and significantly negatively related to R&D and innovation outputs. Investments are decreasing in firm size (assets) because investments are normalized by assets. Larger firms produce more patents across more fields—a one standard deviation increase in asset size is associated with a 62% increase in patent grants.[31] Following a 100% increase in internal innovation ($NewPatStock$), firms typically make more intangible investments. Such an increase is associated with R&D that is 13 percentage points higher, CAPX that is 18 percentage points lower, and 28% more patents grants.

Next, I estimate the distributed lag model of Equation 1.2 and present results for patent acquisitions and divestments in Table 1.5.[32] Estimations show that innovation output is higher after acquiring patents. A 100% increase in patent acquisitions is associated with 0.6% more citations three years after the acquisition.[33] Such increases are large—they are more than 25% of the sample mean. Patent ac-

[31] The marginal effect is given by $exp(0.483) - 1 = 0.62$.

[32] Results for the dynamic model are robust to including four lags of all independent variables. Because this generates high variance inflation factors due to multicollinearity and does not change the conclusions, I keep the specified model. I also test the above lag model against nested models with fewer lags (e.g. $k = 1, 2, 3$). The results, reported in Table B.5, reject the subcases with one and two lags, but do not reject a three lag model. I keep the four lag structure, as the conclusions match a model with three lags.

[33] Years two and four are positive but insignificant.

quisitions are also related to large increases in CAPX three and four years later, but not immediately. Specifically, a 100% increase in acquisitions is related to percentage point increases in CAPX of 4.2 (t-stat of 2.15) in year three and 4.3 (t-stat of 1.97) in year four (relative to the sample mean CAPX of 6%). Across both tables, advertising generally has no statistical relationship with patent acquisitions. These results are consistent with R&D maturing into patented innovation and production in the long run.

[Insert Table 1.5 about here]

Overall, the complementary view is supported by the pattern of estimates. To further ensure that the central coefficient in the R&D model of Table 1.4 is estimated accurately, additional evidence is reported in Appendix B.2. Table B.2 assesses different functional forms, more general fixed effects, and inclusion of additional patent activity controls. Table B.3 examines different normalizations of R&D. Table B.4 explores the sample selection criteria. Both the statistical and economic conclusions are robust to each of these tests.[34]

[34] One exception is when all firms—including those that disclose zero R&D and those that do not disclose R&D—are included in the sample. The correlation remains economically large but is not statistically significant. The reduced statistical relationship is expected since non-R&D firms are unlikely to pay large fixed costs to begin R&D programs even after acquiring patent rights.

1.2.5 Interpretation

The results strongly conform to expectations in the case of a complementary relationship with R&D and thus do not support a view of patent acquisitions driven strictly by commercialization, a desire to displace imitative R&D (i.e. buying the patent instead of making a similar one), or legal hold up. Moreover, because CAPX does not respond immediately, it is unlikely that the immediate increase in R&D is due to idiosyncratic shocks to investment opportunities not captured by Q.

Instead, the strong but delayed increase in CAPX suggests that the nature of the transactions is related to R&D programs which mature into physical investment and, presumably, products. Higher CAPX in the long run is thus indirect evidence that runs counter to the survey evidence in Feldman and Lemley [2015] suggesting patent transactions are a sideshow. The results here, however, are not directly comparable to Feldman and Lemley [2015], as their results concentrate on non-exclusive licenses used by non-practicing entities (commonly called "patent trolls") to extract payment from as many firms as possible. Patent sales, which constitute the majority of acquisitions in my dataset, are less likely to reflect the type of legal actions that lead to Feldman and Lemley's result. The reasoning for this claim is simple: Sellers can not target other firms for legal action with patents they no longer hold. This distinction contributes to the debate about the efficacy of market transactions by highlighting that contract form matters in assessing the nature of the transactions.

Higher innovation and research investment following patent purchases is supported by research in other contexts where firms obtain access to intellectual capital (e.g. Ma [2015] in the case of venture corporate capital; Bena and Li [2014] in the case of mergers). Relatedly, there is a large literature on cumulative innovation examining whether patent rights hinder related research. Among others, Murray, Aghion, Dewatripont, Kolev, and Stern [2009], Williams [2013], and Galasso and Schankerman [2014] report evidence that intellectual property decreases follow-on work (i.e. citations to patents and scientific publications) by around 30%. Recently, Sampat and Williams [2016] find no evidence of such a reduction. This can be reconciled with the evidence here if patents depress research by the likely eventual buyers but not firms with lower valuations of the patent. Such dispersion in the "hindering" effect is reasonable if firms with valuable patents have limited ability to litigate threats and therefore concentrate on reducing infringement at the firms most likely to benefit from infringement. This economic story could result both in my findings, where buying a patent is related to increases in research, and also the findings in Sampat and Williams [2016], where the overall hindering effect is diluted and near zero across the economy.

Given the consistent findings and corroborative evidence, the question is: What drives the complementary relationship in this setting? The literature on cumulative innovation suggests that the complementaries I find are driven by litigation risk. Long standing economic theory on the nature of R&D suggests another reason litigation risk reduces R&D: Much of R&D is intangible investments in highly skilled employees. If a firm reduces R&D and these employees leave, the firm loses the employees' knowledge capital. To avoid this, managers smooth increases in

R&D spending (Hall, Griliches, and Hausman [1986] and Lach and Schankerman [1989]). Given the intangible nature of R&D spending, the threat of large damages, settlements, and injunctions from lawsuits can exacerbate the reluctance to increase R&D.[35] In buying a patent, a firm eliminates in part or in whole the litigation risk in a particular area and thus reduces the probability of having to reduce R&D in the future.

However, there are other potential channels for this complementary relationship. Buying patent rights might unlock new opportunities to build on and synthesize ideas. Additionally, patents can be useful in lowering the effective tax rate of a firm and thus the marginal cost of capital, which can prompt higher investments in research.[36] The next section looks to examine these channels by estimating the determinants of activity in the patent markets.

1.3 Participation in patent markets

Why do firms buy and sell patent rights? Are the decisions related to factors that could drive the complementary relationship, such as synergies, competition,

[35] Irreversible investments increase the threat posed by injunctions. In NTP v. Research in Motion, the jury awarded NTP $33.5 million in damages. To avoid the possibility of injunction, Research in Motion, the maker of Blackberry mobile phones, settled for $612.5 million—approximately 18 times the jury award (Lemley and Shapiro [2006]).

[36] Dischinger and Riedel [2011], Griffith, Miller, and O'Connell [2014], Karkinsky and Riedel [2012], and Sikka and Willmott [2010] describe and evaluate the tax shifting mechanism. Among others, Hall and Jorgenson [1969], Hall and Van Reenen [2000], Hall and Lerner [2010] and Bloom, Schankerman, and Van Reenen [2013] illustrate the role of taxes in R&D decisions.

litigation, growth opportunities, or financing benefits? This section evaluates the determinants of participation in the secondary patent market and asks how these relationships inform our understanding of the complementary relationship established in the previous section. I also evaluate whether patent acquisitions are just another form of patent generation. Put differently, should we think of the motives in the secondary patent market as distinct from motives in patent generation?

Section 1.3.1 introduces the cross sectional specification and the empirical hypotheses. Section 1.3.2 reports and discusses the findings. Section 1.3.3 examines within firm variation in the decision to buy and sell patents to assess sources of possible selection bias in the main results. Section 1.3.4 discusses the interpretation of the findings.

1.3.1 Designing the cross sectional determinant tests

Which firms are most active in patent markets? Pulling from related literature to select theoretically motivated determinants of patent transactions, I estimate the following specification on the firm-year sample:

$$\begin{aligned} y_{i,j,t} = {} & \beta_1 \times \text{Internal Innovation}_{i,t-1} + \beta_2 \times \text{Competition and Litigation}_{i,t-1} \\ & + \beta_3 \times \text{Growth Opportunities}_{i,t-1} + \beta_4 \times \text{Financial Factors}_{i,t-1} + \eta_{j,t} + u_{i,j,t} \end{aligned} \tag{1.3}$$

The dependent variables $y_{i,j,t}$ are logged measures of patent acquisitions, divestitures, net acquisitions, and internal patents generated.[37] Industry by year fixed

[37] In the table, "PatAcq" stands for $Log(1 + PatAcq(CW))$, "PatDiv" stands for $Log(1 + PatDiv(CW))$, "NetAcq" equals $PatAcq - PatDiv$, and "Grants" stands for $Log(1+FlowGrants)$

effects ($\eta_{j,t}$) absorb industry specific (and economy wide) time trends. Standard errors are clustered by firm.

Each of the independent variables are plausible causes of the complementary relationship. First, I include as controls for *Internal Innovation* the firm's stock of intangible capital (K_{INT}) and the firm's citation-weighted stock of internally generated patents (*PatStock (CW)*). Including these factors, established in the literature as powerful predictors of patenting, serves two purposes. One, the technical expertise embodied in these measures of intellectual capital compose an important source of potential synergies with the acquired patent. Firms with more intellectual capital should thus be more likely to purchase patents. Two, if patent purchases are simply analogues of patent generation, then these variables should be positively related to patent acquisitions with the same sign, significance, and magnitude as they predict patenting. The stock of intellectual capital and patents should also positively predict sales for the simple reason that patents can not be sold without producing patents.

I also include R&D ratio among the proxies for internal innovation. While the stock variables capture the amount of intellectual capital at a firm, the R&D ratio captures firm focus on research and is an established predictor of patenting (e.g. Hausman, Hall, and Griliches [1984] find a strong positive relationship). Thus, it is necessary to include it in the base model in order to compare the cross-sectional and within-firm determinants of patent purchasing and patent generation. The relationship between pre-acquisition R&D intensity and acquisitions can inform two important questions. One, including R&D in within-firm tests (discussed later in

Section 1.3.3 and reported in columns 5-8 of Table 1.6) provides a way to check for pre-treatment trends. Two, including R&D in cross-sectional tests establishes whether buyers, holding intellectual capital levels fixed, are more or less research focused. The relationship between the preexisting R&D intensity of a firm and purchasing decisions is subtly distinct from the main question of this paper (the role of previously acquired patents in R&D) and taps into a larger literature on how the nature of R&D affects the way firms acquire new knowledge. Firms with more focus on R&D might monitor technological opportunities better and have higher capacity to absorb knowledge, or have less need to acquire outside knowledge. Rationales for the former are provided by Arora and Gambardella [1994] and Cohen and Levinthal [1990]. Rationales for the latter are provided by Arora, Belenzon, and Rios [2014], Hitt, Hoskisson, and Ireland [1990], and Katz and Allen [1982].

Second, I assess the role of competitive threats and litigation. If patents owned by other firms hinder valuable research and development efforts, then firms under higher threat levels have larger benefits from acquiring patents and therefore, they should consequently buy more patents. Conversely, firms facing larger threats should sell fewer patents. I include two variables to capture these forces. The first variable, *ProdMktFluidity*, is based on product descriptions in firm filings and captures dynamic shifts in competition, product offerings, and technology (Hoberg, Phillips, and Prabhala [2014]). The second variable, *Log(1+TechPeerLit)*, directly captures litigation in the technology space a firm occupies. This variable increases in the litigiousness of technology peers and, likely, expected litigation risk in a given technology field. I define technology peers as all pairs of firms that cite each other, and then count the number of patents of a firm's peers that appear in district court litigation over the prior five years.[38] Importantly, both variables are constructed to have considerable *within* SIC3 variation and can thus be identified even with the industry-year fixed effects.[39]

Third, I consider proxies for growth opportunities at a firm. Controlling for possible synergies due to the technological expertise of the firm, firms with unexplored growth opportunities should be more (less) likely to acquire (divest) patent rights. I argue that these growth opportunities compose the other dimension of possible synergies between the firm and patent. The three proxies I consider—*Total Q*, *Sales Growth*, and *Patent Growth*—examine different possible dimensions of growth opportunities. Peters and Taylor [Forthcoming] developed *Total Q* to explicitly account for intangible capital and show in a neoclassical framework that it proxies for intangible investment opportunities. *Sales Growth* and *Patent Growth* proxy for growth in the product market and technological environments, respectively, under the assumption growth expectations match recent history.

[38] Data on litigation comes from Henry, McGahee, and Turner [2013]. Most patent litigation does not reach the district court level, and thus the cases that reach the district court level select on both the value of the dispute and the attitudes towards litigation of the parties involved. Therefore, while this mismeasures the raw number of disputes, it captures a large share of the total value of legal disputes in the area.

[39] *ProdMktFluidity* is based on the text based TNIC definition of peers developed in Hoberg and Phillips [2016] wherein peers can change over time (unlike static SIC groups) and thus has large variation both within SIC3 and within firm. *TechPeerLit* is based on technology peers defined by citations, whereas SIC3 peers are defined by products and services.

Fourth, I consider two financial components motivated by evidence from the field and theory. Financial constraints (*DelayCon*) could motivate increased patent sales to raise money (as in Pulvino [1998]).[40] Another potentially powerful motive in this setting is due to the role of patents as a tax management tool.[41] By placing patents in low tax subsidiaries and licensing the use of patents to high tax subsidiaries, income is shifted to the low tax subsidiary and taxes are reduced. This technique is at the heart of controversies including Google, Apple, and others. Thus, firms with higher marginal taxes (*MTR-BCG*, from Blouin, Core, and Guay [2010]) have more to gain from acquiring patents and less to gain from selling patents.

1.3.2 Cross sectional determinants of patent market activity

Columns 1-3 of Table 1.6 present the estimates of the cross sectional model of Equation 1.3. To ease interpretation, all dependent variables are standardized.

[Insert Table 1.6 about here]

Column 1 assesses the drivers of patent acquisitions. I find no evidence that

[40] A relationship between constraints and the amount of patent purchases is unlikely because licensing can be used to facilitate sales to constrained buyers. Indeed, I show in the next section and Eisfeldt and Rampini [2009] shows in a lease vs. buy setting for equipment purchases, more financially constrained firms are able to use licenses and leases acquire the use of assets despite contraints. Thus, more constrained firms can shift the contractual method of acquisition instead of reducing acquisitions.

[41] See Dischinger and Riedel [2011], Griffith, Miller, and O'Connell [2014], Karkinsky and Riedel [2012], and Sikka and Willmott [2010].

competitive threats and financial motives are related to patent purchases. However, firms with more potential synergies (via intellectual capital or growth options) acquire more patents, on average, and these relationships are economically large. Specifically, across firms, a one standard deviation increase in intangible capital stock (K_{INT}) and patent stock are associated with increases in patent purchases of 59% and 25%, respectively.[42] Moreover, one standard deviation increases in *Total Q* and sales growth are associated with 4% and 3.3% increases in acquisition activity.

Controlling for amount of intellectual capital with those two stock variables, firms with higher R&D ratios buy fewer patents. This relationship is driven by firms whose lower R&D to lagged asset ratio is due to having more assets, however, and not due to more R&D per se. In unreported tests, I directly add firm size to the estimation or interact firm size with R&D and find that larger firms acquire more and the relationship with R&D intensity disappears.[43]

Column 2 assess the drivers of patent divestitures and broadly finds support for each set of hypothesized forces. More intellectual capital is related with increased selling (as it is with buying), and again, these relationships are economically large. A one standard deviation increase in intangible capital stock (K_{INT}) and patent

[42] The marginal effects of the coefficients are given by $exp(.464) - 1 = 0.59$ and $exp(.221) - 1 = 0.25$.

[43] Asset size is not explicitly included in the baseline model because of high multicollinearity with intangible stock. Nevertheless, in the unreported specification, all other coefficients maintain their sign, magnitude, and significance. This check motivates and assures the removal of assets from the model and the primary interpretation of the intangible capital variable. (Intangible capital stock is related to firm size, yet its coefficient is little changed by including firm size.)

stock are associated with increases in patent sales of 60% and 47%, respectively.[44] Firms with higher R&D ratios sell fewer patents controlling for amount of intellectual capital. As with column 1, I directly add firm size to the estimation in unreported tests and find other coefficients unaffected. However, in contrast to the firm size test for patent acquisitions, size is unrelated to selling, conditional on the other controls. Multicollinearity issues complicate interpretation of this, but it is worth noting that the negative relationship between R&D ratio and selling remains significant.

While sellers tend to have more intellectual capital to sell, they tend to have deteriorating sales and innovation trajectories. A one standard deviation *decrease* in sales growth and patent growth is associated with increased patent sales of 3% (t=-3.1) and 2% (t=-1.9), respectively. This suggests that patent sales transfer patents to firms more able to exploit the patents.[45]

Litigation risks and taxes are related to patent sales (though not purchasing). A one standard deviation increase in litigation by technology peers is related to 7% fewer patent sales. Increasing tax rates by one standard deviation is related to an 11% reduction in sales, consistent with firms holding on to assets that can reduce taxes. This latter result holds within firm as well (column 6): as a firm's marginal tax rate increases, it becomes less likely to sell.

[44] The marginal effects of the coefficients are given by $exp(.475) - 1 = 0.6$ and $exp(.382) - 1 = 0.47$.

[45] Serrano [2011] estimates large gains from trade in the market for patents. He argues that this is consistent with patent trades reallocating patents to firms more effective at commercializing patents, building on them, or resolving patent disputes outside of courts (as in Galasso, Schankerman, and Serrano [2013]).

Column 3 reports that larger intellectual stocks, specifically patent stocks, make firms 16% (t-stat of 8.4) more likely to sell than to buy. Meanwhile, firms buy more on net as their growth prospects increase. A standard deviation increase in each of the three growth proxies—*Total Q*, *Sales Growth*, and *Patent Growth*—is related to about 5% more net patent purchases, with p-values for each coefficient below 1%. Finally, driven by reductions in selling, firms become net buyers as peer litigation and marginal tax rates increase.

Overall, the cross sectional evidence indicates that patent acquirers appear to have more potential synergies, in that they have more technical expertise and growth options. Acquisition activity is not related to competitive threats or industry patent litigation. The lack of a relationship between patent acquisitions and litigation might be due to seller factors; sellers in litigious industries substantially reduce patent sales. Patent sellers have more intellectual capital, as expected, but deteriorating performance in the product and technology markets.

1.3.3 Within firm determinants of patent market activity

Columns 5-7 of Table 1.6 repeat the estimations of columns 1-3 with the addition of firm fixed effects. Here, I focus on the coefficients whose significance differs from the cross sectional regression.

The largest departure of the within firm specifications is the finding on R&D. Importantly, I find no relationship between lagged R&D and patent acquisitions, sales, or net acquisitions. This implies that the main result (patent acquisitions are related to future R&D expenditures) is not driven by preexisting trends within the firm.

For patent acquisitions (column 5), following the addition of fixed firm characteristics, only intangible capital and sales growth remain statistically significant. This motivates a robustness test for the main results which controls for possible selection effect dues to within firm changes in intangible capital and sales growth that may also be correlated with R&D changes. This test is included in Table B.2 and the main results are unchanged. The difference between the within firm and cross sectional estimates shows that firms with larger patent stocks and higher Q relative to peers buy more patents on average, but increases in patent stock and Q *within a firm* are not related to purchasing.

Meanwhile, other than the coefficient on R&D, the determinants of patent divestitures (column 6) remain largely the same after firm fixed effects are added.[46]

1.3.4 Patent acquisition vs. patent generation

Should we think of acquired patents simply as outsourced patents? Are the motives for patent creation different from buying a patent, and how? This question is related to a large literature in finance, strategy, and economics relating to the management of innovation and research strategies and to the make or buy literature in organizational economics.[47] These literatures describe how the organization

[46] The t-statistic on patent growth, 1.61, is marginally below the cutoff.

[47] E.g., Acharya, Baghai, and Subramanian [2014], Arora, Belenzon, and Rios [2014], Hackbarth, Mathews, and Robinson [2014], Manso [2011], and Robinson [2008].

of firms, including the organization of R&D programs, relates to how firms buy (principally via mergers), make, or ally with others to obtain innovation.

Column 4 of Table 1.6 reports the determinants of patent generation in this sample. As is standard in the literature, I find that patent generation is statistically higher for firms with larger stocks of intellectual capital (proxied by K_{INT} and *PatStock*), firms with more growth opportunities, and firms whose technological area is more contested (proxied by *ProdMktFluidity* and *TechPeerLit*).[48]

Column 4, contrasted against column 1, highlights that, at a basic level, patent creation (column 4) is not driven by the same factors as patent acquisition (column 1). For example, increased technological risks (measured by *ProdMktFluidity* and *TechPeerLit*) are associated with more patents grants, but not more patent purchases. If firms typically respond to litigation and competitive risk by applying for and receiving more patents, why do they not also increase patent acquisitions? Purchasing certainly has advantages relative to applying for a new (but similar) patent; it is faster than the USPTO approval process, the value of a current patent is more certain than a prospective patent, and purchased patents can no longer be used by competitors. This question, in turn, brings up a larger point. The difference in the coefficients implies that either the costs and benefits of patent acquisition differ in important ways from patent creation (thus leading to economically rational differences in the point estimates) or, if patents are simply outsourced patents (and the coefficients "should" be equivalent), that market frictions such as search costs prevent firms from taking the desired purchasing actions. A full examination of this

[48] See Ziedonis [2004], among others, on patenting in contested areas.

point is outside the scope of this paper, but interesting.

More important is an implication based on the estimates in column 3 and column 7 (net patent acquisitions). Currently, virtually all studies that use patent stock as a variable of interest measure it under the assumption that patents are not transferred, largely due to the challenge of identifying the firms in a patent transaction; precisely the challenge this paper overcomes. Yet, approximately 20% of all patents are sold or licensed at some point. Moreover, in my sample, 9% of firm-years that buy a patent have no patent stock under the standard method of accounting for patents! At the firm level, 64% of firms buy or sell a patent. Thus, measurement error in patent stock is pervasive. The implication of column 3 is that this measurement error is not random noise. In principle, this means other variables in a regression (including this table in this paper) with patent stock could be biased in any direction if the residual is correlated with determinants of net patent purchasing.

That 20% of patents are traded during their lifetime prompts the measurement problem, but there is a cheap solution available to researchers to reduce the scope of the problem using standard patent data without explicitly accounting for all transactions. Figure 1.1 shows that about 10% of patents are transacted within the first four years (roughly about 2.5% a year), a number similar to those reported by Serrano [2010] using an earlier version of the dataset. This suggests that a patent stock measure based on recently received patents will have less measurement error. In unreported tests, I replace the patent stock measure with recent grants (e.g. grants received in the prior three or four years) and find similar results.

1.4 Contracting decisions in patent markets

Until this point, the focus has been on investment and innovation *after* a patent transaction and on the characteristics of firms *before* a patent transaction. This section analyzes the transaction itself to shed light on the contractual dynamics at play and the key considerations of the parties involved. Specifically, this section investigates the determinants of the type of the transaction contract—purchase or license. The distinction between these two contract types is important, as the choice essentially allocates ownership of the patent going forward; in a license, the patent is retained by the seller but used by the buyer. Thus, the contract sets firm boundaries, and this setting offers a novel way to understand the dynamics that contribute to organizational form.

The primary difference between purchase and license agreements, besides ownership, is that patent licenses often contain royalties. Royalties are typically a share of revenue or income.[49] Because royalties effectively transfer equity in products based on the patent from the buyer to the seller, licenses decrease the marginal product of investment for the buyer. Given the results of the previous sections—that buyer complementarities are related to acquisitions and acquisitions are related to future investment—and the investment distorting nature of royalties, it is natural to wonder how the contract type is related to buyer complementarities.

The richness of this (unexplored) contracting setting allows us to observe di-

[49] Survey evidence suggests that approximately three quarters of license agreements include some form of royalty.

rect relationships between buyers and the assets involved based on on citation links. Preexisting citations reveal a type of complementarity, in that buyers citing a patent have existing research capacity directly related to the patent. Because the contractual form is the outcome of a negotiation between two parties, this section analyzes the complementarity/ownership relationship within a contracting framework. Section 1.4.2 describes the baseline model, which is estimated in Section 1.4.3. Section 1.4.4 adds a measure of buyer citations to the baseline model and estimates whether the contractual form of acquisition is related to buyer complementarities.

1.4.1 Transaction level sample

Section 1.2.2 described the construction of a patent transaction level dataset with firm identifiers for buyers and sellers. Each transaction is a purchase or a license agreement. This dataset was used in prior sections to construct firm-year variables. In this section, I instead modify the dataset to focus on questions at the *transaction level* itself.

To highlight the forces relevant to the contract choice, I remove transactions with more than one assignor or more than one assignee. These are cross-licenses and bundled sales and including them will introduce many confounding legal and product market forces that drive large scale agreements outside the core focus of this section. Because firm information is required, I keep transactions where either side of deal is a public firm. I discuss these restrictions in more detail later. Additional details on the transaction sample are detailed in Appendix B.1.

[Insert Table 1.7 about here]

Table 1.7 describes this transaction level dataset. Panel A tabulates the number of transactions and patents transfered and highlights the dominance of patent sales as a method of transfer; only 2% of transactions are licenses. Of the 215,972 transactions, 91,262 are between two public firms, 72,970 involve a public buyer, and the remaining 51,740 involve a public seller.

Panel B highlights some differences between the patents involved in sales and licenses. On average, sales involve 0.8 more patents and sold patents are one year younger. These univariate differences are statistically significant. Patents involved in licenses are less cited on average but their grants were associated with larger market impacts. Sellers cite licensed patents 1% less frequently. Notably, technological links between the buyers and the patents are explicit and surprisingly frequent given the young age of patents: 25% of all transactions involve patents cited by the buyers before the transaction. These linkages provide an asset level validation of the notion that buyers enter the market to acquirer complementary assets.

1.4.2 Theoretical discussion and empirical contracting model

The goal of this section is to construct an empirical model that can be estimated and then augmented with a new buyer complementarity factor. While the buy or license decision is an under explored empirical setting, much theoretical literature exists to guide the selection of factors for a baseline contracting model. The question of ownership has antecedents in studies of vertical integration[50] and buy or lease decisions.[51] Both strands of literature start with the assumption that the contracting parties endogenously choose the contract which maximizes overall surplus (subject to individual constraints) and then bargain over the surplus by setting contract terms. The literatures differ somewhat in the contracting factors that receive attention, however, and in this study, the baseline model includes the union of key factors. The vertical integration literature highlights the role of patent productivity, transaction costs, and bargaining power. The financing literature highlights the role of tax differentials and buyer financing constraints.

To capture that richness and analyze what factors drive the license versus sale decision, the baseline specification focuses on cross-sectional variation within a technology area–year. To allow flexible fixed effects, I specify the following linear probability model on the transaction-level sample

$$\begin{aligned}
\text{License}_{c,b,s,p,t} = &\beta_1 \times \text{Patent value}_{p,t} \\
&+\beta_2 \times \text{Tax differential (seller-buyer)}_{b,s,t} \\
&+\beta_3 \times \text{Buyer financial constraints}_{b,t} \\
&+\beta_4 \times \text{Relative licensing transaction costs}_{c} \\
&+\beta_5 \times \text{Bargaining power}_{b,s,t} + \alpha_{p,t} + u_{c,b,s,p,t} \qquad (1.4)
\end{aligned}$$

where c is a contract that transfers some patents p between a buyer b and seller s in year t. Technology area-year fixed effects $\alpha_{p,t}$ absorb variation due to technology

[50] E.g. Grossman and Hart [1986], Aghion and Tirole [1994], and Hart and Moore [1990].

[51] E.g. Eisfeldt and Rampini [2009], Sharpe and Nguyen [1995], and Rauh and Sufi [2012].

specific time trends.[52] To the extent that industries cluster by technology area of patent classification, this specification also absorbs time-varying industry forces.[53] Standard errors are heteroskedastic robust. The appendix describes the construction of the contract level sample.

Because the sample includes private-to-public and public-to-private transactions (in addition to public-to-public), buyer and seller variables (including potential industry fixed effects, firm fixed effects, and firm based clustering) will not have full coverage. I adopt the convention of setting missing values of variables to zero and including dummy variables equal to one if missing to absorb the variation in these observations. As a check of possible collinearity issues induced by this procedure, I examine variance inflation factors and find that they are below standard thresholds across all specifications. Additionally, I perform robustness tests (unreported, but available from the author) on subsamples that do not have missing information and find similar results. These robustness tests address concerns with both the missing variable procedure and possible bias induced by including private firms (which might include patent trolls).

One final issue is the nature of license agreements in the sample. Non-exclusive

[52] Technology class based on the CPC standard used by the USPTO, and there are more than 470 classes. This choice—using technology class instead of industries—has the advantage of being fully defined across observations; nevertheless, I run industry (buyer, or separately, seller) $\times$ year fixed effects in unreported robustness tests and find similar results. For transactions with multiple patents, I pick the classification of the most cited patent.

[53] In robustness tests, I directly include buyer industry $\times$ year fixed effects and find similar results. Using seller industry instead also yields similar results.

licenses have substantively different contracting dynamics, and their impact on the coefficient estimates and the appropriate interpretation is unclear. The imposed sample restriction keeping transactions with only one assignor and one assignee deals with this challenge in part by ruling out obvious cross-licenses and bundled sales. This restriction is conservative, in that some patents might be licensed non-exclusively at different times and remain in the base sample. In unreported tests, I make an aggressive assumption that rules out any transactions whose patents were licensed out by the same party multiple times within five years. This restriction is aggressive, in that some patents might be licensed for short terms. Conclusions remain unchanged.

Next, I discuss the received economic and empirical theory for how each of the baseline variables relate to the likelihood of using a license contract. The first two factors, patent value and tax differential, have theoretically ambiguous predictions. Thus, their average relationship with the contract type is an empirical question with an unknown answer. The latter three factors have unambiguous predictions and testing them empirically provides a sensibility check for this setting.

1.4.2.1 Patent value and contracting

The impact of patent value on contractual form is theoretically ambiguous for the following reasons. Royalties induce investment distortions because they function as a tax on production and investment. Thus under a license contract, the buyer will invest less than they would under a sale agreement (wherein the buyer has full equity and perfectly aligned incentives). Aghion and Tirole [1994] show that this underinvestment increases as the value of the patent increases. The tradeoff is that, for any given level of investment, more valuable patents will generate more licensing income.

Thus, β_3 will be positive (more valuable patents are more likely to be licensed) if the licensing revenue effect dominates the investment distortions and negative if investment distortions dominate. Empirically, I proxy for ex-ante patent value with two variables.[54] The first, *Patent XRET*, is the excess stock return of the firm that receives the initial grant on the day of the patent's announced grant (following Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming]). The second, *Previous Cites*, is based on citations and measured as the logged number of preexisting citations for patents in the deal.[55]

1.4.2.2 Taxes

The classic implication of tax differentials between the buyer and seller in a buy or lease setting is to allocate the asset to the high tax party. This maximizes tax savings, which the parties then bargain over. However, the lessons of those studies need not apply directly to patent ownership because licensing royalties shift the legal incidence of income. This gives the parties the ability to increase tax savings by shifting income to the *low* tax party.

Consider an example where the seller's marginal income is taxed at 10% while

[54] Including them separately or together does not alter any conclusions.

[55] For both of these measures, I choose the maximum across the patents in the deal.

the buyer's marginal income is taxed at 30%. In a lease setting, allocating the asset to the buyer via a sale rather than in a lease means that the tax shields generated by the asset are three times larger. Such tax shields would predict that β_2 is positive: when the seller has higher (lower) taxes, the contract should be a lease (sale), so that the high tax seller (buyer) produces the maximum tax shields.

The picture changes if instead the same parties were considering transferring a patent that was expected to generate \$1M a year in taxable income. Suppose the licensing royalty rate is exogenously set at 50% and based on income. Then, in choosing between a license and a purchase agreement, the parties essentially choose who realizes \$500,000 in profits. By selecting a license, the parties reduce the taxes by \$100,000 annually. This royalty effect would instead predict that β_2 is negative.

Because patents are depreciable, it is reasonable to expect both effects—tax shields and income shifting—to be present in the data. Therefore, a significant relationship between the tax differential of the parties implies that one effect dominates the other. I measure *Tax differential (Seller-Buyer)* using data based on the procedure in Blouin, Core, and Guay [2010].[56]

[56] In robustness tests, I also measure the differential with rates from Graham [1996]. Additionally, for both sets of tax measures, I decompose the differential into the seller and buyer rate. The conclusions are consistent, as seller tax rates have a positive relationship with licensing and buyer tax rates have a negative relationship with licensing. Robustness to this procedure is important because while *Tax differential (Seller-Buyer)* is a direct measure, it only uses variation in the public to public subsample. The decomposed tests are less direct measure, but use variation across more observations. For example, seller tax rates is defined for the public to public *and* public to private subsamples.

1.4.2.3 Financial constraints

Eisfeldt and Rampini [2009] finds that more constrained buyers prefer to lease because leases allow the seller to, in effect, extend a type of financing due to the differential bankruptcy treatment of owned and leased property. As the institutional rules that drive their finding also apply to licensing, I expect that constrained buyers, measured with the *Buyer DelayCon* variable, are more likely to enter into a license than a sale agreement ($\beta_3 > 0$).

In a second test, I decompose financial constraints into equity constraints (*Buyer Eqty DelayCon*) and debt constraints (*Buyer Debt DelayCon*). Hoberg and Maksimovic [2015] report that firms facing equity constraints have "material undisclosed proprietary information" and are constrained from funding growth opportunities. If the finding of Eisfeldt and Rampini [2009] applies to patent contracts, equity constraints are precisely the type of constraint that would spur a firm to use licensing. I thus hypothesize that any relationship between constraints and licensing is concentrated among firms facing equity constraints.

1.4.2.4 Transaction costs

Relative to a sale agreement, licenses mandate monitoring costs by the seller to ensure appropriate payment, prevent infringement, stop invalid sublicensing, and ensure licensees do not sidestep the agreement by acquiring new patents that displace the seller's. As monitoring costs increase, overall surplus of licensing decreases relative to sales, and licensing becomes less likely. These costs should increase as the number of patents in a deal increases, so I proxy for *Relative licensing transaction costs* with the logged number of patents in the transaction (*Patents in Deal*), and expect a negative estimate for β_4.

1.4.2.5 Bargaining power

In contracting models, contract form is typically set by the surplus and then bargaining power either sets terms of the negotiation or price terms. However, firms may have preferences over control of the patent that impact their perceived surplus from a proposed deal. Patent transactions are one such setting where that preference is likely to occur, because patents have increasing returns to scale properties in that average patent values increase with the number of patents in a pool. There are several reasons for this, including library licensing and litigation "war chests." This reasoning implies that firms likely assign value to controlling the patent rights. War chests have value because they affect litigation (firms can tie up opposing parties in more litigation, for example) and it is likely that litigation war chests can be used in negotiations over patent transactions. I thus account for bargaining power of each party with their measured patent stock and hypothesize that licensing is likely when a buyer has more patents and more likely when a seller has more patents.[57]

1.4.3 Empirical determinants of licensing vs. selling

Table 1.8 reports the estimates of Equation 1.4. To ease interpretation, all continuous independent variables are standardized. The dependent variable is scaled by

[57] *Buyer PatStock* and *Seller PatStock* are each citation weighted and in log form.

100, so continuous coefficients should be interpreted as percentage point changes for a one standard deviation change in that independent variable. The economic magnitude of coefficients can be judged relative to the unconditional licensing propensity of 2.1%.

[Insert Table 1.8 about here]

Column 1 estimates the base model, which unanimously supports the unambiguous predictions, both statistically and economically. I find statistically significant evidence that licensing is decreasing in the number of patents in a deal, which proxies for relative transaction cost of licensing (-0.20 percentage points (p.p.), t-stat of 4.7), and *Buyer PatStock*, which proxies for buyer bargaining power (-0.09 p.p., t-stat of 1.8).[58] *Seller PatStock*, which proxies for seller bargaining power, is significantly and positively related to the use of licensing (0.28 p.p., t-stat of 4.0), as is buyer financial constraints (.23 p.p., t-stat of 3.8). The finding on buyer constraints extends the buy or lease findings of Eisfeldt and Rampini [2009] to patent licensing. Decomposing the constraints into equity and debt components in column 2, I find an even stronger association with licensing as firms become more equity constrained, as expected. A one standard deviation increase in equity constraints is related 0.33 p.p. more licensing, which is 16% of the sample mean licensing rate.

The two sets of variables that have theoretically ambiguous predictions—tax differentials and patent value—have statistically significant coefficients empirically.

[58] "p.p" means percentage points.

Consistently, I find that more valuable patents are licensed more often and thus remain the property of their sellers. In fact, both *Patent XRET* and *Previous Cites* are highly significant across all models and in unreported tests where they are included separately. The estimates are robust to separately including seller citations to the patents in the deal.[59] Moreover, the estimates are economically large. A one standard deviation increase in XRET and logged citations is associated with license propensity increases equivalent to 16% and 8% of the sample mean licensing rate. The positive relationship between patent value and licensing indicates that, on average, as transaction patents become more valuable, the increase in licensing income outpaces the value loss due to underinvestment.

Empirically, tax differentials also turn out to be a powerful predictor of the contract form. As the tax rate of sellers increase and buyers decrease (i.e. *Tax Differential (Seller-Buyer)* increases), transactions are more likely to be a sale. Specifically, a one standard deviation increase in the seller's relative marginal tax rate is associated with a 0.56 p.p. (t-stat of 6.96) decrease in the likelihood of a license. This is the largest coefficient in magnitude, equivalent to 27% of the sample mean rate of licensing. This result is robust to decomposing the tax coefficient into separate buyer and seller components or using the tax measure from Graham [1996]. Thus, in the data, ongoing revenue tax considerations dominate depreciation tax

[59] This would matter if *Previous Cites* was loading on citations from sellers. In this case, the interpretation would be that sellers keep patents in which they have conducted follow on research. These patents might be considered core patents sellers have large financial interests in maintaining. The coefficient on *Seller Cites* is significant and positive.

shields on average. This squares with recent evidence from press articles and SEC disclosures that licensing revenues can be substantial.[60]

1.4.4 Acquirer-patent complementarities

Conceptually, acquirer complementarities with patents represent a form of acquirer specific value. Thus, the expectation is that relationship between buyer complementarities and contract type matches that of the common value measures. Hence, a positive relationship is expected.

Columns 3 and 4 of Table 1.8 test this hypothesis. I create two direct measures of acquirer-patent complementarities revealed by citations. In column 3, I include *Log(1+Buyer Cites)*, which counts the number of times the acquirer cited the patent before the transaction. In column 4, I include *Frac. Cites from Buyer*, which is the number of times the acquirer cited the patent divided by overall citations it received. These two variables capture different dimensions of acquirer complementarities. The former indicates the extent to which the buyer has preexisting related research. The latter indicates how specific the asset is to the buyer.

As predicted, both enter positively and significantly; higher acquirer-patent complementarities is associated with a higher likelihood of a license agreement. The coefficients are of the same magnitude as financial constraints. The positive relationship support the interpretation of these complementarities as a buyer specific

[60] IBM and TI were the first firms to break $1 billion in annual royalty revenue. Recently, Qualcomm has generated revenue in excess of $6 billion, while Microsoft and Ericsson earned more than $2 billion a year.

version of value.

Column 5 asks whether this relationship is related to product market competition. To the extent that buyers that place a higher valuation on patents, fully aligned incentives (and thus purchasing) becomes more important. Indeed, column 5 reports that this is the case when firms are not competitors.

1.5 Conclusion

The main message of this paper is that transactions in the secondary patent market are important to understand, with implications for investment policy, patent design, and organizational economics. Over the last three decades, the secondary patent market has grown rapidly to become a pervasive part of the life of a patent and the innovation management of firms. At the same time, the changing nature of the economy, in which more firms can scale production and deliver electronically at near zero cost, makes the reallocation of ideas and patents especially powerful.[61]

These trends make it crucial to advance our understanding of transactions in the market for patents and markets for other forms of intellectual property. This paper takes a step in that direction by linking patent transactions to information on firm investment. At the same time, my findings and conclusions bring up new

[61] Consider the recent example of Niantic, the maker of the game Pokemon Go. Within a year of spinning out of Google in the fall of 2015: (1) Niantic licensed the use of Pokemon characters from Nintendo; (2) Launched a product—a game for smartphones—globally; and (3) Earned approximately \$16 million in sales based on 40 million daily customers in the first month of release. This rapid success was facilitated by the pivotal license with Nintendo.

questions for future research that I plan to explore. First, this paper focuses on patent transactions in isolation. Yet, firms have a menu of contract options that bring new and innovative projects into the firm, including mergers, strategic alliances, and corporate venture capital.[62] How do patent acquisitions fit into the set of options available and what tradeoffs dominate the choice? Second, I highlight the widespread use of patent markets by public firms and corroborate the large fraction of patents that are traded first documented by Serrano [2011]. The unexplored implication of the large fraction of patents that are traded is that patent stocks, as measured in the literature, rely on a false assumption: patents granted to a firm stay with the firm until expiration. Exploring this assumption offers a promising avenue to clarify existing puzzles in the literatures on innovation and investment. Finally, linking patent transactions to firm identifiers makes new research on peer-to-peer technology relations, competition, cooperation, and investment possible.

[62] E.g. Arora, Belenzon, and Rios [2014], Hackbarth, Mathews, and Robinson [2014], Ma [2015], and Robinson [2008].

Appendix

A.1 Variable Definitions

Column 1 contains the name of the variable in the table. Column 2 contains the definition. Unless noted, all-cap variables indicate Compustat mnemonics. Patent citation data comes Google Patent Grants as of 2014. Unless time is explicitly subscripted, variables are defined contemporaneously, although they might be used with a lag. Lags are explicitly noted in tables. Variable transformations, when used, (e.g. log and standardizations) are described in the relevant tables. If noted, variables are winsorized at the 1% level annually.

Table 1.1 – firm year level

Assets	AT/GDPDEF (Normalized to 2009$ via BLS series GDPDEF)
Profits	OIBDP/GDPDEF
Patent Grants	From Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming]
R&D Expenses	XRD/GDPDEF

Table 1.2 – firm year level

$\text{R\&D}_t/\text{AT}_{t-1}$	$\text{XRD}_t/\text{AT}_{t-1}$ (winsorized)
$\text{CAPX}_t/\text{AT}_{t-1}$	$\text{CAPX}_t/\text{AT}_{t-1}$ (winsorized)
$\text{Advertising}_t/\text{AT}_{t-1}$	$\text{XAD}_t/\text{AT}_{t-1}$ (winsorized)
Grants	Flow, from Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming]
Cites	Flow of citations received on internally granted patents
Originality	Originality of citations of new patent grants, computed from Google Patent Grant citations, as defined in Hall, Jaffe, and Trajtenberg [2001]
PatAcq (CW)	Flow of patents acquired (bought or in-licensed), citation weighted pre-transaction
PatDiv (CW)	Flow of patents divested (sold or out-licensed), citation weighted pre-transaction
NewPatStock	Flow of patent grants plus flow of citations (*Grants* + *Cites*)
Total Q	From Peters and Taylor [Forthcoming] (winsorized)
Cash Flow	As defined in Chang, Dasgupta, Wong, and Yao [2014] (winsorized)
Log(Assets)	Log(AT)

Table 1.4 – firm year level

See Table 1.2 for remaining variables

R&D	XRD_t/AT_{t-1} (winsorized)
CAPX	$CAPX_t/AT_{t-1}$ (winsorized)
Advertising	XAD_t/AT_{t-1} (winsorized)
Log(Assets)	Log(AT)

Table 1.5 – firm year level

All variables defined under Table 1.4

Table 1.6 – firm year level

See Table 1.2 for remaining variables

PatAcq	Shorthand for PatAcq (CW) (See Table 1.2)
PatDiv	Shorthand for PatDiv (CW) (See Table 1.2)
NetAcq	PatAcq minus PatDiv
K_{INT}	From Peters and Taylor [Forthcoming] (winsorized before log transformation)
PatStock (CW)	Citation weighted patent stock. Depreciation set to 15%.
R&D	XRD_t/AT_{t-1} (winsorized)
ProdMktFluidity	From Hoberg, Phillips, and Prabhala [2014]
TechPeerLit	Count of patents litigated in years $t-4$ to t among TNIC peer firms. TNIC peer groups from Hoberg and Phillips [2010] and Hoberg, Phillips, and Prabhala [2014]. Litigation data from Henry and Turner [2006] and Henry, McGahee, and Turner [2013].
Sales Growth	($SALE_t$ / $SALE_{t-1}$) -1 (winsorized)
Patent Growth	Log($Grants_t$) - Log($Grants_{t-3}$)
Delaycon	Hoberg and Maksimovic [2015]
MTR-BCG	From Blouin, Core, and Guay [2010]

Table 1.7 – transaction level

Variables set to zero if missing and a variable is included for these observations and zero otherwise.

Patents in Deal	Number of patents in the transaction
Patent Age	Average age of the patents in the transaction
Previous Cites	Total number of preexisting citations received by patents in the deal as of the transaction date
Patent XRET	Borrows from Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming]. For each patent in a deal, obtain the grant firm's XRET (from CRSP) on the patent grant day. If the firm obtained multiple patents,given each equal credit for XRET, i.e. XRET/# patents firm was granted that day. For the transaction, take the maximum of this across patents in the deal.
Any Previous Cites by Buyer	1 if the buyer had cited any patents in the deal by the transaction date
Any Previous Cites by Seller	1 if the seller had cited any patents in the deal by the transaction date

Table 1.8 – transaction level

Variables set to zero if missing and a variable is included for these observations and zero otherwise.

License	1 if license, 0 if sale (see Appendix B.1)
Tax Differential (Seller-Buyer)	Marginal tax rate of seller minus buyer. Data from Blouin, Core, and Guay [2010].
Buyer DelayCon	Buyer's Delaycon (See Table 1.6)
Buyer Eqty DelayCon	Buyer's Eqty Delaycon, from Hoberg and Maksimovic [2015]
Buyer Debt DelayCon	Buyer's Debt Delaycon, from Hoberg and Maksimovic [2015]
Buyer PatStock	Buyer's PatStock (See Table 1.6)
Seller PatStock	Seller's PatStock (See Table 1.6)
Buyer Cites	Number of times the buyer had cited the patents in the deal before the deal
Frac. Cites from Buyer	Number of times the buyer had cited the patents in the deal before the deal divided by total citations the patents had received
Diff. SIC3	1 if the buyer and seller were in a different SIC3 code. Defined for public-public transactions only.

Bibliography

Viral V Acharya, Ramin P Baghai, and Krishnamurthy V Subramanian. Wrongful discharge laws and innovation. *Review of Financial Studies*, 27(1):301–346, 2014.

Philippe Aghion and Jean Tirole. The management of innovation. *Quarterly Journal of Economics*, 109(4):1185–1209, 1994.

Gautam Ahuja and Riitta Katila. Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. *Strategic Management Journal*, 22(3):197–220, 2001.

Ashish Arora. Licensing tacit knowledge: Intellectual property rights and the market for know-how. *Economics of innovation and new technology*, 4(1):41–60, 1995.

Ashish Arora and Alfonso Gambardella. The changing technology of technological change: General and abstract knowledge and the division of innovative labour. *Research Policy*, 23(5):523–532, 1994.

Ashish Arora, Andrea Fosfuri, and Alfonso Gambardella. *Markets for technology: The economics of innovation and corporate strategy.* Boston: MIT press, 2004.

Ashish Arora, Sharon Belenzon, and Luis A Rios. Make, buy, organize: The interplay between research, external knowledge, and firm structure. *Strategic Management Journal*, 35(3):317–337, 2014.

Jan Bena and Kai Li. Corporate innovations and mergers and acquisitions. *Journal of Finance*, 69(5):1923–1960, 2014.

Amar V Bhide. *The origin and evolution of new business.* New York: Oxford University Press, 2000.

Nicholas Bloom, Mark Schankerman, and John Van Reenen. Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4):1347–1393, 2013.

Jennifer Blouin, John E Core, and Wayne Guay. Have the tax benefits of debt been overestimated? *Journal of Financial Economics*, 98(2):195–213, 2010.

Xin Chang, Sudipto Dasgupta, George Wong, and Jiaquan Yao. Cash-flow sensitivities and the allocation of internal cash flow. *Review of Financial Studies*, 27(12): 3628–3657, 2014.

Henry William Chesbrough. *Open innovation: The new imperative for creating and profiting from technology*. Boston: Harvard Business Press, 2006.

Iain M Cockburn and Megan J MacGarvie. Patents, thickets and the financing of early-stage firms: Evidence from the software industry. *Journal of Economics & Management Strategy*, 18(3):729–773, 2009.

Wesley M Cohen and Daniel Levinthal. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1):128–152, 1990.

Matthias Dischinger and Nadine Riedel. Corporate taxes and the location of intangible assets within multinational firms. *Journal of Public Economics*, 95(7): 691–707, 2011.

Andrea L Eisfeldt and Adriano A Rampini. Leasing, ability to repossess, and debt capacity. *Review of Financial Studies*, 22(4):1621–1657, 2009.

Timothy Erickson and Toni M Whited. Measurement error and the relationship between investment and q. *Journal of Political Economy*, 108(5):1027–1057, 2000.

Steven M Fazzari, R Glenn Hubbard, Bruce C Petersen, Alan S Blinder, and James M Poterba. Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, 1988(1):141–206, 1988.

Robin Feldman and Mark Lemley. Does patent licensing mean innovation? *Iowa Law Review*, 101(1):137–139, 2015.

Timo Fischer and Joachim Henkel. Patent trolls on markets for technology–An empirical analysis of NPEs patent acquisitions. *Research Policy*, 41(9):1519–1533, 2012.

Alberto Galasso and Mark Schankerman. Patents and cumulative innovation: Causal evidence from the courts. *Quarterly Journal of Economics*, 130(1):317–369, 2014.

Alberto Galasso, Mark Schankerman, and Carlos J Serrano. Trading and enforcing patent rights. *RAND Journal of Economics*, 44(2):275–312, 2013.

Joshua S Gans, David H Hsu, and Scott Stern. When does start-up innovation spur the gale of creative destruction? *RAND Journal of Economics*, 33(4):571–586, 2002.

Joshua S Gans, David H Hsu, and Scott Stern. The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. *Management Science*, 54(5):982–997, 2008.

Paul Gompers, Josh Lerner, and David Scharfstein. Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *Journal of Finance*, 60(2):577–614, 2005.

John R Graham. Debt and the marginal tax rate. *Journal of Financial Economics*, 41(1):41–73, 1996.

Rachel Griffith, Helen Miller, and Martin O'Connell. Ownership of intellectual property and corporate taxation. *Journal of Public Economics*, 112:12–23, 2014.

Sanford J Grossman and Oliver D Hart. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94(4): 691–719, 1986.

Dirk Hackbarth, Richmond Mathews, and David Robinson. Capital structure, product market dynamics, and the boundaries of the firm. *Management Science*, 60 (12):2971–2993, 2014.

Bronwyn Hall and John Van Reenen. How effective are fiscal incentives for R&D? a review of the evidence. *Research Policy*, 29(4):449–469, 2000.

Bronwyn H Hall and Josh Lerner. The financing of R&D and innovation. In Bronwyn H Hall and Nathan Rosenberg, editors, *Handbook of the Economics of Innovation*, volume 1, pages 609–639. Elsevier, 2010.

Bronwyn H Hall, Zvi Griliches, and Jerry A Hausman. Patents and R&D: Is there a lag? *International Economic Review*, 27(2):265–283, 1986.

Bronwyn H Hall, Adam B Jaffe, and Manuel Trajtenberg. The NBER patent citation data file: Lessons, insights and methodological tools. No. w8498, National Bureau of Economic Research, 2001.

Robert E Hall and Dale W Jorgenson. Tax policy and investment behavior: Reply and further results. *American Economic Review*, 59(3):388–401, 1969.

Oliver Hart and John Moore. Property rights and the nature of the firm. *Journal of Political Economy*, 98(6):1119–1158, 1990.

Jerry Hausman, Bronwyn H Hall, and Zvi Griliches. Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, 52(4): 909–38, 1984.

Matthew Henry, Thomas P McGahee, and John L Turner. Dynamics of patent precedent and enforcement: An introduction to the uga patent litigation datafile. *Available at SSRN 2228103*, 2013.

Matthew D Henry and John L Turner. The court of appeals for the federal circuits impact on patent litigation. *Journal of Legal Studies*, 35(1):85–117, 2006.

Michael A Hitt, Robert E Hoskisson, and R Duane Ireland. Mergers and acquisitions and managerial commitment to innovation in m-form firms. *Strategic Management Journal*, 11(4):29–48, 1990.

Michael A Hitt, Robert E Hoskisson, R Duane Ireland, and Jeffrey S Harrison. Effects of acquisitions on R&D inputs and outputs. *Academy of Management Journal*, 34(3):693–706, 1991.

Gerard Hoberg and Vojislav Maksimovic. Redefining financial constraints: A text-based analysis. *Review of Financial Studies*, 28(5):1312–1352, 2015.

Gerard Hoberg and Gordon Phillips. Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies*, 23 (10):3773–3811, 2010.

Gerard Hoberg and Gordon Phillips. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465, 2016.

Gerard Hoberg, Gordon Phillips, and Nagpurnanand Prabhala. Product market threats, payouts, and financial flexibility. *Journal of Finance*, 69(1):293–324, 2014.

Yael V Hochberg, Carlos J Serrano, and Rosemarie H Ziedonis. Patent collateral, investor commitment, and the market for venture lending. No. w20587, National Bureau of Economic Research, 2014.

David H Hsu and Rosemarie H Ziedonis. Patents as quality signals for entrepreneurial ventures. *Academy of Management Proceedings*, 2008(1):1–6, 2008.

Tom Karkinsky and Nadine Riedel. Corporate taxation and the choice of patent location within multinational firms. *Journal of International Economics*, 88(1): 176–185, 2012.

Ralph Katz and Thomas J Allen. Investigating the not invented here (NIH) syndrome: A look at the performance, tenure, and communication patterns of 50 r & d project groups. *R&D Management*, 12(1):7–20, 1982.

William R Kerr and Ramana Nanda. Financing innovation. No. w20676, National Bureau of Economic Research, 2014.

F Scott Kieff. Property rights and property rules for commercializing inventions. *Minnesota Law Review*, 85:697–754, 2000.

Leonid Kogan, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics*, Forthcoming.

Saul Lach and Mark Schankerman. Dynamics of R&D and investment in the scientific sector. *Journal of Political Economy*, 97(4):880–904, 1989.

Mark A Lemley and Carl Shapiro. Patent holdup and royalty stacking. *Texas Law Review*, 85:1991, 2006.

Song Ma. The life cycle of corporate venture capital. Working paper, Available at SSRN 2691210, 2015.

William Mann. Creditor rights and innovation: Evidence from patent collateral. *Available at SSRN 2356015*, 2015.

Gustavo Manso. Motivating innovation. *Journal of Finance*, 66(5):1823–1860, 2011.

Alan C Marco, Amanda F Myers, Stuart Graham, Paul DAgostino, and Kirsten Apple. The USPTO patent assignment dataset: Descriptions and analysis. USPTO Economics Working Paper, 2015.

Filippo Mezzanotti. Roadblock to innovation: The role of patent litigation in corporate R&D. Working paper, 2015.

Fiona Murray, Philippe Aghion, Mathias Dewatripont, Julian Kolev, and Scott Stern. Of mice and academics: Examining the effect of openness on innovation. No. w14819, National Bureau of Economic Research, 2009.

Ryan H Peters and Lucian A Taylor. Intangible capital and the investment-q relation. *Journal of Financial Economics*, Forthcoming.

Gordon M Phillips and Alexei Zhdanov. R&D and the incentives from merger and acquisition activity. *Review of Financial Studies*, 26(1):34–78, 2013.

Todd C Pulvino. Do asset fire sales exist? An empirical investigation of commercial aircraft transactions. *Journal of Finance*, 53(3):939–978, 1998.

Joshua D Rauh and Amir Sufi. Explaining corporate capital structure: Product markets, leases, and asset similarity. *Review of Finance*, 16(1):115–155, 2012.

Matthew Rhodes-Kropf and David T Robinson. The market for mergers and the boundaries of the firm. *Journal of Finance*, 63(3):1169–1211, 2008.

David T Robinson. Strategic alliances and the boundaries of the firm. *Review of Financial Studies*, 21(2):649–681, 2008.

Paul M Romer. Endogenous technological change. *Journal of Political Economy*, 92(5 pt. 2):71–102, 1990.

Bhaven Sampat and Heidi Williams. How do patents affect follow-on innovation? evidence from the human genome. 2016.

Carlos J Serrano. The dynamics of the transfer and renewal of patents. *RAND Journal of Economics*, 41(4):686–708, 2010.

Carlos J Serrano. Estimating the gains from trade in the market for innovation: Evidence from the transfer of patents. No. w17304, National Bureau of Economic Research, 2011.

Amit Seru. Firm boundaries matter: Evidence from conglomerates and R&D activity. *Journal of Financial Economics*, 111(2):381–405, 2014.

Steven A Sharpe and Hien H Nguyen. Capital market imperfections and the incentive to lease. *Journal of Financial Economics*, 39(2):271–294, 1995.

Prem Sikka and Hugh Willmott. The dark side of transfer pricing: Its role in tax avoidance and wealth retentiveness. *Critical Perspectives on Accounting*, 21(4): 342–356, 2010.

Clifford W Smith and Lee Wakeman. Determinants of corporate leasing policy. *Journal of Finance*, 40(3):895–908, 1985.

Heidi L Williams. Intellectual property rights and innovation: Evidence from the human genome. *Journal of Political Economy*, 121(1):1–27, 2013.

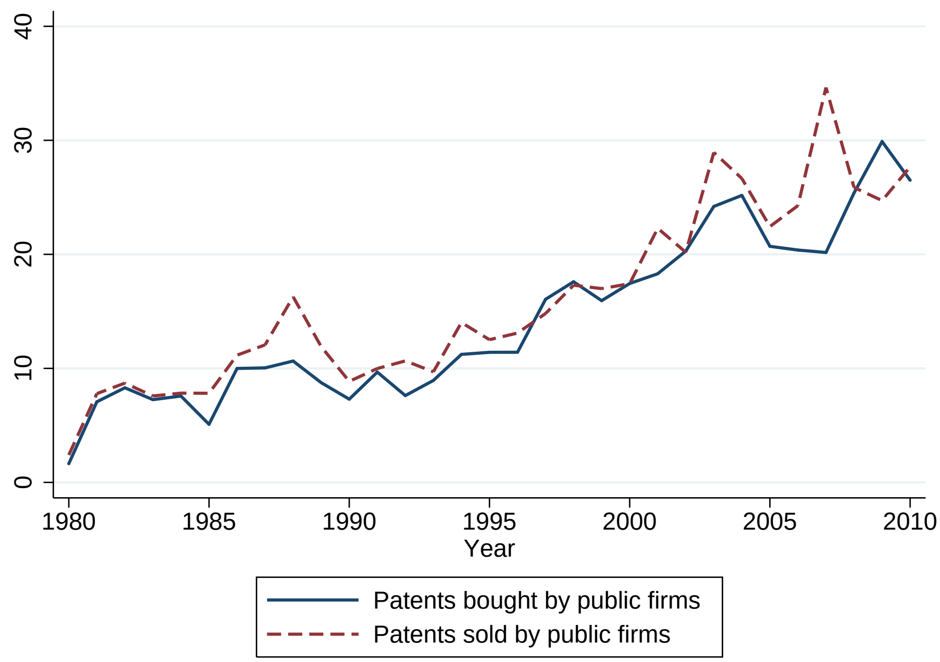
Rosemarie Ham Ziedonis. Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. *Management Science*, 50(6):804–820, 2004.

Figure 1.1: Patent market size

This figure depicts widespread participation in the patent markets. The sample for Panels A and B comprises all Compustat firm-years between 1980 and 2010. I exclude firms in SIC 49, 6, and 9 and firms whose cash flow identity does not hold (see Chang, Dasgupta, Wong, and Yao [2014]). Panel A reports the annual fraction of public firms that buy or sell patents. Panel B reports the total transaction volume, in terms of the number of patents involved, among public firms. The sample for Panel C comprises all US utility patent grants from 1980-2010. For each patent cohort, Panel C reports the fraction of patents that are transacted by 2014 in the solid line and the fraction that are transacted within four years of being granted in the dotted line.

Panel A: Fraction of public firms using patent market

Panel B: Volume of public firm patent market (000s)



Panel C: Fraction of patents sold or licensed, by year of patent grant

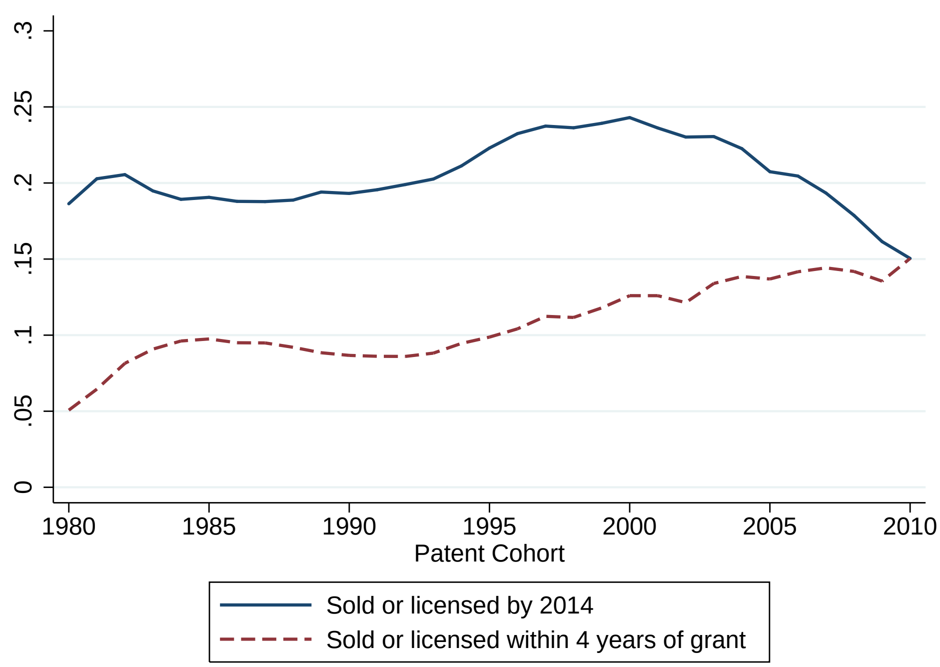


Table 1.1: Fraction of Compustat values in patent transaction years

This table reports widespread participation in the patent markets. The sample comprises all Compustat firm-years between 1980 and 2010. I exclude firms in SIC 49, 6, and 9 and firms whose cash flow identity does not hold (see Chang, Dasgupta, Wong, and Yao [2014]). The firm-life level counts the fraction of public firms that buy patent rights (Column 1), sell patent rights (Column 2), or do either (Column 3) at any point in the sample. Patent right transfers include sale and license agreements. The firm-year level statistics report the fraction of total Compustat values for a given variable that are accounted for by firm-years where the firm buys, sells, or does either. Assets (Compustat variable AT), profits (OIBDP), and R&D (XRD) are inflation adjusted to 2009 values using the GDPDEF series posted on the BLS website.

	Buyer	Seller	Either
Firm-life level:			
Firms	56.0	51.1	63.8
Firm-year level:			
Firm-Years	27.9	24.3	37.7
Assets	58.2	55.1	67.9
Profits	58.5	55.1	68.0
Patent Grants	85.9	82.2	91.6
R&D Expenses	81.4	76.6	87.5

Table 1.2: Summary statistics of firm-year regression sample

This table provides summary statistics for the firm-year sample. The sample comprises Compustat firm-years between 1976 and 2014 with active R&D programs and sufficient data on the control variables. I exclude firms in SIC 49, 6, and 9 and firms whose cash flow identity does not hold (see Chang, Dasgupta, Wong, and Yao [2014]). Data on patent grants come from Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming], citations data from Google Patent Grants, and patent transactions from the USPTO Patent Assignment Dataset. Construction of the sample is described in detail in the Internet Appendix. *R&D, CAPX, Advertising,* and *AT* are Compustat variables XRD, CAPX, XAD, and AT, respectively. *Grants* is the flow of patent grants a firm receives. *Cites* is the flow of citations received by internally granted patents. *Originality* captures the breadth of technologies upon which the new patent grants of a firm rely. A value of 0 means the firms patents only cite one technology field, while a value of 1 means they equally cite all technology fields. *PatAcq (CW)* and *PatDiv (CW)* are the flow of patents acquired and divested, citation weighted on the day before the transaction. *NewPatStock* is new citation weighted patent stock (*Grants* + *Cites*). *Total Q* is from Peters and Taylor [Forthcoming]. *Cash Flow* is cash flow over assets as defined as in Chang, Dasgupta, Wong, and Yao [2014]. The variables defined as ratios (*R&D/AT, CAPX/AT, Advertising/AT, Total Q,* and *Cash Flow*) are winsorized at the 1% level annually.

	N	Mean	SD	P10	P50	P90
Dependent variables:						
$\text{R\&D}_t/\text{AT}_{t-1}$	53,745	0.11	0.15	0.01	0.06	0.25
$\text{CAPX}_t/\text{AT}_{t-1}$	53,745	0.06	0.07	0.01	0.04	0.13
$\text{Advertising}_t/\text{AT}_{t-1}$	53,745	0.01	0.04	0.00	0.00	0.03
Log(1+Grants)_t	53,745	1.21	1.51	0.00	0.69	3.43
Log(1+Cites)_t	53,745	2.18	2.23	0.00	1.61	5.44
Originality_t	53,745	0.31	0.31	0.00	0.30	0.73
Independent variables:						
$\text{Log(1+PatAcq (CW))}_{t-1}$	53,745	0.85	1.50	0.00	0.00	3.22
$\text{Log(1+PatDiv (CW))}_{t-1}$	53,745	0.79	1.51	0.00	0.00	3.14
$\text{Log(1+NewPatStock)}_{t-1}$	53,745	2.11	2.24	0.00	1.61	5.41
Total Q_{t-1}	53,745	1.64	4.03	-0.01	0.65	3.72
Cash Flow_{t-1}	53,745	0.01	0.23	-0.24	0.07	0.18
Log(Assets)_{t-1}	53,745	5.22	2.31	2.45	4.93	8.42

Table 1.3: Predicted relationships with patent acquisitions at time t

This table summarizes the predictions of Section 1.2.3 for Table 1.4 and Table 1.5. The first two columns relate the predictions if patent acquisitions complement future R&D and the last two columns cover the case where patent acquisitions substitute for future R&D.

Case:	If patent acquisitions and R&D are complements		If patent acquisitions and R&D are substitutes	
Time period:	At $t+1$	At $[t+2, t+4]$	At $t+1$	At $[t+2, t+4]$
Coefficients:	$\alpha_{1,1}$	$\{\alpha_{1,2},\, \alpha_{1,3},\, \alpha_{1,4}\}$	$\alpha_{1,1}$	$\{\alpha_{1,2},\, \alpha_{1,3},\, \alpha_{1,4}\}$
Investment types:				
R&D	+	+	−	−
CAPX	0/−	+	+	+
Advertising	0/−	+	+	+
Flow of internal innovation:				
Grants	0/+	+	0/−	−
Cites	0/+	+	0/−	−
Originality	0/+	+	0/−	0/−

Table 1.4: Patent acquisitions, investment, and innovation

This table reports the relationship between patent acquisitions and future investment and internal innovation output. Analysis is based on an OLS estimation of Equation 1.1. The firm-year sample and variable definitions are described in Table 1.2. The dependent variables in the first three columns (under *Investment Ratios*) are normalized by lagged assets and multiplied by 100, so coefficients for those models should be interpreted in terms of percentage points. To facilitate interpretation, *Total Q, Cash Flows,* and *Log(Assets)* are standardized. Industry by year fixed effects absorb industry specific time trends, and firm fixed effects are included. T-statistics are reported in parentheses. Standard errors are clustered by firm. The symbols ***, **, and *, indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Investment Ratios (×100)			Internal Innovation (Flow variables)		
	R&D	CAPX	Advertising	Log(1+Grants)	Log(1+Cites)	Originality
Log(1+PatAcq (CW))$_{t-1}$	0.079***	0.026	-0.009	0.028***	-0.003	0.003***
	(2.69)	(1.39)	(-1.13)	(8.01)	(-1.19)	(3.31)
Log(1+PatDiv (CW))$_{t-1}$	-0.105***	-0.085***	0.015*	-0.003	-0.002	-0.001
	(-3.42)	(-4.27)	(1.84)	(-0.71)	(-0.73)	(-1.32)
Log(1+NewPatStock)$_{t-1}$	0.134**	-0.183***	0.024	0.249***	0.740***	0.012***
	(2.02)	(-4.59)	(1.17)	(25.35)	(116.12)	(8.09)
Total Q$_{t-1}$	0.445***	0.780***	0.104***	-0.013***	-0.032***	-0.007***
	(4.64)	(12.22)	(3.79)	(-3.24)	(-10.01)	(-4.84)
Cash Flow$_{t-1}$	-2.798***	0.369***	-0.005	-0.041***	-0.012***	-0.003
	(-19.28)	(5.91)	(-0.18)	(-7.08)	(-2.72)	(-1.35)
Log(Assets)$_{t-1}$	-9.884***	-3.569***	-0.626***	0.483***	0.199***	0.082***
	(-25.29)	(-17.93)	(-6.82)	(18.49)	(13.15)	(12.21)
Year × SIC3 FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	51,983	51,983	51,983	51,983	51,983	51,983
Adj. R^2	0.74	0.40	0.74	0.81	0.93	0.41

Table 1.5: The dynamic relationship between patent acquisitions, investment, and innovation

This table reports the dynamic relationship between patent acquisitions and future investment and internal innovation output. Analysis is based on an OLS estimation of Equation 1.2. The firm-year sample and variable definitions are described in Table 1.2. The dependent variables in the first three columns (under *Investment Ratios*) are normalized by lagged assets and multiplied by 100, so coefficients for those models should be interpreted in terms of percentage points. Firm-year level controls *NewPatStock, Total Q, Cash Flows,* and *Log(Assets)* are not reported for brevity. Industry by year fixed effects absorb industry specific time trends, and firm fixed effects are included. T-statistics are reported in parentheses. Standard errors are clustered by firm. The symbols ***, **, and *, indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Investment Ratios (×100)			Internal Innovation (Flow variables)		
	R&D	CAPX	Advertising	Log(1+Grants)	Log(1+Cites)	Originality
Log(1+PatAcq (CW))$(t-1)$	0.071**	0.011	-0.010	0.026***	-0.002	0.003***
	(2.44)	(0.60)	(-1.26)	(7.32)	(-0.73)	(3.25)
$(t-2)$	0.004	0.028	0.003	0.021***	0.003	0.002
	(0.14)	(1.48)	(0.32)	(5.75)	(0.88)	(1.58)
$(t-3)$	0.088***	0.042**	-0.001	0.017***	0.006*	0.001
	(2.94)	(2.15)	(-0.11)	(4.87)	(1.91)	(1.38)
$(t-4)$	0.036	0.043**	0.008	0.015***	0.002	0.001
	(1.09)	(1.97)	(0.99)	(3.88)	(0.74)	(0.83)
Log(1+PatDiv (CW))$(t-1)$	-0.101***	-0.091***	0.005	-0.004	0.001	-0.001
	(-3.41)	(-4.79)	(0.70)	(-1.15)	(0.44)	(-0.59)
$(t-2)$	-0.022	-0.040**	0.010	-0.008**	-0.002	-0.001
	(-0.74)	(-2.11)	(1.32)	(-2.25)	(-0.54)	(-1.42)
$(t-3)$	-0.061**	-0.050***	0.007	-0.015***	-0.011***	-0.002*
	(-2.07)	(-2.66)	(0.91)	(-3.77)	(-3.62)	(-1.82)
$(t-4)$	-0.003	-0.036*	0.004	-0.019***	-0.011***	-0.003**
	(-0.08)	(-1.81)	(0.44)	(-4.78)	(-3.32)	(-2.48)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year × SIC3 FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	40,159	40,159	40,159	40,159	40,159	40,159
Adj. R^2	0.75	0.43	0.77	0.83	0.93	0.43

Table 1.6: Determinants of patent market activities

This table presents the OLS estimates on the firm-year sample described in Table 1.2 of Equation 1.3

$$y_{i,j,t} = \beta_1 \times \text{Internal Innovation}_{i,t-1} + \beta_2 \times \text{Competition and Litigation}_{i,t-1} + \beta_3 \times \text{Growth Opportunities}_{i,t-1} + \beta_4 \times \text{Financial Factors}_{i,t-1} + \eta_{j,t} + u_{i,j,t}$$

where $y_{i,j,t}$ is a flow measure of patent market activities or internal patent grants. Patent acquisitions (*PatAcq*) is defined as Log(1+PatAcq (CW)). Patent divestitures (*PatDiv*) is defined as Log(1+PatDiv (CW)). Net patent acquisitions (*NetAcq*) is defined as PatAcq minus PatDiv. Internal patent grants (*Grants*) is defined as Log(1+Flow Grants). *Internal Innovation* includes the stock of intangible capital (K_{INT}), the stock of patents (*PatStock (CW)*), and the R&D ratio. *Competition and Litigation* includes product market fluidity (*ProdMktFluid*) and a measure of litigation among technology peers (*TechPeerLit*). *Growth Opportunities* includes *Total Q, Sales Growth*, and *Patent Growth*. *Financial Factors* includes a measure of financial constraints (*Delaycon*) and the marginal tax rate (*MTR-BCG*). Variables are defined formally in Appendix A.1. To facilitate interpretation, all independent variables are standardized. Industry by year fixed effects ($\eta_{j,t}$) absorb industry specific time trends. T-statistics are reported in parentheses. Standard errors are clustered by firm. The symbols ***, **, and *, indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Cross sectional				Firm fixed effects			
	PatAcq (1)	PatDiv (2)	NetAcq (3)	Grants (4)	PatAcq (5)	PatDiv (6)	NetAcq (7)	Grants (8)
$\text{Log}(1+\text{K}_{INT})_{t-1}$	0.464***	0.475***	-0.011	0.534***	0.385***	0.193***	0.192***	0.656***
	(17.37)	(17.50)	(-0.58)	(23.13)	(6.25)	(2.71)	(3.36)	(15.93)
$\text{Log}(1+\text{PatStock (CW)})_{t-1}$	0.221***	0.382***	-0.161***	0.723***	0.055	0.126***	-0.071**	0.386***
	(8.59)	(14.05)	(-8.41)	(31.46)	(1.52)	(3.47)	(-2.12)	(12.60)
$\text{R\&D}_{t-1}$	-0.031**	-0.039***	0.008	0.061***	-0.008	-0.012	0.004	0.015**
	(-2.11)	(-2.81)	(0.66)	(6.10)	(-0.58)	(-0.86)	(0.23)	(1.96)
$\text{ProdMktFluidity}_{t-1}$	0.007	-0.027	0.035	0.070***	0.036	0.045	-0.008	0.017
	(0.27)	(-0.99)	(1.61)	(3.84)	(1.38)	(1.52)	(-0.30)	(1.05)
$\text{Log}(1+\text{TechPeerLit})_{t-1}$	-0.019	-0.071***	0.052***	0.204***	-0.034	-0.092***	0.057**	0.200***
	(-0.95)	(-3.30)	(3.17)	(14.06)	(-1.31)	(-3.28)	(2.11)	(15.00)
Total Q_{t-1}	0.040***	-0.004	0.043***	0.059***	0.016	0.003	0.012	0.020**
	(2.64)	(-0.27)	(3.59)	(4.83)	(1.10)	(0.27)	(0.81)	(2.32)
Sales Growth_{t-1}	0.033***	-0.028***	0.061***	0.011**	0.020**	-0.029***	0.049***	-0.005
	(3.52)	(-3.12)	(5.66)	(2.24)	(2.17)	(-3.25)	(4.25)	(-1.25)
Patent Growth_{t-1}	0.011	-0.019*	0.031***	0.326***	-0.000	-0.015	0.015	0.264***
	(1.12)	(-1.93)	(2.80)	(38.18)	(-0.05)	(-1.61)	(1.32)	(31.55)
Delaycon_{t-1}	-0.013	-0.021	0.008	0.040***	0.025	0.009	0.016	0.017*
	(-0.64)	(-1.02)	(0.49)	(3.20)	(1.35)	(0.45)	(0.79)	(1.72)
MTR-BCG_{t-1}	-0.016	-0.113***	0.097***	-0.027**	0.010	-0.072***	0.082***	-0.003
	(-0.93)	(-6.65)	(6.45)	(-2.16)	(0.58)	(-3.84)	(3.99)	(-0.29)
Year × SIC3 FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE					Yes	Yes	Yes	Yes
Observations	40,394	40,394	40,394	42,411	40,106	40,106	40,106	42,011
Adj. R^2	0.17	0.21	0.01	0.71	0.33	0.35	0.05	0.84

Table 1.7: Transaction level sample

This table presents summary statistics for the transaction level dataset, where either the buyer or seller is a public firm. The sample contains patent purchases and licenses over 1980-2014. I restrict the sample to deals with only one assignor and one assignee deals to rule out obvious cross-licenses and bundled sales. Construction of the dataset is described in Section 1.4.1 and Appendix B.1. Panel A tabulates the number of transactions based on whether the buyer and/or seller are publicly listed firms. Panel B reports the average of different transaction-level statistics separately for purchases and licenses, the difference of the means, and t-tests for difference of the means. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. *Patents in Deal* is the number of patents in the transaction. *Patent Age* is the average age of the patents in the transaction. *Log(1+Previous Cites)* is the total number of preexisting citations received by patents in the transaction. *Patent XRET* is the maximum (across the patents in the deal) grant day excess return for the firm granted the patent. *Any Previous Cites by Buyer* and *Any Previous Cites by Seller* are equal to one if any of the patents in the deal were previously cited by the buyer and seller, respectively. Variables are defined formally in Appendix A.1.

Panel A: Frequency of transactions by public status of firms

	Purchase	License
Public to Public	89,297	1,965
Private to Public	71,578	1,392
Public to Private	50,805	935
Total	211,680	4,292

Panel B: Transaction level statistics

	Purchase	License	Difference	T-Stat
Patents in Deal	4.19	3.39	-0.80***	3.5
Patent Age	3.86	2.81	-1.05***	13.4
Log(1+Previous Cites)	0.99	0.94	-0.04*	1.9
Patent XRET	0.01	0.02	0.01**	3.6
Any Previous Cites by Buyer	0.25	0.25	0.00	0.1
Any Previous Cites by Seller	0.26	0.25	-0.01**	2.2

Table 1.8: Determinants of licensing

This table presents the estimates of a linear probability model of Equation 1.4,

$$\begin{aligned}\text{License} = \alpha_{FE} + &\beta_1 \times \text{Patent value} + \beta_2 \times \text{Tax differential (seller-buyer)}\\ &+\beta_3 \times \text{Buyer financial constraints} + \beta_4 \times \text{Relative licensing transaction costs}\\ &+\beta_5 \times \text{Bargaining power} + u\end{aligned}$$

where the unit of observation is a patent transaction. The dependent variable is 100 if the contract is a license and 0 if it is a purchase, meaning that all coefficients should be interpreted in percentage point terms. The sample is a contract level database containing patent purchases and licenses over 1980-2014. Construction of the dataset is described in Section 1.4.1 and Appendix B.1. *Patent value* is measured based on information from the stock market (*Patent XRET*) and citations (*Previous Cites*). *Tax differential (seller-buyer)* is measured using data from Blouin, Core, and Guay [2010]. *Buyer financial constraints* is measured using *Buyer DelayCon.* Column 2 decomposes constraints into equity and debt components. *Relative licensing transaction costs* is measured with *Patents in Deal.* *Bargaining Power* is measured with the stock of patents created by the buyer (*Buyer PatStock*) and the seller (*Seller PatStock*). Columns 3 to 5 include direct measures of buyer complementarity with acquired patents. *Buyer Cites* is the number of cites from the buyer to deal patents and *Frac. Cites from Buyer* is *Buyer Cites* divided by *Previous Cites.* Column 5 restricts attention to Public to Public deals. *Diff. SIC3* is an indicator equal to one if the buyer and seller were in a different SIC3. Variables are defined formally in Appendix A.1. To facilitate interpretation, all continuous independent variables are standardized. For each independent variable, missing values are set to zero and a dummy variable equal to one for these missing observations are included. The dummy variables are not included. Variance inflation tests are computed but not reported. Patent technology classification by year fixed effects ($\eta_{j,t}$) are included to absorb technology specific time trends. Heteroskedastic robust standard errors are reported in parentheses. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Predicted sign	Dependent variable: License, ×100 (1)	(2)	(3)	(4)	(5)
Patent XRET	?	0.331***	0.330***	0.295***	0.299***	0.415***
		(3.40)	(3.40)	(3.03)	(3.07)	(2.63)
Log(1+Previous Cites)	?	0.163***	0.163***	0.382***	0.515***	-0.108
		(3.81)	(3.83)	(5.19)	(7.96)	(-1.48)
Tax Diff. (Seller-Buyer)	?	-0.562***	-0.585***	-0.570***	-0.574***	-0.205*
		(-6.97)	(-7.22)	(-7.07)	(-7.11)	(-1.89)
Buyer DelayCon	+	0.234***		0.228***	0.226***	0.321***
		(3.79)		(3.70)	(3.68)	(3.54)
Buyer Eqty DelayCon	+		0.327***			
			(5.05)			
Buyer Debt DelayCon	0		-0.059			
			(-0.92)			
Log(1+Patents in Deal)	-	-0.203***	-0.203***	-0.277***	-0.279***	-0.115*
		(-4.70)	(-4.70)	(-6.18)	(-6.21)	(-1.80)
Log(1+Buyer PatStock)	-	-0.094*	-0.090*	-0.121**	-0.126**	0.008
		(-1.75)	(-1.69)	(-2.25)	(-2.34)	(0.10)
Log(1+Seller PatStock)	+	0.280***	0.285***	0.281***	0.284***	0.172*
		(4.02)	(4.08)	(4.02)	(4.07)	(1.94)
Log(1+Buyer Cites)	?			0.179***		0.459***
				(3.25)		(4.90)
Frac. Cites from Buyer	?				0.213***	
					(4.06)	
Diff. SIC3	?					0.291
						(1.40)
Log(1+Buyer Cites) ×Diff. SIC3	?					-0.497***
						(-4.41)
Year × Pat. Class FE		Yes	Yes	Yes	Yes	Yes
Observations		200,203	200,203	200,203	200,203	83,157
Adj. R^2		0.064	0.065	0.065	0.065	0.102
Avg. License Rate		2.071	2.071	2.071	2.071	2.192

Appendix to “Innovation Acquisition, Investment, and Contracting” for the Internet

This appendix contains additional material not reported in the paper to preserve space. Section B.1 describes the construction of the dataset containing patent transactions. Section B.2 reports additional tests and data.

B.1 Data

Section B.1.1 defines which records in the USPTO Patent Assignment Dataset (PAD) are sales and which are licenses. Section B.1.2 discusses the algorithm to merge COMPUSTAT firm identifiers for buyers and sellers. Section B.1.3 describes the construction of firm year variables from the transaction level data.

B.1.1 Identifying sale and license transactions in PAD dataset

The USPTO Patent Assignment Dataset (PAD) is a newly released set of Stata files, containing the universe of patent assignments reported to the USPTO between 1980 and 2014. The dataset is described in a companion paper (Marco et al. [2015]). The dataset contains about 6.3 million patent assignments. Assignments record changes in ownership, collateral interests, merger transfers, sales, corrections to the record, and most commonly, initial assignments from the inventor-employee to their employing firm.

For researchers, the good news about the data is that patent transactions are usually recorded because assignments are not legally binding unless they are filed with the USPTO. In particular, "By statute, failure to record an assignment in the USPTO renders it void against any subsequent purchaser or mortgagee" (Marco, Myers, Graham, DAgostino, and Apple [2015, p. 6]). Furthermore, Serrano [2010] reports that interviews with patent lawyers support recording patent transfers as best practice. The bad news about the data is that many records are not sales or licenses, and the dataset does not contain links to firm identifiers, such as PERMNO, CUSIP, CIK, or GVKEY. Additionally, transactions are not flagged as sales or licenses explicitly.

Thus, the first crucial task is identifying which of the 6.3 million assignments are sales and licenses. First, I identify transactions that are licensing agreements. These are assignments containing "LICENS" in the assignment description, and whose conveyance type (a PAD variable called *convey_ty*) is "other". I manually examine a random subset of these and find no false positives based on available information. Second, I identify transactions that are sales. Sales are a subset of the are remaining transactions in the dataset whose conveyance type is "assignment", "missing", and "other". However, these conveyance types contain non-sale transactions. In particular, most of these assignments are employee-firm transfers and some are transfers within a firm. The USPTO requires that patent applications filed before Sept 16, 2012 be issued to a human (Marco, Myers, Graham, DAgostino, and

Apple [2015]). Hence, firms have the inventor employee receive the grant and immediately assign the patent to the firm. To remove these non-sales from the list of possible sales, I adapt the strategies of Ma [2015] and Serrano [2010], who use an earlier versions of the PAD dataset. Note that an *assignee* is receiving the patent rights in the assignment from the *assignor*. I standardize the name fields for the assignor and assignee and use the following procedure:[63]

1. I keep transactions whose conveyance type (*convey_ty*) is listed as "assignment", "missing", and "other". I drop any assignments already identified as licenses.

2. PAD contains an indicator variable equal to one if Marco et al. [2015] designated the assignment as an employer assignment based on the assignment's text description. However, they note that they designed this variable in a conservative way, meaning that while the false positive rate is low, the false negative rate (assignments that are employer transfers but not indicated as such) may be high. Thus, I take additional measures to eliminate internal firm transfers.

3. Following Serrano [2010], I drop transactions recorded on the day the patent is granted. These are virtually always employer assignments. However, this

[63] I use *stnd_compname*, a Stata command available at http://www-personal.umich.edu/~nwasi/programs which standardizes strings and is designed specifically to deal with firm names. I augment the program to handle foreign firm names and extraneous address information common to PAD entity names.

fails to account for employee-firm transfers if they file the assignment after the grant day. This prompts two more steps.

4. I drop transactions that meet three criteria. If the transaction (1) covers a single patent and (2) is that patent's first transaction, I check whether (3a) the assignee name corresponds to the patent grant firm (as listed in the USPTO grant dataset, which ignores inventors that immediately transfer the patent) or (3b) the assignor name corresponds to the inventor (as listed in the HBS Patent Inventor Dataset). If all three conditions are satisfied, I drop the transaction. To compare the transaction assignee with the USPTO grant assignee and the transaction assignor with the inventor, I standardize all strings. In this step, I consider (3a) and (3b) as satisfied if the strings are exact matches.

5. I repeat the last step, but allow for small spelling errors. Specifically, I accept fuzzy matches where the Levenshtein distance is less than 10% of the average length of the two strings and where the average length of the two strings is larger than 10.

6. Finally, I drop transactions that are likely internal transfers. These are observations where the standardized assignee and assignor names are exact matches or fuzzy matches, using the same algorithm as before.

The resulting dataset contains approximately 350,000 transactions with 1.4 million patents. Some transactions are bundled sales and cross licenses. I deal with this issue in Section B.1.3 while producing firm-year counts of patent transactions. Next, I obtain firm identifiers for the assignors and assignees.

B.1.2 Merging in firm identifiers

The PAD dataset has string variables containing the names of the assignment parties. To map these to firm identifiers, such as GVKEY, I utilize several sources of information containing firm names and firm identifiers. In each dataset, I standardize company names using the same procedure I use to clean the names in the PAD. Before attempting to merge in GVKEYs, I take the following steps:

1. I build a dataset mapping the *raw* name of the assignee in the PAD dataset to a GVKEY for a range of years. The basic idea is: For the 5.3 million assignments we previously discarded as inventor-firm transactions, we can merge in the firm identity from other common patent data sources. This will produce a large list of PAD names connected to firm identifiers, where even if the firm name is misspelled, without conducting any string matching. This *map* can then be applied to the 350,000 sales and licenses to obtain firm identiers. Call this the ***Direct Grant Map***.

    (a) I select transactions with one patent and coded as employer assignments.

    (b) Optimally, the *Direct Grant Map* will be error free. To reduce possible errors, I keep transactions with only one assignee, and whose patent is only transferred only once during the sample. This is deliberately restrictive.

    (c) I connect the patent number to a PERMNO via the comprehensive firm listing provided by Kogan, Papanikolaou, Seru, and Stoffman [Forthcoming].

   (d) I connect the PERMNO to a GVKEY using a link table provided by WRDS. If a PERMNO has multiple GVKEYs, I pick the one whose data range includes the patent grant. If multiple options remain, discard.

   (e) Standardize firm names and reduce the resulting list to unique assignee name-GVKEY units. For each pair, a name is matched to a GVKEY for a date range that covers inventor-employer assignments.

2. I build a dataset mapping the comprehensive set of firm names in "coleft.cik.c" to a GVKEY for a range of years. Call it the ***Coleft Map***.

   (a) Join in possible GVKEYs for each CIK in "coleft.cik.c" using a link table provided by WRDS.

   (b) Append a name-GVKEY list from the same WRDS link table.

   (c) Standardize firm names, and reduce to unique name-gvkey pairs. For each pair, a name is matched to a GVKEY for a valid date range given by the WRDS the link table.

3. I load all standardized PAD names. For each name, I compare it to all firm names in the *Coleft Map* that begin with the same first there characters. This comparison is done with a large set of fuzzy match functions. Some of these functions have very low false positive rates and some have very low false negative rates. After a random sampling procedure, I select a rule that uses multiple fuzzy match functions to designate which name pairs are considered

matches. The key insight in this process is that some functions have even lower false positive rates conditioning on prior filters. The resulting map designates accepted matches. Call it the ***Scored Map***.

With these maps constructed, I load the dataset containing sales and licenses and merge in GVKEYs for assignees using the following steps, which are repeated for assignors:

1. Using the *Direct Grant Map*, find exact matches and merge in their GVKEY for the transaction year. This is responsible for about 50% of all accepted matches.

2. Using the *Coleft Map*, find exact matches for assignees and merge in their GVKEY for the transaction year. If their GVKEY is still blank, update it. This is responsible for about 10% of all accepted matches.

3. Using the *Scored Map*, find fuzzy matches for assignees and merge in their GVKEY for the transaction year. If their GVKEY is still blank, update it. This is responsible for about 40% of all accepted matches.

The resulting dataset, at the transaction-patent level, contains firm identifiers for the buyers in 66% of transactions and for the sellers in 60% of transactions.

B.1.3 Building firm-year transaction variables

The central tests in the paper are based on a firm year panel. To construct firm-year variables capturing patent market activities, I load the patent transaction dataset. I merge into this dataset the number of pre-transaction citations received by each transfered patent. Then, I count the number of patents and the number of citation weighted patents in which a firm receives rights (as the assignee in the transaction) in a given year.[64] I call these variables *PatAcq* and *PatAcq (CW)*, respectively. Repeating this for assignors results in *PatDiv* and *PatDiv (CW)*.

B.1.3.1 A small issue

The above step contains a few checks to ensure firms get appropriate credit for patent right transactions. The main issue is that this dataset contains bundled sales and cross-licenses. Imagine firm A sells patent 1 to firm B while firm B sells patent 2 to firm A at the same time. If the firms combined both sales into one assignment, the assignment likely lists each firm as both assignor (sell-side) and assignee (buy-side). Thus, the data for this assignment would look like:

Buyer	Seller	Patent
A	B	1
B	A	1
A	B	2
B	A	2

In this case, we might not know which parties are selling which patents and which parties are buying which patents. Thus, I take two steps to remove possible sources of erroneous attribution. First, in counting patent acquisitions, I drop any acquisition if the assignee received the initial grant of the patent because of the low likelihood that a firm sells a patent and buys it back later. Second, in counting

[64] For these variables, I do not make the distinction between complete ownership and receiving licensing rights.

patent divestitures, I drop any assignor where we can positively identify another seller as the original grant holder (and thus the prohibitively likely seller of that patent). In the above example, assuming A received the grant for patent 1 and B received the grant for patent 2, this procedure would completely solve the issue. In fact, this procedure will completely address all instances of erroneous attribution except cases where the patent is being sold a second time, in which case information about the original owner is irrelevant.

A second, smaller issue exists in the dataset, for which I implement a simple solution. In some instances, a sale is submitted separately by both parties. I drop repeated transactions between the same parties for the same patent and keep only the first instance. Additionally, I drop duplicates when a firm is listed as acquiring or divesting the same patent multiple times in a year.

B.2 Additional results

This section presents extra results and robustness tests to support the main paper.

- Table B.1 presents the industry composition of patent acquisitions, sales and grants. The broad pattern is industries that receive the most patent grants also buy and sell patent rights most often. High R&D industries are well represented, as most acquisitions are in electronic components, computer equipment, drugs, scientific instruments, and computing services.
- Table B.2 presents robustness tests for the main result on R&D in Table 1.4

by varying the specification used. The main results are robust to: replacing the main variables with non-citation weighted versions; replacing *NewPatStock* with the main R&D determinants from the literature (*PatStock* and a measure of intangible capital); including the variables that predict firm-year acquisitions in a firm fixed effect specification (intangible capital and *Sales Growth*, per Table 1.6); accounting for quadratic firm age to capture life cycle issues; and including prior R&D growth.

- Table B.3 presents robustness tests for the main result on R&D in Table 1.4 by varying how the independent variable, R&D, is defined.

- Table B.4 presents robustness tests for the main result on R&D in Table 1.4 by varying the treatment of missing R&D observations and the sample selection criteria. Results become attenuated as the sample includes more firms for which R&D is less important, as would be expected if large fixed costs must be incurred to begin R&D programs.

- Table B.5 presents robustness tests for the dynamic estimation results in Table 1.5 by varying the number of included lags in the R&D model. I test the alternative models against each other and find that models with one or two lags are rejected in favor of three or four lags.

- Table B.6 presents robustness tests for the dynamic estimation results in Table 1.5 by varying the number of included lags in the CAPX model. I test the alternative models against each other and find that models with one or two lags are rejected in favor of three or four lags.

Table B.1: Industry shares of public firm innovation activity

This table reports the share of patent acquisitions, patent divestitures, and patent grants among public firms that are attributable to different industries.

		Acquisitions		Divestitures		Patent grants	
SIC Code	Industry	Fraction	Rank	Fraction	Rank	Fraction	Rank
	Panel A: By SIC2						
36	Electronic and Other Electrical Equipment and Components	26%	1	24%	1	28%	1
35	Industrial and Commercial Machinery and Computer Equipment	17%	2	17%	2	20%	2
28	Chemicals and Allied Products	15%	3	13%	4	14%	3
38	Measuring Instruments, Optical Goods, and Clocks	13%	4	15%	3	9%	4
73	Business Services	8%	5	7%	6	8%	6
37	Transportation Equipment	7%	6	8%	5	9%	5
34	Fabricated Metal Products	2%	7	2%	9	1%	11
48	Communications	2%	8	3%	7	2%	8
29	Petroleum Refining and Related Industries	1%	9	2%	8	3%	7
26	Paper and Allied Products	1%	10	1%	10	2%	9
	Panel A: By SIC3						
367	Electronic Components And Accessories	12%	1	9%	2	13%	2
357	Computer And Office Equipment	10%	2	11%	1	16%	1
283	Drugs	8%	3	6%	5	6%	4
737	Computer Programming, Data Processing, And Related Services	8%	4	7%	4	8%	3
366	Communications Equipment	7%	5	8%	3	6%	5
384	Surgical, Medical, And Dental Instruments And Supplies	5%	6	3%	8	2%	14
382	Laboratory Apparatus And Related Instruments	4%	7	3%	9	2%	16
371	Motor Vehicles And Motor Vehicle Equipment	4%	8	4%	7	5%	6
386	Photographic Equipment And Supplies	3%	9	5%	6	5%	7
372	Aircraft And Parts	3%	10	3%	10	3%	9

Table B.2: Robustness of main R&D specification

This table reports alternative specifications of Model 1 in Table 1.4, which focuses on the relationship between patent acquisitions and future investment and internal innovation output. Analysis is based on an OLS estimation of Equation 1.1. The firm-year sample is described in Table 1.2. The dependent variable is R&D, normalized by lagged assets and multiplied by 100, so coefficients for these models should be interpreted in terms of percentage points. *PatAcq* and *PatDiv* are not citation weighted. *Grants* is the flow of grants a firm receives and is the *NewPatStock*'s analogue of not using citation weights in *PatAcq*. *Age* is defined as the fiscal year minus the first year with price data in Compustat. *R&D Growth*$_{t-1}$ is defined as $Log(R\&D_{t-1}/AT_{t-2}) - Log(R\&D_{t-2}/AT_{t-3})$ and is winsorized. The remaining variables are defined in Appendix A.1. *Total Q, Cash Flow,* and *Log(Assets)* are included but not reported. To facilitate interpretation, *PatStock (CW)*, $Log(1 + K_{INT})$, *Sales Growth* and *R&D Growth* are standardized. Industry by year fixed effects absorb industry specific time trends, and firm fixed effects are included. T-statistics are reported in parentheses. Standard errors are clustered by firm. The symbols ***, **, and *, indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	R&D normalized by lagged assets, $\times 100$							
Log(1+PatAcq (CW))$_{t-1}$	0.081***		0.079***		0.060**	0.051*	0.082***	0.085***
	(2.73)		(2.69)		(2.09)	(1.78)	(2.80)	(2.84)
Log(1+PatDiv (CW))$_{t-1}$	-0.102***		-0.105***		-0.102***	-0.088***	-0.102***	-0.100***
	(-3.31)		(-3.42)		(-3.46)	(-3.07)	(-3.32)	(-3.25)
Log(1+NewPatStock)$_{t-1}$		0.131**	0.134**			-0.255***	0.079	0.128*
		(1.97)	(2.02)			(-3.95)	(1.15)	(1.80)
Log(1+PatAcq)$_{t-1}$				0.093**				
				(2.04)				
Log(1+PatDiv))$_{t-1}$				-0.148***				
				(-3.53)				
Log(1+Grants)$_{t-1}$				0.512***				
				(6.56)				
Log(1+PatStock (CW))$_{t-1}$					-0.375***			
					(-5.20)			
Log(1+K$_{INT}$)$_{t-1}$					4.567***	4.346***		
					(19.07)	(16.95)		
Sales Growth$_{t-1}$						0.259***		
						(3.39)		
Firm Age$_{t-1}$							-0.153	
							(-0.00)	
Firm Age$^2_{t-1}$							-0.002**	
							(-2.01)	
R&D Growth$_{t-1}$								1.033***
								(17.61)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year $\times$ SIC3 FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	51,983	51,983	51,983	51,983	51,983	47,349	51,983	43,076
Adj. R^2	0.74	0.74	0.74	0.74	0.75	0.75	0.74	0.76

Table B.3: Robustness to the definition of R&D

This table reports alternative specifications of Model 1 in Table 1.4, which focuses on the relationship between patent acquisitions and future investment and internal innovation output. Analysis is based on an OLS estimation of Equation 1.1. The firm-year sample is described in Table 1.2. In column 1, the dependent variable is R&D normalized by lagged CAPX. In column 2, the dependent variable is R&D normalized by lagged sales. In columns 1 and 2, the dependent variable is multiplied by 100, so coefficients for these models should be interpreted in terms of percentage points. In column 3, the dependent variable is $Log(1 + R\&D)$. The first six independent variables are defined in Table 1.4 and Appendix A.1. *Log(Sale)* is the log of sales. K_{INT} is defined in Table 1.6 The variables defined as ratios (*R&D/CAPX, R&D/Sales, Total Q,* and *Cash Flow*) are winsorized at the 1% level annually. To facilitate interpretation, *Total Q, Cash Flows, Log(Assets), Log(Sale)*, and $Log(1 + K_{INT})$ are standardized. Industry by year fixed effects absorb industry specific time trends, and firm fixed effects are included. T-statistics are reported in parentheses. Standard errors are clustered by firm. The symbols ***, **, and *, indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	100×R&D$_t$/CAPX$_{t-1}$	100×R&D$_t$/Sales$_{t-1}$	Log(1+R&D)$_t$
Log(1+PatAcq (CW))$_{t-1}$	0.207**	0.044***	0.006***
	(2.28)	(4.28)	(2.87)
Log(1+PatDiv (CW))$_{t-1}$	-0.074	-0.017	-0.016***
	(-0.90)	(-1.51)	(-6.21)
Log(1+NewPatStock)$_{t-1}$	0.896***	0.239***	-0.022***
	(5.70)	(9.46)	(-4.34)
Total Q$_{t-1}$	-0.185*	-0.020	0.109***
	(-1.91)	(-0.91)	(18.15)
Cash Flow$_{t-1}$	-0.671**	0.146***	0.069***
	(-2.01)	(3.23)	(12.26)
Log(Assets)$_{t-1}$	-10.777***		
	(-11.62)		
Log(Sale)$_{t-1}$		-5.117***	
		(-17.21)	
Log(1+K$_{INT}$)$_{t-1}$			1.619***
			(55.36)
Year × SIC3 FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Observations	51,328	51,256	51,983
Adj. R^2	0.44	0.57	0.95

Table B.4: Sample selection and R&D treatment

This table reports alternative specifications of Model 1 in Table 1.4, which focuses on the relationship between patent acquisitions and future investment and internal innovation output. Analysis is based on an OLS estimation of Equation 1.1. The firm-year sample and all variables are described and defined in Table 1.2. Column 1 replicates the baseline model. Each column then adds a new set of firms, where the the restriction of active R&D programs is gradually relaxed. In column 2, observations with zero R&D are allowed if the firm reported positive R&D at any other point. In column 3, all observations with zero R&D are allowed. In column 4, observations are allowed if the firm reported positive R&D at any point and missing R&D is set to zero. In column 4, all observations are allowed and missing R&D is set to zero. To facilitate interpretation, *Total Q, Cash Flows, Log(Assets)*, *PatStock (CW)*, and $Log(1 + K_{INT})$ are standardized. Industry by year fixed effects absorb industry specific time trends, and firm fixed effects are included. T-statistics are reported in parentheses. Standard errors are clustered by firm. The symbols ***, **, and *, indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	R&D normalized by lagged assets, ×100				
Log(1+PatAcq (CW))$_{t-1}$	0.079***	0.080***	0.079***	0.050*	0.024
	(2.69)	(2.69)	(2.68)	(1.78)	(1.07)
Log(1+PatDiv (CW))$_{t-1}$	-0.105***	-0.108***	-0.106***	-0.098***	-0.070***
	(-3.42)	(-3.52)	(-3.48)	(-3.45)	(-3.03)
Log(1+NewPatStock)$_{t-1}$	0.134**	0.126*	0.121*	0.072	-0.001
	(2.02)	(1.89)	(1.83)	(1.18)	(-0.03)
Total Q$_{t-1}$	0.445***	0.459***	0.461***	0.416***	0.422***
	(4.64)	(4.82)	(4.83)	(4.55)	(5.18)
Cash Flow$_{t-1}$	-2.798***	-2.807***	-2.800***	-2.760***	-2.601***
	(-19.28)	(-19.38)	(-19.34)	(-19.41)	(-19.94)
Log(Assets)$_{t-1}$	-9.884***	-9.658***	-9.580***	-8.125***	-5.877***
	(-25.29)	(-24.91)	(-24.89)	(-23.72)	(-22.30)
Year × SIC3 FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Observations	51,983	52,530	53,556	59,562	75,198
Adj. R^2	0.74	0.74	0.74	0.74	0.76
Sample condition:	R&D$_t > 0$	R&D> 0 during sample	R&D not missing	R&D > 0 during sample	All
Treatment of missing R&D:	Missing	Missing	Missing	Zero	Zero

Table B.5: Dynamic lag selection in the R&D model

This table reports alternative specifications of Model 1 in Table 1.5, which focuses on the dynamic relationship between patent acquisitions and future investment and internal innovation output. Analysis is based on an OLS estimation of Equation 1.2. The firm-year sample and variable definitions are described in Table 1.2. The dependent variable is R&D normalized by lagged assets and multiplied by 100, so coefficients for these models should be interpreted in terms of percentage points. Likelihood ratio p-values are reported containing tests of against each nested model. Rejecting the null implies the larger model fits the data better. Firm-year level controls *Total Q, Cash Flows,* and *Log(Assets)* are not reported for brevity. Industry by year fixed effects absorb industry specific time trends, and firm fixed effects are included. T-statistics are reported in parentheses. Standard errors are clustered by firm. The symbols ***, **, and *, indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	R&D			
	(1)	(2)	(3)	(4)
Log(1+PatAcq (CW))$_{t-1}$	0.079***	0.075**	0.091***	0.071**
	(2.69)	(2.56)	(3.04)	(2.44)
Log(1+PatAcq (CW))$_{t-2}$		0.043	0.016	0.004
		(1.51)	(0.57)	(0.14)
Log(1+PatAcq (CW))$_{t-3}$			0.108***	0.088***
			(3.50)	(2.94)
Log(1+PatAcq (CW))$_{t-4}$				0.036
				(1.09)
Log(1+PatDiv (CW))$_{t-1}$	-0.105***	-0.099***	-0.106***	-0.101***
	(-3.42)	(-3.29)	(-3.53)	(-3.41)
Log(1+PatDiv (CW))$_{t-2}$		-0.046	-0.036	-0.022
		(-1.44)	(-1.22)	(-0.74)
Log(1+PatDiv (CW))$_{t-3}$			-0.068**	-0.061**
			(-2.23)	(-2.07)
Log(1+PatDiv (CW))$_{t-4}$				-0.003
				(-0.08)
Log(1+NewPatStock)$_{t-1}$	0.134**	0.101	0.105	0.091
	(2.02)	(1.47)	(1.46)	(1.23)
Controls	Yes	Yes	Yes	Yes
Year × SIC3 FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	51,983	47,813	43,893	40,159
Adj. R^2	0.74	0.75	0.75	0.75
Likelihood ratio p-values:				
Relative to null of model (1):		0.78	0.01	0.03
Relative to null of model (2):			<0.01	0.01
Relative to null of model (3):				0.51

Table B.6: Dynamic lag selection in the CAPX model

This table reports alternative specifications of Model 1 in Table 1.5, which focuses on the dynamic relationship between patent acquisitions and future investment and internal innovation output. Analysis is based on an OLS estimation of Equation 1.2. The firm-year sample and variable definitions are described in Table 1.2. The dependent variable is CAPX normalized by lagged assets and multiplied by 100, so coefficients for these models should be interpreted in terms of percentage points. Likelihood ratio p-values are reported containing tests of against each nested model. Rejecting the null implies the larger model fits the data better. Firm-year level controls *NewPatStock, Total Q, Cash Flows,* and *Log(Assets)* are not reported for brevity. Industry by year fixed effects absorb industry specific time trends, and firm fixed effects are included. T-statistics are reported in parentheses. Standard errors are clustered by firm. The symbols ***, **, and *, indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	CAPX			
	(1)	(2)	(3)	(4)
Log(1+PatAcq (CW))$_{t-1}$	0.026	0.015	0.019	0.011
	(1.39)	(0.82)	(1.03)	(0.60)
Log(1+PatAcq (CW))$_{t-2}$		0.036*	0.032*	0.028
		(1.93)	(1.74)	(1.48)
Log(1+PatAcq (CW))$_{t-3}$			0.049**	0.042**
			(2.49)	(2.15)
Log(1+PatAcq (CW))$_{t-4}$				0.043**
				(1.97)
Log(1+PatDiv (CW))$_{t-1}$	-0.085***	-0.087***	-0.087***	-0.091***
	(-4.27)	(-4.56)	(-4.58)	(-4.79)
Log(1+PatDiv (CW))$_{t-2}$		-0.048**	-0.047**	-0.040**
		(-2.43)	(-2.42)	(-2.11)
Log(1+PatDiv (CW))$_{t-3}$			-0.052***	-0.050***
			(-2.75)	(-2.66)
Log(1+PatDiv (CW))$_{t-4}$				-0.036*
				(-1.81)
Controls	Yes	Yes	Yes	Yes
Year × SIC3 FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	51,983	47,813	43,893	40,159
Adj. R^2	0.40	0.41	0.42	0.43
Likelihood ratio p-values:				
Relative to null of model (1):		0.07	$<$0.01	$<$0.01
Relative to null of model (2):			$<$0.01	$<$0.01
Relative to null of model (3):				0.15

Chapter 2: What's your Identification Strategy? Innovation in Corporate Finance Research

With Laurent Frésard and Jérôme P. Taillard

2.1 Introduction

How do individuals and organizations make decisions about the use of new technologies? This question is important as the diffusion of innovation is believed to be a key channel through which productivity growth is achieved (e.g. Aghion and Howitt [1992]). Simple intuition suggests that profitable innovations should be adopted almost instantaneously. Yet, various frictions can act as barriers that slow the diffusion of innovation. The rate of adoption of newly invented technologies can vary substantially across agents that differ in their knowledge of the costs and benefits of adopting, their access to the technologies, their learning ability, financial resources, previous experiences, human capital, and exposure to the technology through peers (e.g. Hall and Khan [2002] or Foster and Rosenzweig [2010]).

In this paper, we study the diffusion of a particular innovation: Econometric techniques designed to estimate causal relationships in empirical corporate finance research. The use of such techniques has recently become a widespread tool for corporate finance researchers, mirroring the trend observed in other areas of economics (e.g. labor or development).[1] Proponents claim that these econometric techniques offer a "credibility revolution" for researchers (see Angrist and Pischke [2010]), thus enhancing their ability to make causal statements, policy recommendations, and gain credibility in the marketplace for ideas.

Studying this specific innovation allows us to shed new light on the forces that drive the diffusion of innovation in financial research and the establishment of new academic norms. We construct a unique sample that comprises the *content* of all empirical corporate finance articles published in the leading ("top-three") finance academic journals between 1970 and 2012.[2] This sample is comprised of 1,796 articles written by 1,880 distinct authors for whom we hand-collect detailed biographical information. We consider five main identification techniques: Instrumental variables; difference-in-differences (and natural experiments); selection models; regression discontinuity design; and randomized experiments. We treat these identification techniques as innovations, and taken together, they form what we coin the "identification technology." We exploit the fine granularity of our dataset to examine in detail the patterns and determinants of technological adoption in

[1]Attesting to the increasing focus on identification in corporate finance research, several recent methodology surveys outline current best practices to address issues related to endogeneity in corporate finance (e.g. Li and Prabhala [2007], Roberts and Whited [2013], Strebulaev and Whited [2012]), and several recent keynote speeches at major conferences have concentrated on identification (e.g. Toni Whited at the 2014 SFS Cavalcade or Wei Jiang at the 2015 SFS Cavalcade).

[2]These are the Journal of Finance (JF), the Journal of Financial Economics (JFE), and the Review of Financial Studies (RFS).

academia.

We first document a recent and fast adoption of the identification technology in corporate finance. The fraction of articles using the identification technology is close to zero until the end of the eighties and then steadily rises to reach more than 50% in 2012. The adoption pattern is consistent with an S-shaped diffusion curve, similar to diffusion patterns observed for "harder" types of technology (e.g. Griliches [1957], Mansfield [1961], and more recently, Comin and Hobijn [2010]). By fitting logistic diffusion curves to the data and using the conventional 5% adoption threshold, we estimate that the year 1998 marks the emergence of the identification technology in finance journals.

We then use various adoption thresholds to compare the emergence of the identification technology in top finance and top economics journals.[3] Focusing on the language of identification, we document that the identification technology emerges in finance journals about twenty years after its emergence in the top economics journals. Given the relatively blurred boundaries between economics and corporate finance research, the long adoption lag between the two fields suggests the presence of important frictions that impede the diffusion of the identification technology across research fields.

To shed light on the possible frictions at play, we examine the determinants of adoption and measure how adoption rates vary across researchers. We conjecture

[3]Following the classification of Card and DellaVigna [2013], we focus on the top-three economics journals: the American Economic Review, the Journal of Political Economy, and the Quarterly Journal of Economics.

that a given researcher will adopt as soon as the expected net gains from using the identification technology are positive. Estimating such expected net gains is empirically challenging however, as we do not observe the relevant inputs in each researcher's utility function. We posit that researchers attempt to maximize their lifetime scientific impact and recognition, which we approximate by using citations and publications (see Hamermesh and Pfann [2012]).

We estimate that identification articles attract 22% more citations on average, after controlling for various author and paper-level characteristics. Such an identification "premium" started in 1995 for finance articles and as soon as 1982 for economics articles. In parallel, we document a substantial change in the composition of editorial boards at the leading finance journals, with a substantial increase in the fraction of board members who have adopted the technology in their own research. This shift can arguably increase the odds of publication for papers relying on the identification technology. The aggregate citation premium, combined with the change at the editorial board level, suggests that the benefits of adopting the identification technology increased over our sample period.

The costs and benefits of adopting the technology vary across researchers. Using a survival analysis to estimate the determinants of time-to-adoption across researchers, we show that knowledge about the net gains from adopting is related to the diffusion of the identification technology. In particular, our estimates indicate a faster adoption among researchers holding a PhD degree in economics compared to finance PhDs. The magnitude is large as researchers with PhDs in economics adopt almost 30% faster. This significant differential suggests that doctoral training in economics increases researchers' exposure to the identification technology, which in turn improves their inference about the net gains of adopting the innovation. Relatedly, we document that authors that have previously published articles in economics journals – whom we call "straddlers" – are among the early adopters in finance journals. These findings highlight that the diffusion of innovation can be accelerated when researchers with a given training migrate to a neighboring field of research (see Stoyanov and Zubanov [2012]).

We also find that the adoption speed is related to an author's research network. Specifically, we show that the adoption rate is positively and significantly related with the extent of prior adoption by a researcher's network of current and previous colleagues. Moreover, researchers adopt the identification technology significantly faster when a larger fraction of alumni researchers from their PhD-granting institutions have already adopted the technology. Both results are consistent with networks creating positive externalities that facilitate the diffusion of knowledge among researchers and reduce the informational constraints associated with learning about the new technology.[4]

We find no evidence that the school ranking of a researcher's PhD-granting institution is related to the likelihood of adopting the identification technology.

[4] Learning through peers and similar networking effects have been documented in other settings. For instance, Bayer, Hjalmarsson, and Pozen [2009] finds strong peer learning effects among juveniles in correctional facilities. In academia, Azoulay, Graff Zivin, and Wang [2010] shows that superstar researchers generate significant positive spillover effects among their network of collaborators.

In contrast, the ranking of a researcher's affiliation at the time of publication is significantly related to the speed of adoption. Researchers employed at top-tier institutions adopt the identification technology significantly sooner than their peers at lower-ranked institutions. The relationship is not necessarily a causal one: The faster adoption of researchers at top institutions could indicate that the larger financial and organizational resources allocated to research at top schools plays a role in explaining the observed "top-down" diffusion of identification techniques. But this finding could also suggest that top schools are simply better at selecting innovators and early adopters.

We estimate that seniority is significantly associated with the diffusion of the identification technology. Controlling for cohort and year effects, younger authors tend to adopt earlier. This age effect is consistent with models of vintage human capital implying that early-stage researchers are the primary adopters of new technology (e.g. Chari and Hoenhayn [1991] or MacDonald and Weisbach [2004]). The age effect could be due to the more flexible minds of young scholars (e.g. Darwin [1859]), their larger exposure to recent innovation, or higher incentives and greater amount of time to learn new techniques (e.g. Diamond [1980]). Alternatively, more senior scholars could have more vested interests that may render them less receptive to innovation (e.g. Cohen [1985]). Although they adopt later and less frequently on average, a significant number of senior researchers do adopt the identification technology over our sample period. We show that seniors are more likely to adopt by co-authoring papers and that their coauthors are younger on average for identification articles.

While we primarily focus on the diffusion of the identification technology, we also document changes *within* the technology itself. In particular, we show that best practices for the implementation of the technology – or simply the "norms" – are evolving. Specifically, we track the evolution of several technological refinements (e.g. the appearance of weak instrument tests for the implementation of instrumental variables) and confirm that the boundaries of the identification technology change significantly over time. This dynamic process is consistent with the notion that most innovations and scientific improvements are the result of experimentation and refinements (e.g. Popper [1959] or Kuhn [1962]). Moreover, the genuinely transient nature of best practices makes the judgment of the proper implementation of the identification technology at any given point in time a somewhat delicate exercise.

This study makes several contributions. First, it adds to the growing literature that studies the determinants of the diffusion of scientific advances and new technologies. Mokyr [2002] argues that while the stock of existing knowledge is important for creating new knowledge, the effective diffusion of knowledge across researchers requires them to be not only aware of existing knowledge, but also to be able to access it. Confirming this idea, existing research indicates that the reduction of informational, institutional, or legal barriers to knowledge spurs the creation and diffusion of new ideas. For instance, Agrawal and Goldfarb [2008], Ding, Levin, Stephan, and Winkler [2010], Furman and Stern [2011], and Kim, Morse, and Zingales [2009] show that improvements in communication technologies render researchers more productive, foster knowledge accumulation, and increase collaboration. Other studies indicate that the cost of accessing existing ideas, as measured by geographical proximity (e.g. Azoulay, Graff Zivin, and Sampat [2012] and Head, Li, and Minondo [2015]) or protection of intellectual property (e.g. Williams [2013] and Teodoridis [2015]), is related to the flow of academic knowledge and inventive activities. Our study is complementary as we examine the forces that explain the diffusion of a *soft* innovation among researchers in a setting where legal and financial barriers are low. We find that human factors related to the cost of accessing the existing knowledge (e.g. PhD training and research networks) are empirically important and help explain the diffusion of new technologies across the boundaries of research fields.

Our work also relates to the "burden of knowledge" concept (e.g., Jones [2009] and Teodoridis [2015]). Jones [2009] shows that, as the stock of knowledge grows in the economy, researchers need to spend more time in training and narrow their field of study to reach the ever-expanding frontier of knowledge. As a consequence, researchers become more specialized and collaborate more. Our evidence is supportive of his theory. Researchers specialized in corporate finance adopt only with a significant lag the econometric techniques established outside their area. Further, the diffusion of the identification technology comes in part through early adopters in the field of economics who migrate and bring the technology to finance journals. We also highlight a mutually beneficial "cross-generational" collaboration, whereby seniors team up with juniors to overcome the extension of the frontier of knowledge.

Our findings further contribute to the literature studying the evolution of academic research in finance. This introspective and rhetorical literature is relatively thin in finance compared to economics and other fields in social sciences. This line of inquiry used to be more prevalent in the top-three finance journals.[5] The recent studies in this tradition concentrate on evaluating the adequacy of statistical tools used in corporate finance research.[6] Our focus is distinct, as we study the shift towards identification techniques in corporate finance research and the resulting changes in the profession.

Finally, we need to underline two limitations of our analysis. First, perhaps paradoxically, our paper is not an identification paper. Our study concentrates on analyzing patterns of adoption and highlighting factors that are significantly related to these adoption patterns, but not necessarily in a *causal* manner. Second, while we believe that our framework is useful to shed new light on some of the forces that drive technological adoption, we recognize that our academic setting is atypical and hence, extrapolation of our results to other contexts deserves careful thought.

The paper proceeds as follows. In Section 2.2, we provide the conceptual

[5] Weber [1973] and Keenan [1991] study trends in finance publications. Kaufman [1984], Niemi [1987], and Alexander and Marby [1994] evaluate the performance of departments and journals. Zivney and Bertin [1992] and Borokhovich, Bricker, Brunarski, and Simkins [1995] examine the productivity of finance researchers. Corrado and Ferris [1997] analyze the influence of published articles on finance doctoral education. There are a few notable recent studies: Karolyi [Forthcoming] examines general trends in finance publications, Welch [2014] analyzes the refereeing process in finance and economics journals, and Brogaard, Engelberg, and Parsons [2014] examines how proximity with an editor influences research productivity.

[6] Examples include Atanasov and Black [Forthcoming], Campello, Galvao, and Juhl [2013], Gormley and Matsa [2014], Karolyi [2011], Petersen [2009], Roberts and Whited [2013], and Thompson [2011].

framework for our study. In Section 2.3, we describe the construction of our sample. In Section 2.4, we show general diffusion patterns of the identification technology. Section 2.5 explores the determinants of diffusion at the researcher-level. Section 2.6 highlights technological refinements over time. Section 2.7 concludes.

2.2 Conceptual Framework

How do finance researchers decide to adopt or not the identification technology? To help organize the discussion and guide our empirical analysis, we follow the theoretical literature on technology adoption. The backbone of this literature is an empirical regularity: When the number of users of a new innovation is plotted versus time, the resulting curve is typically an S-curve (or ogive distribution). This general pattern appears natural as one imagines adoption proceeding slowly at first, accelerating as it spreads, and slowing down as the relevant population becomes saturated (e.g. Comin and Hobijn [2010]).

In an attempt to explain this pattern, existing models of technology diffusion typically consider the decision of an agent (e.g. a firm, an organization, or an individual) to adopt a new technology at a given date. The agent chooses to adopt when the expected net benefits of adopting are positive. The central prediction of existing models is that potential adopters will adopt at different dates (or not adopt at all). The asynchronous adoption pattern originates from differences in agents' inherent characteristics (e.g. skills, experience, or beliefs) and/or information about the expected net benefits associated with the technology (see Rogers [2003]).

Early work typically models the diffusion of a new technology using epidemic theories according to which, perfectly homogeneous agents decide to adopt a new technology when they (exogenously) receive information relating to its existence and its associated net benefits (e.g. Griliches [1957] or Mansfield [1961]). In these models, the pattern and speed of adoption is entirely driven by the awareness of the new technology (or lack thereof) and the velocity of information spreading. In contrast, more recent theoretical work considers heterogeneous agents that can differ on various dimensions and focuses on the decision to adopt a new technology (e.g. Hall and Khan [2002] or Foster and Rosenzweig [2010]). Potential adopters have different characteristics and differential access to information about the technology and, as a result, form distinct expectations about the costs and benefits from adopting the new technology.[7] In this framework, the pattern of adoption is dictated by intrinsic differences among potential adopters that can include, for instance, differences in education levels, access to information, financial capital, personal experiences, social connections, or risk aversion.

Estimating the expected return to adoption of a new technology is notoriously challenging. Our specific academic setting is no exception. To make our empirical analysis tractable, we assume that finance researchers are economic agents who make decisions that maximize the present value of their expected lifetime utility.[8]

[7]See Kremer and Miguel [2007] for a specific example of heterogeneous agents modeling.

[8]As emphasized by Foster and Rosenzweig [2010], the estimation of return to adoption is a general problem in the literature. In the case of technology used by profit maximizing entities, it is clear that technology profitability is a good measure of return to adoption. For technologies that improve an agent utility (e.g. those that improve health, happiness, or scientific impact)

Based on this assumption, they adopt a new technology when doing so increases their expected utility. Arguably, the utility function of researchers is complex and features many different inputs (and associated weights) that can vary significantly across individuals. For instance, financial gains associated with additional grant funding, salary increases, or tenure decisions can all increase the utility of researchers. But, there can also be non-pecuniary elements, such as titles, honorary memberships (e.g. editorial board positions, Distinguished Fellowships), solicitations to give talks or serve as experts in the media and courts, or simply the pursuit of true scientific endeavors.

In this context, a scholar's decision to adopt a technological innovation in his research is made by trading off the expected adoption costs and the increased expected utility associated with the use of the innovation. While it is a priori unclear what elements increase a researcher's utility, we posit that a researcher's utility is directly linked to her reputation and impact in the scientific community.[9] Based on existing studies that examine the determinants of researchers' career paths (e.g. Hamermesh, Johnson, and Weisbrod [1982], Moore, Newman, and Turnbull [2001], or Hamermesh and Pfann [2012]), we assume that researchers' scientific impact and recognition is determined by the number of peer-reviewed publications obtained and citations they garner. Therefore, our analysis builds on the assumption that

measurement of returns is much less straightforward. Agents choose to use a technology based on gains in terms of welfare, which cannot be easily observable.

[9]Hamermesh and Pfann [2012] show that reputation is positively related to both pecuniary and non-pecuniary rewards that provide researchers with utility.

a researcher adopts the identification technology in her research when she expects that the use of the new technology will increase her overall research impact, net of adoption costs.

Based on this simple conceptual framework, our empirical analysis first describes the general pattern and speed of adoption of the identification technology in empirical corporate finance research. In a second step, we concentrate on the determinants of adoption to understand what factors drive – or hinder – the diffusion of the identification technology. We focus on characteristics of researchers and their research networks that could affect their awareness of the identification technology and/or its potential benefits, their incentives to adopt, and their adoption cost. In doing so, we explicitly recognize the heterogeneous nature of the pool of researchers we study.

2.3 Sample Construction

This section details the sample. We first explain the construction of the main sample of empirical corporate finance articles and authors. Second, we outline the procedure we followed to define and identify articles that rely on the identification technology.

2.3.1 Empirical Corporate Finance Research

We start by constructing a dataset containing detailed information on all published articles in the Journal of Finance (JF), Journal of Financial Economics (JFE), and Review of Financial Studies (RFS) for the period 1970 to 2012. Using the JSTOR database for JF and RFS articles and ScienceDirect for JFE articles, we collect the following variables for each article: Year of publication, title, authors, journal, volume, issue, and page range. We screen for errata, comment, correction, reply, or discussion articles. We apply further screens to filter common issue fillers (e.g. "Back Matter"), results related to annual meetings, and abstracts of doctoral dissertations. The initial sample contains 3,579 JF articles, 1,969 JFE articles, and 1,249 RFS articles for a total top-three sample of 6,797 articles.[10]

We then identify empirical corporate articles as follows. First, we define an article as empirical if it performs statistical tests on real data (see Hamermesh [2013]). With this broad definition, the only excluded articles are: (1) Pure theory articles; (2) articles that use only simulated data; and (3) articles that calibrate a model based on prior publication results and not self-generated inference.

Second, we define corporate finance articles using the Journal of Economic Literature (JEL) codes and resort to a manual classification for the subset of articles for which JEL codes are not available. All JFE articles since 1993 and most RFS articles since 2006 directly contain JEL codes (1,994 articles). Of the remaining articles (4,803), we search the EconLit database and find JEL codes for 2,402 articles. We assign articles into the corporate finance group if they contain at least one

[10]The Internet Appendix presents some general trends in publication between 1970 and 2012. Overall the patterns in our sample are consistent with the findings of other articles focusing on publication trends in economics journals (e.g. Hamermesh [2013], Card and DellaVigna [2013], and Ellison [2002]).

JEL code starting with G3 ("Corporate Finance and Governance").[11] For articles without JEL codes (2,401 articles), we manually define them as corporate finance or non-corporate finance. To enhance integrity and accuracy, each article was independently evaluated by two of the authors.[12] We identify a total of 1,796 empirical corporate finance articles in the top-three finance journals between 1970 and 2012.

For each article we gather citation data from Thomson Reuters' Web of Knowledge (WOK). We collect the "Times Cited" variable, which tracks the total number of citations received by the article across all databases indexed by WOK, as of November 2013. Additionally, we collect the year, title, authors, and journal of each citation. From this, we are able to track the citations a paper receives every year. 99.6% of our sample contains citation information from WOK.[13]

Finally, we collect detailed biographical information on authors from their personal or professional websites. Between 1970 and 2012, 1,880 distinct authors published at least one empirical corporate article in a top finance journal. In par-

[11] See http://www.aeaweb.org/jel/guide/jel.php for a description of JEL classifications.

[12] Note that we assign "banking" articles into the corporate finance group if the main variable of interest is a corporate decision (e.g. capital structure or the size of the board) made by non-financial firms. To illustrate the outcome of our classification, Table IA.1 in the Internet Appendix presents the assignment of the two most cited articles every year into empirical (or not) and corporate (or not), as well as the source for the classification (JFE, RFS, or EconLit JEL codes or Manual). A visual inspection of this table reveals that our classification performs well. Out of 90 cases, the only potential misclassification is the article by Bansal and Yaron [2004] that is classified as corporate finance based on the JEL codes from EconLit.

[13] Excluding JFE articles in 1974 and 1975 and RFS articles in 1988 and 1989, for which WOK did not have records as of the time of our data collection.

ticular, we identify the field of PhD (e.g. finance or economics) for 1,610 authors (85.6%), year of PhD graduation for 1,620 authors (86.2%), PhD-granting institution for 1,718 authors (91.4%), and the employment history (i.e. career path) for 1,554 authors (82.7%).

2.3.2 The Identification Technology

We devise a simple classification procedure to determine whether an article uses an econometric technique designed to identify causal relationships. We consider five widely-used techniques: (1) Instrumental variables (IV), (2) difference-in-differences and (quasi-)natural experiments (DD), (3) selection models, (4) randomized experiments (field or lab), and (5) regression discontinuity designs (RDD). Henceforth, we use DD as shorthand that includes quasi-natural experiments. We label these five techniques as identification techniques. Together they form the "identification technology."[14]

We start by flagging articles that *may* use each technique by searching the full text of every article for a list of keywords (a dictionary) associated with each technique. As an example, we use the following terms to form the IV-specific dictionary: "instrumental variable", "two stage least square", "three stage least square", "2SLS", and "3SLS" (allowing for all common permutations, hyphenation, and plural form). The dictionary for all five techniques is reported in Table C.1 of Ap-

[14] In unreported analyses, we also considered matching techniques and structural estimations as part of the identification technology. The results presented throughout the paper are virtually identical with the addition of these two techniques.

pendix C.1. The dictionaries are designed to catch all articles that refer to the technique—that is, we purposefully err on the side of encouraging false positives to reduce the number of false negatives. In a second step, we verify manually each flagged article and remove all false positives. We then classify each corporate finance articles as "identification" if the paper uses the identification technology, and "non-identification" if it does not.[15] We provide more details about the classification in Appendix C.1.

2.3.3 Descriptive Statistics

Table 2.1 provides a description of our sample of published papers in the top-three finance journals between 1970 and 2012. Empirical corporate finance articles represent 26% of all published articles in the top-three finance journals over the entire sample period. Panel A of Table 2.1 highlights that this fraction has largely increased over time; specifically, the share of empirical corporate finance articles has grown from 8% in the 1970s to 39% during 2010-2012.

[Insert Table 2.1 about here]

Panel B of Table 2.1 provides details on the use of the identification technol-

[15]Note that while we classify articles based on the econometric technique they use, we do not attempt to evaluate whether these techniques are applied "correctly". For instance, we remain agnostic as to whether a given natural experiment provides a truly exogenous source of variation in the data or whether an instrument satisfies the exclusion restriction. We believe this exercise is outside of the scope of this paper (see Section 2.6 for further discussion). We manually checked a random sample of 100 non-flagged articles, and found only four false negatives. We conclude that our empirical approach minimized the potential impact of false negatives.

ogy within the subset of empirical corporate finance papers over the sample period. We uncover 408 articles that adopt the identification technology. Taken together, identification (ID) articles represent 22.7% of all published empirical corporate finance articles over the 1970-2012 period. These articles are written by 632 distinct authors. IV is the most established technique, hence it is perhaps not surprising to find that it is the most represented technique with more than half of all ID articles using an IV approach.[16] The difference-in-differences category (which includes all natural experiments) comes in second, followed by selection models, regression discontinuity designs, and, lastly experiments (with only five articles). Regression discontinuity is the most recent identification technology to emerge in top finance journals.

2.4 Diffusion Patterns

We start the analysis by examining the overall rate of adoption of the identification technology in corporate finance. We then measure the adoption lag between finance and economics journals.

2.4.1 The Evolution of Identification

Based on our classification, we simply define p_t as the fraction of the population (i.e., empirical corporate finance articles) that has adopted the identification

[16]Note that an identification article can be associated with several techniques, and as such, the techniques listed in this section are not necessarily mutually exclusive.

technology in year t. We plot p_t over the period 1980-2012 in Figure 2.1.[17] We observe a substantial increase in the adoption rate over the sample period (the solid line) from virtually zero in the eighties to more than 50% in 2012.

[Insert Figure 2.1 about here]

Similar to many hard innovations, the adoption path of the identification technology resembles a logistic distribution (S-shape curve), which is marked by a relatively long period of slow diffusion that is followed by a period of rapid diffusion. To formally estimate the speed of adoption and date the emergence of the identification technology in corporate finance, we fit a logistic diffusion model to the data. Specifically, we estimate the following equation:

$$p_t = \frac{k}{1 + e^{-(\alpha + \beta \times t + \varepsilon_t)}}, \tag{2.1}$$

where α and β are parameters that determine the position and slope of the S-curve, and ε_t is a normally distributed error term. The variable k represents the "ceiling" value of the S-curve. It captures the equilibrium fraction of the population that will eventually adopt the identification technology. Because the adoption is still under way (it reaches 52% in 2012) and will likely not achieve 100%, we cannot easily estimate k. In our baseline estimates, we set $k = 1$, but we also show results for $k = 0.8$ and $k = 0.6$ and conclude that the choice of k has little bearing on our conclusions. Moreover, we obtain a close fit when superimposing the fitted value $\tilde{p}_t$ on top of p_t in Figure 2.1.

[Insert Table 2.2 about here]

[17] Although we identify one IV article prior to 1980, we decided to be consistent throughout and present results, tables, and figures using the sample years 1980-2012.

Following the convention in the technology diffusion literature, we determine that a technology has "emerged" when it becomes adopted by more than 5% of the population ($p_t > 5\%$). Using this definition, the date of emergence is given by the implied year $\tilde{t}_{5\%}$ that corresponds to the year when the fitted S-curve crosses the 5% threshold ($\tilde{p}_t = 5\%$).[18] We report the estimated value for $\tilde{t}_{5\%}$ in the first column of Table 2.2 (Panel A). Our baseline estimation indicates that 1998 marks the emergence of the identification technology in finance journals. As a robustness check, panels B and C reveal that this estimate is fairly insensitive to our choice of k. The 5% adoption threshold is crossed in 1996 if we set $k = 0.8$ and in 1995 if we set $k = 0.6$.

Our estimates reveal a fast diffusion rate. Although we do not have a good benchmark (i.e., estimates of the diffusion speed of related innovation), the first column of Table 2.2 (Panel A) indicates that the diffusion rate doubles in only two years after its emergence. Starting from the baseline emergence level of 5% in 1998, the diffusion rate increases to 10% in 2000 ($\tilde{t}_{10\%} = 2000$). It doubles again to 20% over the following three years ($\tilde{t}_{20\%} = 2003$). Overall, the diffusion rate increases from 5% to 50% in only 11 years ($\tilde{t}_{50\%} = 2009$).

The remaining columns in Table 2.2 further report the date of emergence of several of the ID techniques, namely instrumental variables, difference-in-differences, and selection models. We concentrate on these three techniques because we cannot

[18] This number is computed using the 5% threshold of the cumulative distribution of our estimated normalized logistic function. Mathematically, $\tilde{t}_{5\%} = 1980 + \frac{-2.94-\hat{\alpha}}{\hat{\beta}}$. The year 1980 corresponds to the first year of the sample period over which we estimate the model.

estimate similar dates of emergence for regression discontinuity design and experiments since their fitted adoption curves do not cross the 5% adoption threshold during the sample period. Panel A indicates that the implied emergence year for the IV approach is 2001; it is 2006 for difference-in-differences, and 2011 for selection models. (Results from Panel B and C are similar to those in Panel A and hence are not discussed for brevity.) Remarkably, the speed of adoption is roughly similar across the three techniques going from 5% to 10% in the space of about two years, and reaching 20% approximately six years after reaching the 5% threshold.

2.4.2 Origins and Lags

The identification technology was not originally developed to tackle corporate finance questions but rather originated as a set of solutions to endogeneity problems in other subfields of economics. To better characterize the diffusion of the identification technology in corporate finance, we contrast its emergence with that observed in its original terrain: Top economics journals.

We complement our sample by including *all* articles published in the top-three economics journals between 1970 and 2012 (i.e., the American Economic Review (AER), the Journal of Political Economy (JPE), and the Quarterly Journal of Economics (QJE)). For this exercise, our original sample of 1,796 empirical corporate finance articles is augmented by 8,156 articles published in the top-three economics journals. To simplify our task of contrasting the emergence of the identification technology in both sets of journals, we focus only on the *language* associated with the identification technology (instead of the actual use of identification techniques). We rely on a dictionary of distinct keywords that are unambiguously associated with the concept of identification (presented in Table C.2 in Appendix C.1). The dictionary comprises (1) general terms (e.g. exogenous variation) as well as (2) the same technique-specific terms we use to flag potential identification article in the first step of our classification scheme. For each article, we create a dummy variable that equals one if an article uses at least one of the keywords and aggregate these dummies by year for both economics and corporate finance articles.

[Insert Table 2.3 about here]

The results from estimating diffusion-curve models on each sample separately are reported in Table 2.3. We observe a clear upward trend in the frequency of the vocabulary associated with identification within both economics and empirical corporate finance articles. In the first column, we estimate that the 5% and 10% adoption thresholds are reached in economics journals in 1963 and 1974, respectively. These estimates confirm that the identification technology has been used in economics journals for a very long time. For instance, Stock and Trebbi [2003] indicate that the instrumental variable regression technique was first introduced by Wright [1928] to estimate the slopes of supply and demand curves for animal and vegetable oils. Relatedly, our textual search identifies seven articles referring to instrumental variables in economics journals in 1970 (the first year of our sample). In contrast, the 5% adoption threshold is reached in finance only in 1992, while the 10% threshold is crossed in 1996. This implies adoption lags of approximately 29 and 22 years for corporate finance (each difference is statistically significant). The average lag taken across the four adoption thresholds (5% to 20%) is approximately 20 years.[19]

Taken together, these estimates provide an empirical validation to the widely held belief that finance "lags" economics when it comes to adopting new methodological tools. While the identification vocabulary emerges with a significant lag in corporate finance, we observe that the adoption rate is significantly faster within finance journals. Table 2.3 reveals that the adoption rate progressed from 5% to 20% in 23 years in economics journals (1963-1986). In contrast, it took only eight years to reach the 20% adoption threshold in corporate finance (1992-2000).

2.4.3 Empirical Corporate Finance in Economics Journals

Our main focus is on empirical corporate finance articles published in the top-three finance journals. Yet, to provide a complete picture of the adoption of the identification technology in corporate finance, we extend our definition of "corporate finance" research to include empirical corporate finance articles published in the top-three economics journals. We identify these articles based on their JEL codes, and classify an article as corporate finance if it contains at least one JEL code starting with G3. We identify 242 corporate finance articles published in the economics journals between 1980 and 2012. We read each article to determine if it is (a) empirical, and (2) uses one of the five identification techniques. We find 160 empirical corporate finance articles, 45 of which use the identification technology.

[19] We obtain similar results if we assume lower saturation levels; namely we estimate the lag at 20.5 years with $k = 0.8$ and at 12.3 years with $k = 0.6$.

[Insert Table 2.4 about here]

Table 2.4 replicates Table 2.2 but expands the set of empirical corporate finance articles (i.e., the population in equation (2.1)) to include both finance and economics journals (1,878 articles). The estimated adoption of the identification technology is notably earlier when we include articles from economics journals. The 5% adoption threshold is crossed in 1990, and the 10% threshold crossed in 1996, as opposed to 1998 and 2000 when we concentrate only on finance journals. With this more inclusive definition of published work in empirical corporate finance, the adoption lag with economics journals is now shorter. This confirms the overall later adoption of the identification technology in finance journals compared to economics journals. Moreover, the estimates suggest that the identification technology in empirical corporate finance first appears through articles published in the top economics journals. We confirm this result in Section 2.5.2.5.3 when we look in greater details at the adoption patterns of a subset of authors that "straddle" the boundary between finance and economics journals.

2.5 The Determinants of Adoption

As highlighted in Section 2.2, an agent's decision to adopt a new technology at a given point in time depends on the expected net gains of adopting. We examine in this section factors that could be associated with such gains. We start by approximating the aggregate expected net gains from adopting using citations and publication odds and then assess how various researcher characteristics influence the speed with which they adopt the identification technology.

2.5.1 Aggregate Benefits

Theory predicts that an agent will adopt a new technology when the (present value of the) expected net gains from adopting are positive. So, to understand agents' decisions and the observed diffusion pattern more broadly, we need to be able to estimate expected net gains for each agent. This is notoriously difficult because the factors (and weight) that enter each agent's utility function are not observed by the econometrician (see Foster and Rosenzweig [2010] for more details about measurement issues). Although we are not able to fully overcome this challenge, we can examine whether the use of the identification technology is associated with tangible aggregate benefits measured across all researchers. To do so, we assume that researchers seek to maximize their lifetime scientific impact and recognition in the profession. Following Hamermesh and Pfann [2012], we thus focus on the number of peer-reviewed publications and citations as metrics for scientific impact. We offer a two-pronged approach to approximate the aggregate net gains from adoption and their evolution over time.

2.5.1.1 Citations

We first concentrate on citations and examine whether articles using the identification technology display a different citation pattern compared to non-identification articles. We want to assess whether researchers in a given year T could have been aware of the difference in citations attracted by identification papers (i.e. the potential benefits).

To do so, we track the citations that identification papers obtain before T, and compare them to that of matched articles that do *not* use the identification technology. We follow Azoulay, Stuart, and Wang [2014] and select a single matched article for each identification article. We use a non-parametric approach designed to guarantee balance on crucial covariates. Specifically, for each identification article, we define the set of potential matches as all the non-identification articles published in the same journal, year, and quarter to ensure that we compare articles of the same vintage. We further restrict to articles where the most prolific author, measured using the number of previous publications, is in the same quantile as the focal identification article. If there is more than one article in this set, we select the non-identification article whose authors have the closest number of accrued citations. Any remaining ties are broken randomly. This procedure generates matches for 75% of identification articles.[20]

[20] This non-perfect matching is expected because our non-parametric matching is prone to a "curse of dimensionality", where the proportion of matched articles decreases with the number of strata that are imposed. In an unreported analysis, we find similar results if we only match on journals and semester. In addition, we alternatively consider a series of expanding regressions where we regress, for each article, the (logarithm of the) number of citations obtained in a year on a dummy variable indicating whether an article uses the identification technology as well as control variables (log of the maximum of the log of accrued citations across the article's authors, the number of authors, the log of the number of pages, as well as fixed effects corresponding to the interaction terms of the year of publication and age of the paper fixed effects). These regressions

[Insert Figure 2.2 about here]

Panel A of Figure 2.2 plots the average difference in (log of) citations between pairs of identification and matched articles for each period starting in 1980 and ending in year T. The shaded area on the graph corresponds to the 95% confidence bounds, where the standard errors are clustered by match-pair. The expanding window nature of our procedure means that the furthest point to the right covers the entire sample from 1980 to 2012. Over the full sample, identification articles attract 22% more citations than matched non-identification articles. This identification "premium" is statistically significant with a t-statistic of 3.40. The premium was first significantly positive in finance in 1995, suggesting the use of the technology translates into some benefits for researchers and that such benefits were apparent as early as the mid-1990s.[21]

We repeat a similar analysis in Panel B of Figure 2.2 for articles published in the leading economics journals (AER, JPE, QJE) but define identification articles using the presence of the identification "language" (see Section 2.4). Interestingly, we observe an identification premium in economics that started as early as 1982. This result is consistent with the earlier adoption patterns of the identification technology observed in the economics journals (see Sections 2.4.2.4.2 and 2.4.2.4.3). Moreover, the magnitude of the identification premium in economics articles is roughly similar to that estimated for empirical corporate finance articles.[22]

use all of the articles in the sample. The results and conclusions are qualitatively similar.

[21] In an unreported analysis, we obtain a very similar pattern if we focus on differences in citations (in levels) instead of differences in the logarithm of citations.

[22] Note however that the larger number of articles in economics journals provide more precise

2.5.1.2 Editorial Boards

To measure the net gains of adopting the identification technology, we would ideally want to gauge whether identification articles are more likely to get published. However, because we only observe published articles (and not all submitted articles), we cannot estimate the likelihood of getting published in finance as a function of whether or not a researcher adopts the technology. Instead, we provide indirect evidence on the likelihood of getting published by focusing on the characteristics of editorial board members at the leading finance journals. We conjecture that the odds of publication for articles using the identification are higher when a larger fraction of editors have already adopted the technology in their own research. The motivation behind this hypothesis is that adopting editorial board members are more likely to be familiar with the identification technology and view more positively studies that also use it.

[Insert Figure 2.3 about here]

We gather information on editors and associate editors directly from the journals' websites (for the RFS and JFE) and annual activity reports (for the JF).[23] The data is available starting in 1990 for the RFS and 1997 for the JFE and JF and includes information on 269 individuals. We define both editors and associate editors as editorial board members and (manually) link each member to our sample of all authors. Then, for each member-year observation, we create a dummy vari-

estimates of the identification premium.

[23] The JFE website only reports the current editorial board, so we use http://archive.org/web/ to retrieve annual snapshots of the website.

able that equals one if the board member has published at least one identification article by that year and zero otherwise.[24] Figure 2.3 displays a substantial change in the composition of journals' editorial boards over our sample period. We find no adopting editors before 1997. The fraction of adopting editors climbs above 20% in 2007 and 40% in 2011 and reaches 52% in 2012. This compositional change is sizable and has occurred across all three top finance journals. This shift in board composition towards identification techniques confirms the widespread diffusion in the field and is consistent with increased benefits for researcher adopting the identification technology over time.

2.5.2 Characteristics of Researchers

Next, we exploit the fine granularity of our sample to examine researcher-level determinants of adoption. We specifically focus on factors that can influence a researcher's awareness about the technology and its expected net benefits, as well as factors related to the cost of adopting it: (1) the doctoral training of researchers, (2) their research network, (3) the ranking of their doctoral and current institution, and (4) their seniority in the profession.

We rely on a survival analysis to shed light on the "time-to-adoption" among finance researchers. We use a linear model to estimate the factors that drive the *timing* of adoption for each individual researcher and identifies whether a given factor

[24] Note that this definition of adopting editors provides a lower bound estimate of adopting editors, since we only use ECF articles in the top-three finance journals to detect the use of identification techniques in their research.

induces faster adoption. For each author-year, the dependent variable (adoption) is equal to zero before the author adopts the identification technology (i.e., publishes her first ECF identification article in a finance journal), and is equal to one if the author adopts the identification technology in a given year.[25] Once an author adopts, we exclude her from the panel as is standard in survival analysis.

The panel of author-years comprises the career years of all authors who publish at least one empirical corporate finance article in the JF, JFE, or RFS between 1980 to 2012. To capture the publishing career of each author, we rely on the detailed biographical information we collect on PhD year and employment. Each author enters the panel the year when their publishing career begins as indicated by the first of the following three events: (1) PhD graduation year; (2) First academic position; (3) First publication in top-three finance journal. We use this starting date for each author to compute their "professional" age. An author leaves the sample when their publishing career ends as indicated by the last of the following two events: (1) Last academic position ends (retirement); (2) Last publication in top-three finance journal.[26] We drop years before 1980 and the sample ends in 2012, so there is left and right censoring in the data.

Because many authors' characteristics mechanically increase over time (e.g.

[25]In unreported test, we define the year of adoption based on publications in ECF articles *as well as* articles in the top-three economics journals (AER, JPE, and QJE). The inferences are unchanged.

[26]The results are robust to altering the conditions based on the year of first and last publication. For example, the results are unchanged if we define the entry year as the first publication year minus three years.

seniority or the set of former colleagues), we saturate our linear model with age, cohort, and calendar year fixed effects. This specification guarantees that our estimates truly isolate the effect of researcher-specific factors on adoption time from any age, cohort, or calendar year effects (see Hall, Mairesse, and Turner [2015] for a discussion of age-cohort-time effects).

2.5.2.1 Doctoral Training

We first consider the type of doctoral degree of each researcher. Based on detailed (hand-collected) biographical information for 1,610 scholars, we create a dummy variable that equals one if a researcher holds a PhD degree in economics. The fraction of economic PhDs in the sample is 31.5%. Results reported in the first column of Table 2.5 reveal a positive association between adoption and a PhD degree in economics.[27] The estimated coefficient is 0.007, indicating that the adoption of economics PhDs is almost 30% stronger than the average researcher (as the unconditional adoption rate in the panel is 0.024). In terms of actual time-to-adoption, our estimates imply that a median researcher with an economics PhD adopts about 2.3 years faster than the same researcher with any other PhD degree. In particular, the average cohort of economics PhDs reaches the 20% adoption threshold in 12.8 years, compared to 10.5 years for researchers with non-economics PhDs.

[Insert Table 2.5 about here]

[27] Non-economics PhDs are coined Finance PhDs in the text for simplicity although we recognize that researchers publishing in empirical corporate finance might have non-economics PhDs from other fields (e.g. accounting, business administration, psychology, mathematics, physics, etc.).

As documented earlier, the identification technology emerged in economics roughly 20 years prior to its emergence in finance. Hence, doctoral training in economics is likely to have exposed and informed researchers about the identification technology both earlier and more prominently. The quicker adoption of economics PhD holders is thus consistent with the idea that increased exposure improves researchers' awareness and inference about the new technology. In addition, estimates from Table 2.5 indicate that the diffusion of innovation across research fields accelerates when researchers with a given training migrate to a neighboring field of research. This result is in line with Stoyanov and Zubanov [2012] which shows that labor mobility produces knowledge spillovers across firms and Azoulay, Graff Zivin, and Manso [2011] which documents similar knowledge spillovers across scientists in the life-sciences.

2.5.2.2 Academic Networks

Next, we consider researchers' academic network. Arguably, knowledge about the expected net gains associated with using the identification technology could be facilitated by interactions with other researchers. As proxies for such interactions, we define three types of social networks. First, we define a network of colleagues (*Adopting Colleagues*) by tracking employment history so as to link a researcher to all other researchers located at the same institution during his tenure at the institution. The network is time-varying and is composed of all past and current colleagues as of year t. We also define two other networks based on alumni ties. In particular, we define a second network that comprises all researchers that receive a PhD from the same PhD granting institution (*Adopting Alumni*) and a third network that links each researcher to all researchers working at his PhD granting institution in the two years preceding his PhD graduation year (*Adopting PhD Faculty*).[28] Because we want to examine the influence of a given researcher's network on his decision to adopt in a given year t, we compute the fraction of a researcher's network members that have adopted the identification technology by the prior year $(t-1)$.

The estimations displayed in column (2) of Table 2.5 indicate that the adoption of the identification technology is related to social interactions among academics. We estimate a faster adoption when a larger fraction of a researcher's academic network has already adopted. This result holds for networks of current and former colleagues, as well as network of alumni. The estimated coefficients for the network variables range between 0.003 and 0.011. Given that each network variable is normalized by its standard deviation, these estimates indicate adoptions of the identification technology that are 30% to 45% faster in response to a one standard deviation increase in the adoption rate of the members of their academic networks.

The positive link between a researcher's adoption and that of its peers is consistent with researchers' networks creating positive externalities that facilitate the diffusion of knowledge among individuals. For instance, Azoulay, Graff Zivin, and

[28]We manually obtain the career path and institutions of employment for 1,554 authors. For the remaining 246 authors, we infer the institutions from the author information on each publication. We fill years in which an author does not publish by carrying forward all institutions from the last year the author published. We carry back the institution of the first publication.

Wang [2010] finds that co-authorship ties play an important role in researchers productivity, and Head, Li, and Minondo [2015] shows that academic linkages facilitate knowledge flows (measured using theorem citations) among mathematicians. Our results suggest that academic networks significantly reduce the informational constraints associated with learning about the net gains of the new technology and foster the diffusion of innovation.

2.5.2.3 Institutions' Rankings

We next examine whether adoption varies across school rankings. We follow the classification of Kim, Morse, and Zingales [2009] to define the top-25 academic institutions. When we focus on the ranking of a researcher's PhD-granting institution, the results reported in the second row of Table 2.5 reveal no evidence that the school ranking of a researcher's PhD-granting institution is related to his time-to-adoption. While we find a strong link between a researcher's alumni-based network and his adoption, adoption appear similar for researchers holding PhDs from top-tier institutions than for researchers with PhDs from lower-ranked institutions.

In sharp contrast, estimates in the third row reveal that researchers adopt the identification technology significantly faster when they are employed at top-tier institutions. The estimated coefficients indicate a difference in adoption speed between 11% (column (2)) and 21% (column (1)) when we compare researchers at the top institutions to researchers in lower-ranked institutions. Kim, Morse, and Zingales [2009] show that since the nineties, researchers at top institutions are not associated with more productive research. Yet, we find that these researchers have adopted the identification technology significantly faster.

This positive effect of institution ranking is strong and could be consistent with several explanations. Although we focus on a soft innovation that requires limited costly research equipment (e.g. a lab or a supercomputer), the larger financial or organizational resources allocated to research in top schools could facilitate the faster adoption of new techniques. Alternatively, differences in research incentives and tenure requirements could induce researchers at top schools to be more innovative and adopt new techniques faster. On the other hand, the relationship between ranking and adoption may go the other way; our estimate could simply indicate that top schools are better at selecting future innovators and early adopters.

2.5.2.4 Seniority

We find evidence that researchers' seniority and status are significantly related to the diffusion of the identification technology. In columns (3) and (4) of Table 2.5, we include authors' career age (and its squared term) as explanatory determinants of adoption speed and limit the set of fixed effects to cohort and calendar years. We observe a negative association between age and adoption time; more senior authors adopt the technology later. Moreover, results in column (4) indicate that the age-adoption relation is concave, with an implied maximum adoption rate approximately nine years into a researcher's career.

This seniority effect is consistent with models of vintage human capital implying that early-stage researchers are the primary adopters of new technology (e.g. Chari and Hoenhayn [1991] or MacDonald and Weisbach [2004]). Several theories could lead to lower adoption costs for younger, less-established researchers. Lower costs could arise from more flexible minds of young scholars (e.g. Darwin [1859]), their larger exposure to recent innovation, higher incentives, and more time to learn new techniques (Diamond [1980]). Alternatively, more senior scholars could have more vested interests that may render them less receptive to innovation (e.g. Cohen [1985]).

Remarkably, the association between age and adoption mirrors the negative link between adoption time and authors' number of citations observed in every specification of Table 2.5. Insofar as citations take time to accumulate, citation count is likely higher for more senior researchers. Yet, citations is also a commonly-used metric to identify "high-status" researchers. In an attempt to disentangle the effects of seniority and status on adoption, we define a subset of researchers as exhibiting high-status if they have received a best-paper award from one of the top journals, or are fellows of the American Finance Association.[29] Results in column (5) confirm that seniority and status are distinct, as we estimate a faster adoption for the subset of high status researchers.

[Insert Table 2.6 about here]

[29] We consider the Smith-Breeden (1989-) and Brattle Group (1999-) Prizes for the *Journal of Finance*, the Jensen (1997-) and Fama-DFA (1997-) Prizes for the *Journal of Financial Economics*, and the Michael J. Brennan (2000-) and Young Researcher (2006-) Prizes for the *Review of Financial Studies*. There are 81 researchers in the high-status category.

To better understand the role of seniority in the diffusion of the identification technology, Table 2.6 provides an analysis of coauthorship teams across different career age groups. In Panel A, we examine how seniority relates to solo-authored identification articles. Columns under "Authorship" report the number of authorships by age group for all published articles ("All"), identification articles ("ID"), all solo-authored articles ("Solo"), and solo-authored identification articles ("Solo ID").[30] Columns under "Percent by Career Age" report the likelihood of a solo-authored identification article among three different subsets of publications. The fraction of solo-authored identification articles among all authorships ("Solo ID/All") is 0.2% for authors with more than 20 years seniority and 3.3% for authors in the first 10 years of their career. The difference in proportions of -3.0% is statistically significant. Furthermore, the difference in proportions of identification articles being solo-authored ("Solo ID/ID") between the senior (row 3) and junior (row 1) authors is -12.2% and statistically significant. The magnitude in the last column, which reports the fraction of identification articles within solo-authored articles, is similar but not significant. These results add to the findings in Table 2.5 by showing that senior researchers are more likely to adopt by co-authoring papers.

Panel B of Table 2.6 investigates the co-authorship structure across researchers at different stages of their careers. Each row computes the average career age difference (between an author and their coauthors) for a given career age category. The first row of Panel B focuses on authors with a career age of less than 10 years.

[30] Authorship is credited per author per paper. For instance, two authors on one paper corresponds to two authorships.

Given the positive age gap we find, we conclude, perhaps not surprisingly, that junior authors tend to collaborate with more senior authors. More interestingly, the average age gap relative to coauthors is greater for published work using the identification technology (7.3 years vs. 6 years) and the difference is statistically significant. Corroborating those results, we find that senior researchers (as measured by a career age exceeding 20 years) have younger coauthors on average. Further, we also find a significantly greater average age gap for identification articles compared to non-identification articles (-12.2 years vs. -10.7). While senior researchers typically co-author with more junior researchers, they team up with even younger researchers when they adopt the identification technology.

2.5.3 Straddling Authors

The aim of this section is to provide a more complete picture of adoption patterns by identifying authors who publish at least one empirical corporate finance article in a finance journal and one article in a economics journal during our sample period. We define these authors as "straddlers". Given the general pattern of diffusion from the field of economics to the field of finance, we expect these straddlers to play a vital role in the transfer of the technology across the two fields. Further, because our estimates of authors' time-to-adoption are based on articles published in finance journals, our interpretation could potentially be biased if a large number of authors in our sample adopt the identification technology earlier, but do so in articles published in the top-three economics journals. To assess the robustness of our findings, we investigate these straddlers and their impact in the diffusion process below.

Table 2.7 presents the characteristics of straddlers. Overall, we identify 291 authors with publications in both disciplines. They represent 16.3% of our sample of authors. For each straddler, we further check whether the article(s) published in the economics journals (a) uses the identification technology, and (b) is in empirical corporate finance (using the same classification as in Section 2.3.2.3.1). Panel A reveals that straddlers are more likely to adopt. While 36.4% of all authors adopt the technology in our sample, the proportion rises to 56.5% among straddlers. Importantly, with regard to potential biases, only a small number of straddlers adopt the technology in an economics article prior to adopting it in a finance article (32 authors). An even smaller number of straddlers (8 authors) adopt the technology in an empirical corporate finance article published in a top economics journal before adopting it in a finance journal. Ultimately, among all authors who adopt the identification technology in finance journals, only a handful adopted the identification technology in published work outside of finance journals beforehand. This result dispels the bias concerns raised above.

[Insert Table 2.7 about here]

Panel B reveals that straddlers are significantly more likely to have PhD degrees in economics and degrees from top schools. The former result is to be expected given that we condition our sample of straddlers on having at least one publication in the top-three economics journals (in addition to having a publication in the top-three finance journals).

In Panel C, we show that straddlers adopt sooner the identification technology than non-straddlers. Specifically, when considering publications in both sets of journals, the average year of adoption among straddlers is 2003 compared to 2007 for non-straddlers. Even considering adoption *only* within finance journals, straddlers adopt earlier (2005 vs. 2007). Both differences are statistically significant. The average year of adoption for straddlers is slightly earlier in economics journals (2002) relative to finance journals (2005). Results from Panel C confirm that straddlers are early adopters of the identification technology and hence play a potentially important role in transferring the identification technology to the field of finance.

We test further this conjecture in a multivariate regression of time-to-adoption that controls for age, cohort, and year fixed effects. The last column of Table 2.5 replicates the baseline time-to-adoption regression reported in column (1) but adds binary variables to capture the notion of straddlers. Including directly a straddler dummy variable, however, induces substantial look ahead bias. Thus, we define a binary variable, *Prev Econ*, which is equal to one if the author has an article published in one of the top economics journals in previous years. We also define *Prev Econ-ID*, which is equal to one if any of the previous economics articles also used the identification technology. Consistent with the idea that straddling authors represent an important vehicle of technology diffusion, we observe a significantly faster time to adoption for authors with previous identification articles published in economics journals. Importantly, we also see that the estimates of the other determinants are unaltered.

2.6 Technological Refinements

Our analysis so far has focused on a “static” definition of innovation. However, innovation is by nature a dynamic concept. As with other technologies, we expect the identification technology to evolve over time as more researchers learn its benefits and limitations. Epistemologists have long argued that scientific progress evolves through experimentation or trial-and-error. As a consequence, scientific norms (or what could be considered “best practices”) also evolve over time as more scientists use and criticize the new techniques (see Popper [1959] and Kuhn [1962]). To buttress the dynamic nature of the identification technology, we provide evidence of the changing nature of “standard” tests, diagnostics, and procedures performed when implementing identification techniques.

In Figure 2.4, we focus on the tests and diagnostics that are performed when applying the two most used identification techniques: Instrumental Variables (IV) and Difference-in-differences (DD). We track IV and DD articles over time and verify whether they contain keywords related to diagnostic tests recommended in survey articles such as Roberts and Whited [2013]. For IV articles, we count the number of articles that mention the following terms: “weak instrument”,“Stock and Yogo”, “overidentification”, “exclusion restriction”, or “Hausman test”. For DD articles, we count the number of articles that mention “falsification test”, “parallel trend”, “placebo test”, or “Bertrand, Duflo, and Mullainathan [2004].” Common permutations of each keyword are allowed.

[Insert Figure 2.4 about here]

Figure 2.4 reveals a slow and gradual implementation (or discussion) of these diagnostic tests. For instance, whereas IV techniques emerge in the mid-nineties, discussion of "weak instrument" or "overidentification" come about 10 years later. Similarly, whereas difference-in-differences appear around 2000, the discussion of "parallel trend" within ECF papers first appears in 2007. Overall, this gradual emergence of different diagnostic tools over our sample period suggests that researchers are learning-by-doing and refine the implementation of these identification techniques over time (see the surveys by Hall and Khan [2002] or Foster and Rosenzweig [2010] for a detailed discussion of learning-by-doing). More generally, the patterns in Figure 2.4 are consistent with changing scientific "norms" over time.

As put forth by Kuhn [1962], such a changing nature constitutes a "healthy" scientific process, whereby new techniques and discoveries are first introduced, then scrutinized and criticized, and as a result improved. The identification technology is no exception. For instance, several recent papers cast doubt on the merits and appropriateness of the identification technology (e.g. Hennessy and Strebulaev [2015]), while others evaluate whether the implementation of the technology is performed "correctly" in practice (e.g. Atanasov and Black [Forthcoming]). This line of research is highly constructive as it provides a description of *current* best practices. Yet, as we document in this section, the evolving nature of scientific discoveries make it difficult to pass judgment on the merits of a particular application at any given point in time.

2.7 Conclusion

This study documents a secular rise of the identification technology in corporate finance research as of the mid-nineties. Although this rise lags the emergence of these techniques in the top economics journals by approximatively 20 years, the adoption rate is steeper in finance, with more than 50% of all published articles in empirical corporate finance having adopted the technology by 2012. We find significant heterogeneity in the speed of adoption across finance researchers. Consistent with this choice being driven by each individual's awareness of the technology and its net benefits, we find that scholars that are less senior, hold a PhD in economics, and work for the top institutions display faster adoption. In addition, we find evidence that a researcher's adoption is related to that of his research networks composed of colleagues and alumni sharing the same alma mater. Our findings highlight new forces that facilitate the diffusion of innovation in academic research.

Our results suggest several interesting avenues for future research. In particular, it would be interesting to examine how the identification technology disseminates into other related research fields, such as accounting, strategy, or management, and estimate general adoption lags across disciplines. More generally, it would be interesting to study the diffusion of other research tools (e.g. textual analysis or computing power) and assess whether similar economic forces are at play.

From a practical perspective, our findings point to important changes in the publication norms in the field of corporate finance. These changes can have potentially important implications, particularly for the career path of young scholars.

Competition for space in the top journals has grown fiercer over time, making it increasingly difficult for young scholars to achieve a given set of publication benchmarks.[31] At the same time, editorial boards are now composed of researchers that have adopted the identification technology. Young authors, in particular those with training in economics and those favored by top schools, seem to have largely responded to the incentives created by this new environment by adopting the identification technology much faster in recent years.

[31]Based on information provided by the journals, the overall acceptance rates for submission at the top-three journals is about one third as high today as it was in the early nineties.

Appendix

C.1 Detailed Classification

This Appendix details our procedure to identify articles that use the identification technology. Table C.1 presents the list of keywords we use for each of the five identification techniques we use to flag articles. Below, we describe the criteria we use to assign flagged articles into each category, and remove all false positives (i.e. flagged articles that do *not* use identification techniques). Table C.2 lists the keywords that we use to define the language associated with the identification technology.

- **Instrumental Variables**. The IV designation requires an article to apply an IV, 2SLS, 3SLS, or GMM estimation and show the results in the article. 215 articles satisfy this requirement. Roberts and Whited [2013] cite Giroud, Mueller, Stomper, and Westerkamp [2012] as a successful implementation of IV in corporate finance. An additional 51 articles are also assigned to the IV category for the following reasons: (1) Untabulated IV results, (2) reduced form IV approach, (3) IV approach without disclosing the instruments used, (4) unusual procedure claimed to be equivalent to IV by the authors, and (5) non-valid instruments as stated by the authors themselves. We find a total of 266 IV articles.

- **Difference-in-Differences**. The DD category includes articles that perform a difference-in-differences estimation as defined in Roberts and Whited [2013]. Examples include Agrawal [2013] and Gormley and Matsa [2011]. The DD classification further includes articles that make use of an exogenous shock as a natural experiment or quasi-experimental framework outside of a difference-in-differences framework. We find a total of 126 DD articles.

- **Selection Models**. The Selection category requires that an article implements a selection model following the traditional two-stage approach of the Roy model or the Heckman model as presented in Li and Prabhala [2007] and report the estimates. Examples include Campa and Kedia [2002] and Bris, Welch, and Zhu [2006]. The Selection category further contains articles that either discuss untabulated results or use a non-standard technique to perform an estimation that corrects for self-selection biases. We find a total of 72 Selection articles.

- **Regression Discontinuity Design**. The RDD category requires that an article explicitly performs a RDD estimation as defined in Roberts and Whited [2013]. Examples include Chava and Roberts [2008] and Roberts and Sufi [2009]. We find a total of 25 RDD articles.

- **Randomized Experiments**. The Randomized Experiment category requires that an article explicitly perform a field or laboratory experiment with a treatment and control condition. Examples include Cox, Smith, and Walker [1984] and Alevy, Haigh, and List [2007]. We find a total of 5 Randomized Experiment articles.

Table C.1: Classifying Identification Articles
We manually identify articles that adopt the identification technology in Section 2.3.1. This table presents the search terms that we use to flag articles for manual inspection. We explicitly provide plural terms and hyphenation during the search process. For brevity, we list here permutations within parentheses.

Technique	Key words used in search
IV	instrumental variable(s); 2sls; two(-)stage(s) least square(s); 3sls; three(-)stage(s) least square(s)

DD	difference(s)(-)in(-)difference(s); diff-in-diff;
	quasi(-)natural; natural experiment(s);
	exogen(e)ous(-)shock(s)

Selection	self(-)selection; switching(-)regression(s)
RDD	regression discontinuity; regression discontinuities

Experiments	random(ized) experiment(ation)(s);
	lab(oratory) experiment(ation)(s);
	field experiment(ation)(s);

Table C.2: Identification Language Terms

In Section 2.4, we compare the adoption rate of the identification *language* in the top-three finance journals (JF, JFE, RFS) to that of the top-three economics journals (AER, JPE, QJE). We define an article as having adopted the identification language if at least one of the terms in the dictionary below or at least one of the terms in the dictionaries listed in Table C.1 is found in the text. We explicitly provide plural terms and hyphenation during the search process. For brevity, we list here permutations within parentheses.

Key words used in search
causal affect(s); causal effect(s); causal impact(s)
exogen(e)ous variation
reverse causality; reverse causation
identification strategy; identification strategies

Bibliography

Phlippe Aghion and Peter Howitt. A model of growth through creative destruction. *Econometrica*, 60(2):323–351, 1992.

Ajay Agrawal and Avi Goldfarb. Restructuring research: Communication costs and the democratization of university innovation. *American Economic Review*, 98(4): 1578–1590, 2008.

Ashwini Agrawal. The impact of investor protection law on corporate policy and performance: Evidence from the blue sky laws. *Journal of Financial Economics*, 107(2):417–435, 2013.

Jonathan E Alevy, Michael S Haigh, and John A List. Information cascades: Evidence from a field experiment with financial market professionals. *The Journal of Finance*, 62(1):151–180, 2007.

John Alexander and Rodney Marby. Relative significance of journals, authors, and articles cited in financial research. *Journal of Finance*, 49(2):697–712, 1994.

Joshua Angrist and Jrn-Steffen Pischke. The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2):3–30, 2010.

Vladimir Atanasov and Bernard Black. Shock-based causal inference in corporate finance and accounting research. *Critical Finance Review*, Forthcoming.

Pierre Azoulay, Joshua S Graff Zivin, and Jialin Wang. Superstar extinction. *The Quarterly Journal of Economics*, 125(2):549–589, 2010.

Pierre Azoulay, Joshua S Graff Zivin, and Gustavo Manso. Incentives and creativity: Evidence from the academic life science. *Rand Journal of Economics*, 42(3):527–554, 2011.

Pierre Azoulay, Joshua S Graff Zivin, and Bhaven Sampat. The diffusion and scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine. In Joshua Lerner and Scott Stern, editors, *The Rate and Direction of Inventive Activity Revisited*, chapter 2, pages 107–155. University of Chicago Press, 2012.

Pierre Azoulay, Toby Stuart, and Yanbo Wang. Matthew: Effect or fable? *Management Science*, 60(1):92–109, 2014.

Ravi Bansal and Amir Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4):1481–1509, 2004.

Patrick Bayer, Randi Hjalmarsson, and David Pozen. Building criminal capital behind bars: Peer effects in juvenile corrections. *Quarterly Journal of Economics*, 124(1):105–147, 2009.

Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. How much should we trust difference-in-differences estimates. *Quarterly Journal of Economics*, 119(1): 249–275, 2004.

Kenneth Borokhovich, Robrt Bricker, Keyy Brunarski, and Betty Simkins. Finance research productivity and influence. *Journal of Finance*, 50(5):1691–1717, 1995.

Arturo Bris, Ivo Welch, and Ning Zhu. The costs of bankruptcy: Chapter 7 liquidation versus chapter 11 reorganization. *Journal of Finance*, 61(3):1253–1303, 2006.

Jonathan Brogaard, Joseph Engelberg, and Christopher Parsons. Networks and productivity: Causal evidence from editor rotations. *Journal of Financial Economics*, 111(1):251–270, 2014.

Jose Manuel Campa and Simi Kedia. Explaining the diversification discount. *Journal of Finance*, 57(4):1731–1762, 2002.

Murillo Campello, Antonio Galvao, and Ted Juhl. Policy heterogeneity in empirical corporate finance. *Working Paper. Cornell University*, 2013.

David Card and Stefano DellaVigna. Nine facts about top journals in economics. *Journal of Economic Literature*, 51(1):144–161, 2013.

V. Chari and H. Hoenhayn. Vintage human capital, growth, and the diffusion of new technology. *Journal of Political Economy*, 99(6):1142–1165, 1991.

Sudheer Chava and Michael Roberts. How does financing impact investment? the role of debt covenant violations. *Journal of Finance*, 63(5):2085–2121, 2008.

Bernard Cohen. *Revolution in Science*. Harvard University Press: Cambridge, MA., 1985.

Diego Comin and Bart Hobijn. An exploration of technology diffusion. *American Economic Review*, 100(5):2031–2059, 2010.

Charles Corrado and Stephen Ferris. Journal influence on the design of finance doctoral education. *Journal of Finance*, 52(5):2091–2102, 1997.

James C Cox, Vernon L Smith, and James M Walker. Theory and behavior of multiple unit discriminative auctions. *The Journal of Finance*, 39(4):983–1010, 1984.

Charles Darwin. *On the origin of species.* John Murray: London, 1859.

Arthur Diamond. Age and the acceptance of cliometrics. *Journal of Economic History*, 40(4):838–841, 1980.

Waverly Ding, Sharon Levin, Paula Stephan, and Anne Winkler. The impact of informaton technology on academic scientists' productivity and collaboration patterns. *Management Science*, 56(9):1439–1461, 2010.

Glenn Ellison. Evolving standards for academic publishing: A q-r theory. *The Journal of Political Economy*, 110(5):994–1034, 2002.

Andrew Foster and Mark Rosenzweig. Microeconomics of technology adoption. *Annual Review of Economics*, 2:395–424, 2010.

Jeffrey Furman and Scott Stern. Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *aer*, 101(5):1933–1963, 2011.

Xavier Giroud, Holger Mueller, Alex Stomper, and Arne Westerkamp. Snow and leverage. *Review of Financial Studies*, 25(3):680–710, 2012.

Todd Gormley and David Matsa. Growing out of trouble? corporate responses to liability risk. *Review of Financial Studies*, 24(8):2781–2821, 2011.

Todd Gormley and David Matsa. Common errors: How to (and not to) control for unobserved heterogeneity. *Review of Financial Studies*, 27(2):617–661, 2014.

Zvi Griliches. Hybrid corn: An exploration in the economics of technological change. *Econometrica*, 25(4):501–522, 1957.

B. Hall and B. Khan. Adoption of new technology. In D. Jones, editor, *New Economy Handbook*, pages 5–20. Elsevier Science, Amsterdam, The Netherlands, 2002.

Bronwyn Hall, Jacques Mairesse, and Laure Turner. Identifying age, cohort, and period effects in scientific research productivity: Discussion and illustration using simulated and actual data on french physicists. *Economics of Innovation and New Technology*, 16(2):159–177, 2015.

Daniel Hamermesh and Gerard Pfann. Reputation and earnings: The roles of quality and quantity in academe. *Economic Inquiry*, 50(1):1–16, 2012.

Daniel S Hamermesh. Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, 51(1):1–11, 2013.

Daniel S Hamermesh, George E Johnson, and Burton A Weisbrod. Scholarship, citations and salaries: Economic rewards in economics. *Southern Economic Journal*, 49(2):472–481, 1982.

Keith Head, Yao Amber Li, and Asier Minondo. Geography, ties, and knowledge flows: Evidence from citations in mathematics. *Working Paper. University of British Columbia.*, 2015.

Christopher Hennessy and Ilya Strebulaev. Beyond random assignment: Credible inference of causal effects in dynamic economies. *NBER Working Paper No. 20978*, 2015.

Benjamin Jones. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Review of Economic Studies*, 76(1):283–317, 2009.

Andrew Karolyi. The ultimate irrelevance proposition in finance? *Financial Review*, 46(4):485–512, 2011.

Andrew Karolyi. Home bias, an academic puzzle? *Review of Finance*, Forthcoming.

George Kaufman. Rankings of finance departments by faculty representation on editorial boards of professional journals: A note. *Journal of Finance*, 39(4):1189–1197, 1984.

Michael Keenan. Fifty years of the American Finance Association. *Journal of Finance*, 46(3):1125–1156, 1991.

Han Kim, E, Adair Morse, and Luigi Zingales. Are elite universities losing their competitive edge? *Journal of Financial Economics*, 93(3):353–381, 2009.

Michael Kremer and Edward Miguel. The illusion of sustainability. *Quarterly Journal of Economics*, 122(3):1007–1065, 2007.

Thomas Kuhn. *The Structure of Scientific Revolutions*. The University of Chicago Press, 1962.

K Li and N Prabhala. Self-selection models in corporate finance. In B Espen Eckbo, editor, *Handbook of Empirical Corporate Finance*, chapter 2, page 37–83. Elsevier, Amsterdam, Netherlands, 2007.

Glenn MacDonald and Michael Weisbach. The economics of has-beens. *Journal of Political Economy*, 112(1 pt. 2):289–310, 2004.

Edwin Mansfield. Technical change and the rate of imitation. *Econometrica*, 29(4): 741–766, 1961.

Joel Mokyr. *The gifts of Athena: Historical origins of the knowledge economy.* Princeton University Press, 2002.

W. Moore, R. Newman, and G. Turnbull. Reputational capital and academic pay.

Economic Inquiry, 39(4):663–671, 2001.

Albert Niemi. Institutional contributions to the leading finance journals, 1975 through 1986: A note. *Journal of Finance*, 42(5):1389–1397, 1987.

Mitchell A. Petersen. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *Review of Financial Studies*, 22(1):435–480, 2009.

Karl Popper. *The Logic of Scientific Discovery.* London: Hutchinson, 1959.

Michael Roberts and Amir Sufi. Control rights and capital structure: An empirical investigation. *Journal of Finance*, 64(4):1657–1695, 2009.

Michael R Roberts and Toni M Whited. Endogeneity in empirical corporate finance, . In George Constantinides, René Stulz, and Milton Harris, editors, *Handbook of the Economics of Finance*, volume 2, Part A, chapter 7, pages 493–572. Elsevier, Amsterdam, Netherlands, 2013.

Everett Rogers. *Diffusion of Innovation.* Free Press: New York, 5th edition, 2003.

James Stock and Francesco Trebbi. Who invented instrumental variable regression? *Journal of Economic Perspective*, 17(3):177–194, 2003.

Andrey Stoyanov and Nikolay Zubanov. Productivity spillovers across firms through worker mobility. *American Economic Journal: Applied Economics*, 4(2):168–198, 2012.

Ilya A Strebulaev and Toni Whited. Dynamic models and structural estimation in corporate finance. *Foundations and Trends in Finance*, 6(1-2):1–163, 2012.

Florenta Teodoridis. Generalists, specialists, and the direction of inventive activity. *Working Paper. University of Southern California.*, 2015.

Samuel B Thompson. Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1):1–10, 2011.

John Weber. A note on keeping abreast if developments in the field of finance. *Journal of Finance*, 28(1):161–165, 1973.

Ivo Welch. Referee recommendations. *Review of Financial Studies*, 27(9):2773–2804, 2014.

Heidi Williams. Intellectual property rights and innovation: Evidence from the human genome. *Journal of Political Economy*, 121(1):1–27, 2013.

Philip Wright. *The Tariff on Animal and Vegetable Oils.* MacMillan, New York, 1928.

Terry Zivney and William Bertin. Publish or perish: What competition is really doing. *Journal of Finance*, 47(1):295–329, 1992.

Figure 2.1: Adoption of the Identification Technology

This figure presents the evolution over time of the fraction of empirical corporate finance (ECF) articles published in the JF, JFE, or RFS that use the identification technology as well as the fitted values obtained from estimating the baseline diffusion model specified in Equation (2.1) and estimated in column (1) of Table 2.2. The sample period is from 1980 to 2012. The five techniques that make up the identification technology are: (1) Instrumental variables; (2) Difference-in-differences; (3) Selection models; (4) Regression discontinuity design; and (5) Randomized experiments. The classification is detailed in Section 2.3.2. The solid line represents the actual fraction, while the dashed line represents the fitted fraction. We include four vertical lines to highlight the years where the fitted fraction crosses the 5%, 10%, 15%, and 20% adoption thresholds.

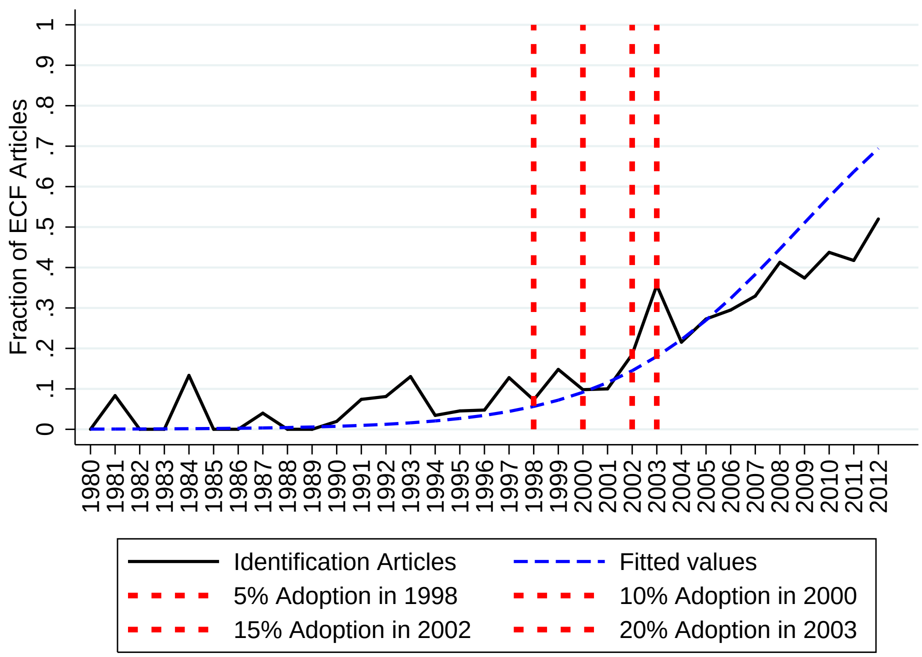
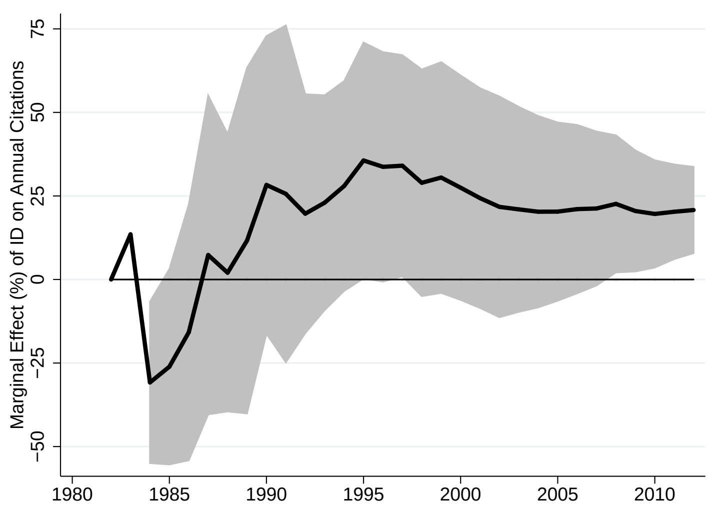


Figure 2.2: Evolving Relationship Between Citations and Adoption

This table presents the results of a citation analysis on two distinct samples. The sample in Panel A consists of all ECF identification articles published in the JF, JFE, and RFS over the period 1980-2012. The sample in Panel B consists of all identification articles published in the top-three economics journals (AER, JPE, and QJE) over the same period. Following Azoulay, Stuart, and Wang [2014], we select one matched non-identification article for each identification article. Details on the matching procedure can be found in Section 2.5.2. For each year t after an identification article i is published, we compute the difference, $\Delta_{i,t}$, in log of citations between the identification article and its match. For each year T between 1980 and 2012, we compute the average of $\Delta_{i,t}$, where $t \in [1980, T]$, and report this as the solid line below. The gray area represents the 95% confidence interval based on standard errors clustered by matched pair. In Panel A, an article is classified as an identification article as detailed in Section 2.3.2. In Panel B, an article is classified as an identification article if the article adopts the language of identification by using any of the identification technology keywords displayed in Table C.2. Citation data comes from Web of Knowledge.

Panel A: Empirical Corporate Finance



Panel B: Economics

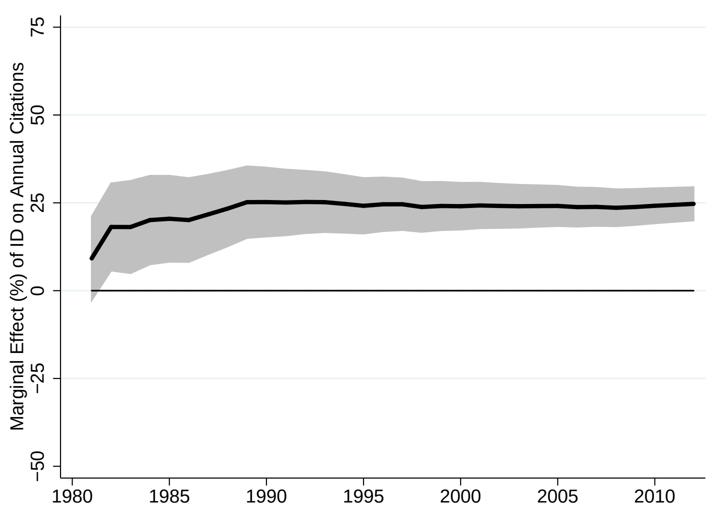


Figure 2.3: Evolution of Adopting Editors

This figure presents the evolution over time of the fraction of editorial board members (editors and associate editors) at the JF, JFE, or RFS that have adopted the identification technology ("Adopting Editors"). The sample period is from 1990 to 2012. We define a board member in a given year as an Adopting Editor if he or she has written at least one paper that uses the technology by that year. The classification of articles into identification and non-identification is detailed in Section 2.3.2.

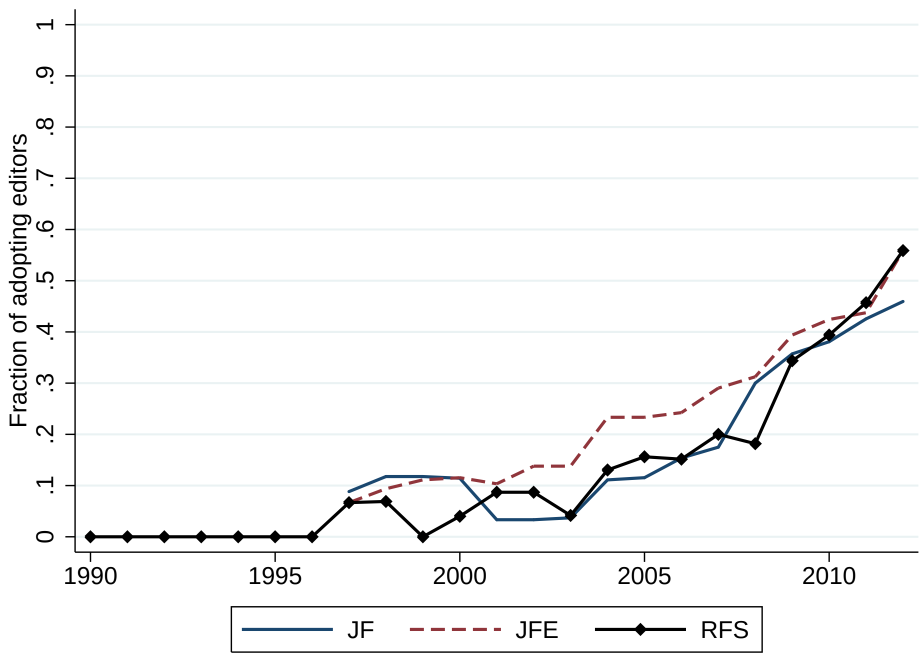
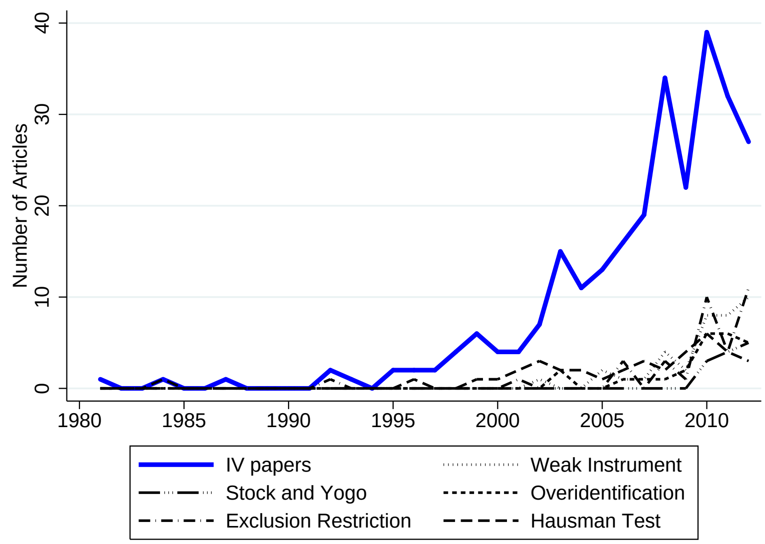


Figure 2.4: Technological Refinements
This figure presents the evolution over time of the number of ECF articles published in the JF, JFE, or RFS that adopt the instrumental variables ("IV") approach or difference-in-differences ("DD') approach, as well as the implementation of diagnostic tests associated with these techniques. The sample period is from 1980 to 2012. Each ECF article is classified into an IV or DD category, depending on whether it adopts the technique (see Section 2.3.2). We further search within these classified articles for keywords associated with diagnostic tests ("weak instruments","Stock and Yogo", "overidentification", "exclusion restriction", or "Hausman test" for IV; and "falsification test", "parallel trend", "placebo test", or "Bertrand, Duflo, and Mullainathan [2004]" for DD). Common permutations of each keyword are allowed. Panel A displays the prevalence of the IV technique (solid line) and its associated diagnostic terms (dashed lines). Panel B displays the prevalence of the DD technique (solid line) and its associated diagnostic terms (dashed lines).

Panel A: IV



Panel B: DD

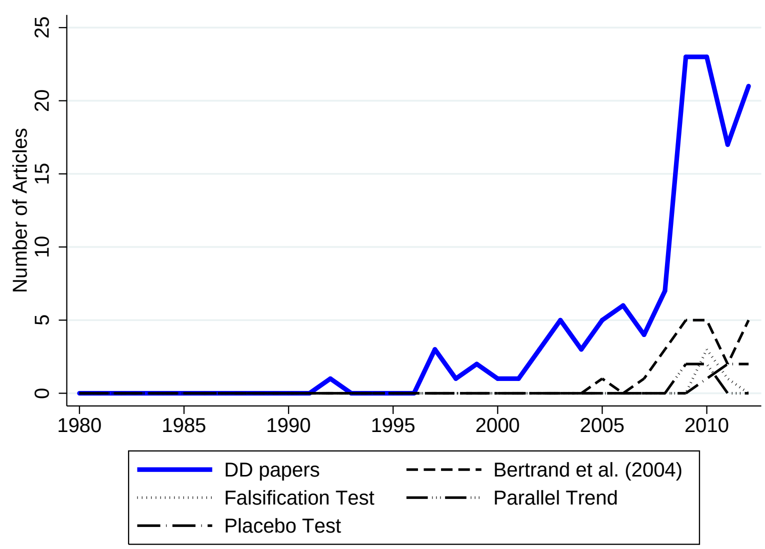


Table 2.1: Descriptive Statistics

The sample comprises all articles published in the Journal of Finance (JF), Journal of Financial Economics (JFE), and Review of Financial Studies (RFS) between 1970 and 2012. Panel A reports the number and fraction of articles that are classified as Empirical, Corporate Finance (CF), and Empirical Corporate Finance (ECF) by decade, as well as for the whole sample period (Section 2.3.1 provides more details). Panel B presents the use of the five techniques that compose the identification technology in ECF articles. Each ECF article is classified as having adopted or not these techniques based on a three-step procedure as detailed in Section 2.3.2. The columns report the number and fraction of ECF papers using these different identification techniques along with the year the technique was first used in ECF.

Panel A: Classification of Sample Articles

Classification:		Empirical		CF		ECF	
	Total	N	(%)	N	(%)	N	(%)
Total	6,797	4,889	72%	2,438	36%	1,796	26%
1970s	976	513	53%	220	23%	78	8%
1980s	1,228	732	60%	356	29%	226	18%
1990s	1,519	1,141	75%	534	35%	409	27%
2000s	2,196	1,766	80%	921	42%	740	34%
2010-2012	878	737	84%	407	46%	343	39%

Panel B: Use of Identification Technologies

	N	(% of ECF)	First Use
Any identification technology (ID)	408	22.7%	1977
Instrumental variables (IV)	266	14.8%	1977
Difference-in-differences (DD)	126	7.0%	1992
Selection models (Selection)	72	4.0%	1990
Regression discontinuity design (RDD)	10	0.6%	2006
Randomized experiments (Experiments)	5	0.3%	1984

Table 2.2: Adoption of the Identification Technology

This table presents the estimates and fitted values of the empirical model of technology diffusion (S-curve) presented in Equation (2.1). The sample corresponds to a time series of 33 annual observations computed from all ECF articles published in the JF, JFE, and RFS over the period 1980-2012. In column (1), the dependent variable (p_t) is the fraction of ECF articles classified as using the identification technology in year t. In columns (2) to (4), we examine separately the three most prominent identification techniques found in the sample (respectively IV, DD, Selection). The classification procedure is detailed in Section 2.3.2. Panels A, B, and C report the estimates and implied adoption thresholds for three different ceiling levels, $k = 1$, $k = 0.8$, and $k = 0.6$ respectively. An implied adoption threshold year is defined as the year in which the fitted values ($\tilde{p}$) cross a given fraction of the sample population (respectively 5%, 10%, 15%, 20%, 50%). For brevity, we report the estimates only for the specification in Panel A. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	ID	IV	DD	Selection
Panel A: Ceiling Parameter k = 1				
β	0.260***	0.263***	0.306***	0.229***
	(5.91)	(5.99)	(18.65)	(10.43)
α	-7.487***	-8.416***	-10.77***	-10.01***
	(-7.21)	(-7.96)	(-24.07)	(-18.12)
R^2	0.598	0.553	0.782	0.473
$\hat{p}$ =5%	1998	2001	2006	2011
$\hat{p}$ =10%	2000	2004	2008	—
$\hat{p}$ =15%	2002	2005	2010	—
$\hat{p}$ =20%	2003	2007	2011	—
$\hat{p}$ =50%	2009	2012	—	—
Panel B: Ceiling Parameter k = 0.8				
$\hat{p}$ =5%	1996	2000	2005	2010
$\hat{p}$ =10%	1999	2003	2007	—
$\hat{p}$ =15%	2001	2004	2009	—
$\hat{p}$ =20%	2002	2006	2010	—
$\hat{p}$ =50%	2007	2011	—	—
Panel C: Ceiling Parameter k = 0.6				
$\hat{p}$ =5%	1995	1998	2004	2008
$\hat{p}$ =10%	1997	2001	2006	2012
$\hat{p}$ =15%	1999	2003	2008	—
$\hat{p}$ =20%	2000	2004	2009	—
$\hat{p}$ =50%	2005	2009	—	—

Table 2.3: Adoption in Economics and Finance Journals

This table presents the estimates and fitted values of the empirical model of technology diffusion (S-curve) presented in Equation (2.1) applied to two distinct samples. The first sample ("Econ.") consists of all articles published in the top-three economics journals (AER, JPE, and QJE). The second sample ("ECF") consists of all empirical corporate finance articles published in the JF, JFE, and RFS over the period 1980-2012. In all estimations, the dependent variable (p_t) is the fraction of articles that adopt the *language* associated with the identification technology in year t. We define that an article adopts the language of identification if it mentions at least one of the keywords associated with the identification technology (see Table C.2 for details). Panel A reports the position (α) and slope (β) estimates of the logistic S-curve. Panel B reports the implied adoption thresholds, which are the years in which the fitted values ($\tilde{p}$) cross several adoption thresholds. We report tests of the difference between the slope coefficient and fitted threshold years across samples (Econ. vs ECF). The last row displays the adoption lag computed as the average difference in number of years for the crossing of the four adoption thresholds. We report estimates and fitted values for three different ceiling levels, $k = 1$, $k = 0.8$, and $k = 0.6$. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Ceiling (k):	$k = 1$			$k = 0.8$			$k = 0.6$		
	Econ.	ECF	Difference	Econ.	ECF	Difference	Econ.	ECF	Difference
Panel A: Estimation of S-curves									
β	0.068***	0.209***	-0.141***	0.080***	0.224***	-0.144***	0.115***	0.255***	-0.140***
	(18.63)	(5.59)		(17.99)	(5.99)		(11.62)	(5.55)	
α	-1.808***	-5.520***		-1.602***	-5.406***		-1.454***	-5.327***	
	(-26.55)	(-7.94)		(-19.35)	(-7.76)		(-8.35)	(-6.86)	
R^2	0.918	0.502		0.913	0.537		0.823	0.524	
Panel B: Implied Adoption Thresholds									
$\hat{p}$ =5%	1963	1992	-29***	1963	1991	-28***	1967	1989	-22***
$\hat{p}$ =10%	1974	1996	-22***	1973	1994	-21***	1974	1992	-18***
$\hat{p}$ =15%	1981	1998	-17***	1978	1996	-18***	1978	1994	-16***
$\hat{p}$ =20%	1986	2000	-14***	1983	1998	-15***	1981	1995	-14***
Average:			-20.5			- 20.5			-17.5

Table 2.4: Broader Definition of Empirical Corporate Finance

This table presents results for the same adoption model (S-curve) as in Table 2.2 estimated on a broader sample of empirical corporate finance (ECF) articles, which also includes ECF articles published in the top economics journals (AER, JPE, and QJE). Classification of the articles in these journals follows the same procedure used for finance articles, which is detailed in Section 2.3.2. For brevity, only the adoption threshold years are reported in the different panels. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	ID	IV	DD	Selection
Panel A: Ceiling Parameter k = 1				
$\hat{p}$ =5%	1990	1996	2004	2010
$\hat{p}$ =10%	1996	2001	2006	—
$\hat{p}$ =15%	1999	2004	2008	—
$\hat{p}$ =20%	2002	2007	2009	—
$\hat{p}$ =50%	—	—	—	—
Panel B: Ceiling Parameter k = 0.8				
$\hat{p}$ =5%	1988	1994	2003	2009
$\hat{p}$ =10%	1993	1999	2005	—
$\hat{p}$ =15%	1997	2003	2007	—
$\hat{p}$ =20%	1999	2005	2008	—
$\hat{p}$ =50%	2010	—	—	—
Panel C: Ceiling Parameter k = 0.6				
$\hat{p}$ =5%	1986	1992	2002	2008
$\hat{p}$ =10%	1991	1997	2004	2011
$\hat{p}$ =15%	1994	2000	2006	—
$\hat{p}$ =20%	1996	2002	2007	—
$\hat{p}$ =50%	2006	2012	2012	—

Table 2.5: Determinants of Adoption

This table presents the results of a time-to-adoption analysis using OLS. The sample comprises the author-years between 1980 and 2012 of the careers of authors who publish at least one ECF article in the JF, JFE, or RFS. See Section 2.5.2 for more details about the panel data construction. The dependent variable is an indicator variable that equals one once an author adopts the identification technology in any article in the ECF sample, and zero before. Observations for an author after adoption are dropped as is standard. Each ECF article is classified as using identification techniques based on a three-step procedure as detailed in Section 2.3.2. *Econ PhD* equals one if the author obtained a PhD in Economics. *TopPhD* equals one if the author holds a PhD from a top-25 institution, as defined by Kim, Morse, and Zingales [2009]. *Prev Papers* is the number of articles the author published in the JF, JFE, or RFS by the prior year. *Prev Cites* is the number of citations (via Web of Knowledge) received by previously published articles up to the prior year. The *Colleagues* network comprises authors that were previously employed at the same institution in the same year. The *Alumni* network comprises authors that previously received a PhD from the same institution. The *PhD Faculty* network comprises authors that were on staff at the PhD-granting institution in the two years preceding graduation with a PhD. *Adopting Colleagues, Adopting Alumni, and Adopting PhD Faculty* are the fraction of each network that adopted the identification technology by the prior year. *Career Age* is the number of years since an author's career began (see Section 2.5.2). *High Status* is a binary variable equal to one if the author has received a best-paper award or is a fellow of the American Finance Association. *Prev Econ* is a binary variable equal to one if the author has an article published in one of the top economics journals in previous years. *Prev Econ-ID* is a binary variable equal to one if the author has an article using the identification technology published in one of the top economics journals in previous years. To facilitate interpretation, *Log(1+Prev Papers), Log(1+Prev Cites)*, and the adopting network variables are standardized. Standard errors are clustered by author. We report t-statistics in parentheses. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Econ PhD	0.007***	0.005*	0.007***	0.007***	0.007***	0.005**
	(2.96)	(1.73)	(3.19)	(2.95)	(3.03)	(2.32)
Top PhD	-0.002	-0.001	-0.002	-0.002	-0.002	-0.002
	(-0.67)	(-0.20)	(-0.73)	(-0.64)	(-0.84)	(-0.74)
Top School (Current)	0.005**	0.003	0.005**	0.005**	0.004*	0.004*
	(2.12)	(1.06)	(2.03)	(2.27)	(1.86)	(1.80)
Log(1+Prev Papers)	0.001	0.001	0.005**	0.002	0.005**	0.002
	(0.43)	(0.37)	(2.20)	(0.76)	(2.05)	(0.70)
Log(1+Prev Cites)	-0.005*	-0.005	-0.009***	-0.007**	-0.010***	-0.006**
	(-1.91)	(-1.59)	(-3.38)	(-2.49)	(-3.65)	(-2.10)
Adopting Colleague		0.011***				
		(4.02)				
Adopting Alumni		0.011***				
		2.84)				
Adopting PhD Faculty		0.003				
		(1.36)				
Career Age			-0.001**	0.003***	-0.001**	
			(-2.13)	(4.54)	(-2.11)	
Career Age^2				-0.000***		
				(-8.22)		
High Status					0.025**	
					(2.23)	
Prev Econ						-0.001
						(-0.22)
Prev Econ-ID						0.023**
						(2.45)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Career Age FE	Yes	Yes	No	No	No	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,383	16,899	21,383	21,383	21,383	21,383
Authors	1,473	1,190	1,473	1,473	1,473	1,473
Adj. R^2	0.051	0.056	0.048	0.050	0.049	0.52
Avg. Adoption Rate	0.024	0.027	0.024	0.024	0.024	0.24

Table 2.6: Seniority

This table presents publication patterns in ECF and the subset of identification articles by career age groups. Panel A focuses on characteristics of solo-authored work, while Panel B focuses on the nature of collaborative work. The sample comprises the author-years between 1980 and 2012 of the careers of authors who publish at least one ECF article in the JF, JFE, or RFS. Each ECF article is included in the "identification" or "non-identification" category based on a three-step procedure as detailed in Section 2.3.2. Each row corresponds to a 10 year career age bin. We set year zero of a career to the minimum of the year of PhD, first publication, or first academic position. We omit observations 30 years after a career begins and any author whose career began before 1970. In columns (1) through (4) of Panel A, *All* and *ID* count the number of authorships among all ECF articles (respectively all ECF identification articles) for each age bin, where every author on an article receives full authorship credit. *Solo* and *Solo ID* count the number of authorships among all solo-authored articles, and respectively all identification solo-authored articles. Columns (5) through (7) show proportions of solo-authored identification authorships relative to all authorships, all authorships of identification papers, and all solo-authorships respectively. The bottom rows *Diff* and *P-value* compute the difference (and statistical significance) in the proportions of solo-authored identification articles between the first age group (Career Age 0-9) and the last age group (Career Age 20-29). Panel B reports the average age difference between an author and coauthors. We compute the average coauthor age for each author-year, and then average the age gap within the given age group. Higher values indicate more senior coauthors. For each group, we compute the average age gap for the subset of coauthors of identification articles (column (1)) and non-identification articles (column (2)). The difference in the average age gap across the two types of articles is computed along with the t-statistic from a test of equality in means in columns (3) and (4), respectively.

Panel A: Solo-Authored Identification Articles

	Authorships				Percent by Career Age		
Career Age	All	ID	Solo	Solo ID	Solo ID/All	Solo ID/ID	Solo ID/Solo
0-9	1,960	485	256	64	3.3%	13.2%	25.0%
10-19	1,208	281	59	15	1.2%	5.3%	25.4%
20-29	460	105	7	1	0.2%	1.0%	14.3%
Diff					-3.0%	-12.2%	-10.7%
P-value					<0.01	<0.01	0.52

Panel B: Avg. Age Difference of Coauthors by Article Type

Career Age	ID	Non-ID	Difference	T-Stat
0-9	7.3	6.0	1.3	3.04
10-19	-2.3	-2.1	-0.2	-0.37
20-29	-12.2	-10.7	-1.5	-1.70

Table 2.7: Straddlers

This table explores the characteristics of authors straddling between economics and finance journals. We define authors as "straddlers" if they have at least one publication in a top-three finance journal and one publication in a top-three economics journal at any point in their career. The sample contains all authors who publish at least one ECF article in the JF, JFE, or RFS during 1980 - 2012. Panel A provides a comparison between straddlers and non-straddlers in terms of adoption rate. *Adopting Author* is a binary variable equal to one if an author adopts the identification technology in at least one published paper over the sample period. Panel B offers descriptive statistics on the academic background (training) of straddlers vs. non-straddlers in terms of field of PhD studies and PhD institution. *Econ PhD* equals one if the author obtained a PhD in Economics. *TopPhD* equals one if the author holds a PhD from a top-25 institution, as defined by Kim, Morse, and Zingales [2009]. In Panel C, we compute in the first row the average first year of adoption for all authors (column (1)), non-straddlers (column (2)) and straddlers (column (3)). The second and third row provide the average first year of adoption only within the subset of the top-three finance journals and top-three economics journals respectively. In column (4), we compute a Z-statistic for the test of equality in proportions across straddlers and non-straddlers in Panel A and B, and a t-statistic from a test of equality in means in Panel C.

Author sample:	All	Non-straddlers	Straddlers	Z-stat/T-stat
		Panel A: Authors		
All	1,780	1,489	291	
Adopting Authors	648	484	164	
Adopting Authors	36.4%	32.5%	56.4%	-7.86
		Panel B: Author Training		
Econ PhD	30.9%	23.7%	65.8%	-14.41
Top PhD	57.7%	53.4%	79.8%	-8.37
		Panel C: Year of Author's First ID Article by Field		
Overall	2006	2007	2003	9.48
Fin	2007	2007	2005	4.11
Econ	2002	N/A	2002	

Chapter 3: Were non-independent boards really captured before SOX?

3.1 Introduction

Despite more than a decade of research on corporate governance following the passage of Sarbanes-Oxley (SOX), consensus still evades researchers about the inference that should be taken from the literature on corporate governance that followed.[1] Some papers suggest that the stipulations of the regulations regarding corporate governance, in particular the mandate that the board of directors be independent, induced beneficial changes in firm values and performance.[2] Other papers support the view that governance structures endogenously arise as solutions to the contracting environment of the firm.[3]

This paper takes a novel approach to investigating how board independence impacts performance. Should we infer that governance structures before SOX led to detrimental agency costs and that shareholders are better off with the regulations? Or should we infer that boards are chosen as endogenous and optimal responses to

[1] 132 papers were published containing either "SOX" or "Sarbanes-Oxley" from 2002 to 2013 in the Journal of Finance, the Journal of Financial Economics, and the Review of Financial Studies.

[2] See, e.g., Bebchuk and Grinstein [2005] and Chhaochharia and Grinstein [2007].

[3] See, e.g., Ahern and Dittmar [2012] and Demsetz and Lehn [1985].

the heterogeneous contracting environment in which a firm exists? This debate has not been settled; central papers supporting each viewpoint are contested.

By exploiting the specific legal stipulations of exchange mandates regarding board structure, this paper provides new, casual, and U.S. based evidence that the constraints on board structure imposed by the law were harmful to firms. The key new insight is the observation that the legal definition of an independent director allows firms to "reclassify" a board member as an independent director—even if the social and economic relationship between the CEO and director are largely unchanged. Thus, the law did not constrain some firms as it might appear—in effect, there was a subgroup of "treatment" firms that received a "placebo".

Based on this observation and manually collected data on directors, I specify a difference-in-difference-in-difference (DDD) empirical design that estimates the differential response of the treatment and placebo groups. The intuition of the test is as follows: Imagine two identical firms, with three independent directors (henceforth called "I") and five non-independent directors (henceforth called "N") who receive a mandate that independent directors must form half of the board. Suppose the eight directors is the optimal board size for these firms. The first firm alters the personnel on its board by hiring an I and removing an N. The second firm takes action to reclassify an N director as an I director. Legally, the second firm is compliant, but its board is exactly the same as before! The relative performance of the two firms after this change likely depends on why the firm was at its original board position. If the $3-I/5-N$ board was optimal, the reclassifying firm should perform better, all else equal, because it avoided any changes while the first firm moved away from the optimal board structure. Conversely, if the board structure imposed agency costs and was not optimal to begin, then the reclassifying firm should perform relatively worse. The crucial aspect of the test design is that the differential performance effect indicates which theory about the ex-ante board structure fits the data better.

I examine which theory dominates in the data using an empirical test designed to match the intuition above. Specifically, I estimate a DDD model which reports the differential response between the two types of treatment firms—those that reclassify directors and firms that alter personnel. I define treatment firms as those with non-independent boards ($N > I$) in 2001. I define "reclassifiers" as firms for whom net reclassifications (total reclassifications from N to I net of total reclassifications from I to N) in the years between the passage of the law and the year in which the requirement was binding makes the board more legally independent without recomposing the board's personnel. The DDD framework controls for time effects and the endogeneity of the choice of board composition and thus produces a causal treatment effect.

I find statistically significant evidence that reclassifiers outperformed similar treatment firms that changed board personnel to meet the independent board mandate. The central estimate in the paper is that the profitability of placebo firms was 4.9 percentage points higher than comparable treatment firms. The magnitude of the difference—ranging between 3-5%—is economically meaningful (the sample average is 14%), implying that the constraints are significant impediments to the conduct of firms targeted by the regulations.

It is important to emphasize that one reclassification can substantially reduce the need for the firm to recompose its board. The alternative option to achieve the equivalent shift in legal independence is to fire (or not renew) one non-independent director and to hire a new independent director. Typically, these non-independent directors have significant personal knowledge about the firm. Outside and independent replacements who are new to the board have adjustment costs as they learn about the firm. The difference in value to the firm of these skill sets could be large (see, e.g. Duchin, Matsusaka, and Ozbas [2010]). The magnitude of this adjustment cost and the value of firm specific capital for board members, in part, is what this paper tests. The central tests indicate that this value is large.

To verify the validity of the main findings, I run a battery of tests. First, I focus on robustness of the main specification. The result is robust to estimation in first differences and firm fixed effects. The baseline model does not include additional time-varying controls, such as R&D expenditures, as these endogenous corporate choices might be altered as a result of the mandate.[4] Nevertheless, all results are robust to including controls that are standard in corporate finance studies.

Second, I conduct a series of tests suggested by Roberts and Whited [2013]. The treatment and placebo groups satisfy the parallel trends assumption on the main outcome of interest and are generally very well balanced on key covariates before the rule is introduced. Additionally, falsification tests add confidence to the

[4] For example, boards composed of outside directors might perceive lower benefits from internal firm R&D and thus advocate for decreased investment in research, which would alter profits in the short term.

empirical design: Tests that assume the mandate's treatment begins in 2001 or 2005 (rather than 2003) find no differential treatment effect. Moreover, the DDD coefficient is severely weakened when the sample is expanded to include firms that receive a lower "dosage". The final falsification test shows that among control firms where the mandate's constraint is not binding, reclassification is not associated with a difference in performance.

The last set of tests explore possible mechanisms for the change in performance and the valuation implications of the performance differential. First, I decompose the DDD estimate regarding profitability into its components. The principle change among placebo firms is an increase in sales, rather than a decrease in expenses or asset size. Second, I investigate two sets of channels that could justify the relative increase in profit margin: investment strategy and changes in the competitive landscape. These tests do not find conclusive evidence supporting either channel.[5] Third, I check if the increased profitability of reclassifying firms flows through to increased valuations or dividends. While I find positive DDD point estimates for market value, Tobin's Q, and dividends, none are statistically significant.

The findings contribute to the debate about the efficacy of the board independence mandate and provide more evidence that it harmed some firms: one size does not fit all when it comes to corporate governance. This conclusion is in line with

[5] While overall SG&A expenses of placebo firms increase more than treatment firms, the *level* of investment (CAPX, R&D, and acquisitions) are not elevated. However, there is some evidence that risk profiles change somewhat, which suggests that the underlying firm assets might be shifting due to changing investment allocations (at least relative to treatment firms).

Cohen, Frazzini, and Malloy [2012], Coles, Daniel, and Naveen [2008], Duchin, Matsusaka, and Ozbas [2010], and Falaye, Hoitash, and Hoitash [2011]. These papers find that board independence is not costless. Non-independent directors typically serve an advising role using their firm specific knowledge. Furthermore, independent directors may be chosen for their favorable views of management, and appointing independent directors can increase horizon-based agency conflicts on the board.

Finally, the paper contributes a new setting to reduce the endogeneity problems common to board structure regressions. To my knowledge, I am the first to utilize a feature of the rules that allows for some firms to escape the bite of the rules; that is, some firms were able to use a loophole. The presence of the "loophole" in the definition of independent directors in effect creates a treatment and placebo group that are semi-exogenously fixed at the passage of the law. This test design could be useful in other settings.

The remainder of the paper is organized as follows. Section 3.2 discusses the rules passed by the exchanges and the definition of independent directors and constructs the key test variables. Section 3.3 conducts the main analysis. Section 3.4 explores other differences between the placebo and treatment groups. Section 3.5 concludes.

3.2 The difference between legally independent and economically independent directors

The section discusses the data underlying the main tests. I first explain the independent board rules passed by the exchangse to show how director reclassifications take place. After describing reclassification in the data, I define the treatment group and the placebo group. Finally, I introduce the additional data needed for the main tests.

3.2.1 Exchange rule background

Political economy issues notwithstanding, writing laws and statutes is difficult because implementation and enforcement inevitably hinge on the definition of terms. Even assuming agreement in principle by the parties involved in writing a statute, consensus on the manner in which a law is operationalized can be difficult to reach due to the unclear mapping from economic and legal concepts to real world outcomes.

Following the corporate scandals that preceded Sarbanes-Oxley (SOX), the NYSE and NASDAQ decided to institute a rule requiring the board of directors of listed firms be comprised of a majority of "independent" directors. The idea behind the new rule was to encourage the creation of boards which were more effective in monitoring and correcting adverse CEO's decisions. The operational legislative step is defining what conditions qualifies a director as an independent director and to do so in a way that ensures such directors are better monitors, at least on average.

Chhaochharia and Grinstein [2007] presents a thorough and concise summary of the events that precipitated SOX and the follow-on rules by the NYSE and NASDAQ exchanges that mandated boards be comprised of a majority of independent directors. The key dates to know about, for the purposes of this paper, are that the rules were introduced and passed in the fall of 2002 and firms had until 2006 to comply with the mandate. Thus, I set 2003 as the year in which the treatment period begins.

The key definition, per the NYSE rule, is that a director is considered independent if (1) the board affirms no material relationship exists (partial list of such relationship: commercial, industrial, banking, consulting, legal, accounting, charitable, and familial), and (2) in the past three years, neither she nor any immediate family member has been an employee of the firm or its auditor(s), received more than $100,000 outside of compensation for the directorship or prior services, been in an interlocked board with an executive, or had outside employment with a firm that did business with the firm of more than one million dollars or 2% of the other companys gross revenues.

It is implicitly assumed by researchers because of the intuition that a director is unlikely to switch from being a lax monitor to being stringent monitor that making a board more (legally) independent requires a combination of hiring independent directors or culling inside directors. Yet, this is not the case in practice. The counterintuitive implication of the rule is that directors can switch classifications from non-independent (I) to independent (N). A lifelong employee on the board that considers the long time CEO a great friend is considered independent in the fourth year after the termination of her primary role at the firm, despite the fact that the social connection and associated conflicts remain (Duchin, Matsusaka, and Ozbas [2010]). If a "material relationship" such as a consulting contract ends, the director classification goes from N to I, despite the fact that the director might anticipate being hired as a consultant again in the future. Even though we might be hesitant to believe that the monitoring incentives and economic relationship vary little around these changes, these legal reclassifications did occur because of the legal implementation of the rule.

Director reclassifications are indeed present in firm filings given to the SEC. For example, in 2001, the chairman of the board of Ralcorp, William P. Stiritz, was paid $1.2 million in advisory fees. Since that arrangement was ended during the year 2001, the 2004 DEF-14A listed him among the independent directors. In 2005, Emmis Communications Corp. added Greg Nathanson to its roster of independent directors "because on that date three years will have elapsed since he was employed by Emmis." The next year, Emmis reported another reclassification, noting that "We expect Mr. Leventhal to qualify as an independent director on June 25, 2005, because on that date three years will have elapsed since Mr. Leventhal's brother-in-law ceased to be one of Emmis' executive officers."

3.2.2 Director data

It is crucially important to obtain director classifications that match the *exchange's* definitions, rather than corporate governance advisory firms. I obtain data on board members from 1998 to 2006 from IRRC. IRRC provides a variable that classifies directors as independent "I", linked-affiliated "L", or employee "E". However, these definitions are based on IRRC's evaluation of independence, not the exchanges. They additionally provide other information about links between the director and the exchanges that can be used to correct this discrepancy.

To define directors in a way that matches the exchanges as closely as possible, I use the classification routine from Guthrie, Sokolowsky, and Wan [2012b], which is as follows: If IRRC defines a director as independent, I do. If IRRC defines a director as linked, I classify them as independent provided that the director is not a relative or interlocked with the CEO and the director is one of the following: not an employee more in the last three years; never an employee; not linked via a business transaction, a professional service, a charity, a director designated (e.g. by a union), or the IRRC variable "otherlink." Because of the "materiality" exception the exchanges provided firms in classifying directors, this approximates the rules of the exchanges as closely as possible using the codified data in the IRRC dataset. The remaining directors are classified as N.

Because of the small number of treatment firms, it is important that the classifications of directors at such firms are accurate. As such, I manually reexamine all director observations for all treatment firms between 1999 and 2006 by reading their DEF-14A filings. For each firm year, I verify that (A) the firm did not comply with the rule as of the treatment date and (B) directors are classified properly in all years. I defer to the firm documentation on which directors are independent when listed explicitly. Otherwise, I apply the law as defined which requires referencing

DEF-14As from the preceding years.

3.2.3 Board composition dynamics

Table 1 reports time trends in the board structure of the average firm.

Panel A reports that the average board size is roughly 9 throughout, composed of 6.6 I and 2.4 N in 1999 and steadily reshuffling to a more independent board of 7.3 I and 1.8 N in 2006.[6] There is a substantive shift to independent directors in the latter years. In particular, the year with the highest churn in board composition based on new hires and departures of both types of directors is 2005, which is the year in which most firms were required to meet all stipulations of the rules.[7] This aligns with summary stats in other papers on the topic.[8]

[Insert Table 3.1 about here]

Panels B and C of Table 1 report turnover that makes a firm more and less legally independent, respectively. Most shifts in independence come from personnel decisions. For the average board in the years before 2003, 0.6 independent directors and 0.2 non-independent directors are hired. As pre-SOX boards are relatively stable, roughly the same amount (of each type of director) depart in a given year. In each year from 2003 to 2006, which is the compliance period for the rule, the average board hired about 0.1 more independent director than departed and retained about

[6] In each year, the median board is 7 I and 2 N.

[7] Firms with classified boards were given until 2006 to comply.

[8] E.g. Chhaochharia and Grinstein [2007], Duchin, Matsusaka, and Ozbas [2010], and Guthrie, Sokolowsky, and Wan [2012a].

0.1 fewer non-independent directors than they hired.

Panels B and C of Table 1 also report that director reclassifications are surprisingly frequent. In the pre-treatment period, the average board reclassified about 0.11 directors per year as independent and 0.06 directors lost their independent designation. In the post-treatment period, boards continued to reclassify directors as independent, even as the pool of non-independent directors shrank. During this same period, the average number of directors losing their independent designation remained flat, at 0.06 directors per year.

Overall, the pool of directors remaining on the board shifted significantly during the sample: Later in the sample, the average board retained fewer non-independent directors but retained more independent directors. This trend makes inferring the likelihood of a director reclassification from the information in Panels B and C difficult. Thus, Panel D directly reports the likelihood that a director classification changes, conditional on remaining on the board. Throughout the sample, reclassification propensity is flat, with 5% of returning non-independent directors reclassifying to independent status. Meanwhile, directors rarely reclassify *away* from independent status. In fact, such reclassifications take place in just 1% of possible occurrences.

Finally, Panel E highlights that reclassification is relatively common at the *firm* level. In an average year, roughly 10% of firms reclassify a director to independent status. This fraction declines during the sample, likely because the available pool of N directors eligible to reclassify falls.

3.2.4 Defining treatment and placebo groups

Based on the timeline provided in Chhaochharia and Grinstein [2007], I set 2001 as the year to determine whether a firm is "treated" by the rules because the proposals were well known in 2002. Firms with $N > I$ in 2001 must alter their boards to comply and are the *Treatment* group. Otherwise, firms are in the control group. Based on the timing

Next, I must define if a firm is a *Reclassifier*, which captures firms that use reclassification in a way that makes their boards more independent without altering personnel. First, I define

$$
\begin{aligned}
\textit{Net Reclassification}_i = & \Sigma_{t=2003}^{2006} \#\text{ of Reclassifications to } I_{i,t} \\
& - \Sigma_{t=2003}^{2006} \#\text{ of Reclassification to } N_{i,t}
\end{aligned} \tag{3.1}
$$

wherein I measure reclassifications during the compliance period of the rules (2003 to 2006). Then, I simply define *Reclassifier* as one if *Net Reclassification*> 0 and zero otherwise. Note that control firms can be defined as a *Reclassifier*.

With these variables defined, it is important to clarify how the variables map into the language I have used to describe the tests. What I have called a "placebo" firm is one for which *Reclassifier=1* and *Treatment=1*. While they are part of the sample for whom the variable *Treatment* is one, they are not a "treatment" firm. A "treatment" firm is a firm for which *Reclassifier=0* and *Treatment=1*.

3.2.5 Sample and other data

To examine the performance of firms, I merge the IRRC board data with data from the CRSP-Compustat Merged (CCM) database. I drop observations with SIC codes 4900-4949 and 6000-6999, total assets less than five million, and observations where DLC plus DLTT exceed AT. The firm must be listed on the NYSE or NASDAQ exchange. All ratios based on CCM and CRSP data are winsorized at the 1% tails.

[Insert Table 3.2 about here]

Table 3.2 defines and reports summary stats for all variables used in the paper, based on the sample in the main regression (Column 1 of Table 3.4). Average profitability is 14% in the sample. 10% of firm-years in the sample are *Treatment* firms, and firms are defined as *Reclassifiers* in 20% of observations. Given that the IRRC data concentrate on large firms, the summary statistics of the sample are standard for finance studies focused on large firms.

3.3 Main analysis

This section conducts the main tests in the paper. First, I design a difference-in-difference-in-difference (DDD) empirical strategy to produce causal estimates of the effect of board structure on firm performance, paying special attention to ex-ante balance between groups. Second, I execute the main test and subject that to robustness tests for specification form and parallel trends. Third, I conduct a series of falsification tests to validate the test design and key findings.

3.3.1 Methodology

The ideal test of the effect of board independence on performance would randomly treat a subset of firms by replacing one non-independent (N) board member with an independent (I) board member. The remaining firms would then be the control group, and a simple comparison of average performance after the treatment would produce an estimate of the effect of board independence.

In lieu of this ideal experiment, financial economists have examined the exchange rules, which deterministically forced non-majority independent boards to recompose membership. The difficulty with this strategy is obvious: These firms do not randomly have such boards. Thus, the standard approach is to use care in designing the comparisons, principally through the use of control variables. However, these controls are *also* not randomly distributed among treatment and control firms. Firms that construct insider dominated boards likely face significantly different business fundamentals than a firm which chooses a completely independent board. For example, Duchin, Matsusaka, and Ozbas [2010] conclude that outside directors are more valuable when the cost of acquiring information about the firm and its business environment is lower.

The problem with a simple diff-in-diff approach to exploiting the rule change motivates looking for a new empirical strategy. One feasible method is to examine variation in treatment dose, wherein the effects of having to change a board with 0%

independence are assumed to be greater than changing a board with 40% independence. This alternative strategy runs into the same problem as a straightforward diff-in-diff: These firms endogenously arrived at these board structures. A second feasible method exploits the variation in treatment dose rather than examine it. In particular, some papers use pre-SOX independence as an instrumental variable for board changes. The same endogeneity concerns apply to this strategy however.

The ability for a director to be reclassified presents another option. Suppose in 2001 a firm has three I and five N. For one firm, an employee director retires. After the 2002 rules, this director will be considered legally independent as of 2005. This scenario has two crucial aspects. First, the change in employment status occurs before the rule, and is thus plausibly exogenous to factors related to firm performance. Second, this board is now more *legally* independent but the individual relationship with the CEO that largely governs the director's monitoring incentives is unchanged (see, e.g. Coles, Daniel, and Naveen [2014]). This firm, in fact, complies with the ruling without any additional board changes (with four N and four I). Thus, this firm receives a "placebo" treatment.

Comparing the placebo firm to a similar treatment firm that actually recomposes its board will generate an estimate of the true performance effect of board independence. Therefore, I specify a DDD test that controls for time trends, industry factors, and the fact that firms whose board has directors that can be reclassified are different from firms whose boards lack that flexibility. The central specification is a standard DDD:

$$\begin{aligned}Profitability_{i,j,t} =& \beta_1 Reclassifier_i \times Treatment_i \times Post_t \\ &+ \beta_2 Reclassifier_i \times Treatment_i + \beta_3 Reclassifier_i \times Post_t \\ &+ \beta_4 Treatment_i \times Post_t + \beta_5 Reclassifier_i + \beta_6 Treatment_i \\ &+ \beta_7 Profitability_{i,t-1} + \phi_j + \phi_t + u_{i,j,t} \end{aligned} \tag{3.2}$$

Reclassifier and *Treatment* are defined in Section 3.2.4. *Post* is equal to one after 2002 and zero otherwise. As noted by Ahern and Dittmar [2012], including additional controls is likely to confound the estimates, so I simply model profitability as a function of its lag. In additional tests, I report first differenced and fixed effect estimators. Fixed effects ϕ_j and ϕ_t are included for industry and year, respectively.[9] Standard errors are clustered at the firm level.

In this specification, β_1 produces an estimate of the *differential* treatment effect between treatment firms that reclassify (which I have called placebo firms) and treatment firms that do not reclassify. In principle, this approach is equivalent to separately running one diff-in-diff on reclassifying firms and another diff-in-diff on non-reclassifying firms and then comparing the two diff-in-diff coefficients. This is precisely what the intuition of the test calls for. However, this requires additional tests for ex-ante balance; in addition to balance between treatment and control firms, we need balance between firms that reclassify and those that do not.

Table 3.3 presents tests for ex-ante balance for the level and growth rate of key corporate variables, the most important of which is the outcome variable *Profitabil-*

[9] The inclusion of year fixed effects is why *Post* is not included as a variable.

ity. I split the sample into two parts. Panel A compares firms near the independence threshold in 2001, that is firms within two members of the 50% cutoff. Specifically, firms with $|I - N| \leq 2$. On all but one dimension and measure, this subsample exhibits good balance.

[Insert Table 3.3 about here]

Panel B compares firms far from the cutoff and reveals that ex-ante balance in terms of profitability is compromised across the reclassification groups. This holds whether we examine the level of profits (where reclassifying firms have 3.8% lower profits) or the growth rate (where profits of reclassifying firms grow 13.1% less). Because of this balance issue, I impose a sample restriction across all main tests and focus on firms near the cutoff.

This sample restriction is prompted by econometric reasoning but also solves another problem: Even if firms far from the cutoff reclassify a director, they will not escape constraints of the mandate. Therefore, these reclassifying firms will be forced to alter their board. In effect, the differential between these reclassifying firms and similar treatment firms is reduced. As such, including firms far from the cutoff in the test would attenuate the main coefficient. This reasoning will justify an additional robustness test.

3.3.2 Central results

Table 3.4 presents the main results from estimating Equation 3.2. In column 1, the key coefficient indicates that, controlling for lagged profitability, placebo firms (treatment firms that reclassified) outperformed comparable treatment firms that do not reclassify by 4.9 percentage points. This difference is economically meaningful as well, equivalent to about a third of the sample mean profitability rate. This estimate implies that the board constraints are significant impediment to the conduct of firms targeted by the regulations.

Row two (*Reclassifier*$\times$ *Treatment*) indicates that the pre-treatment balance maintains in a regression controlling for industry effects. Rows three to six indicate that there are no discernible performance difference across the treatment and reclassification groups in the pre- and post- periods.

The remaining columns report several robustness tests. Column 2 uses free cash flow as the measure of performance. Column 3 includes additional controls despite possible econometric issues highlighted by Ahern and Dittmar [2012]. Column 4 redefines *Reclassifier* based on director reclassifications during the 2000-2003 period. This measure is plausibly more exogenous, as the reclassification pre-date the rule changes, but balance is compromised (as indicated by coefficients in rows two to six). Column 5 presents the most stringent specification, estimated with firm fixed effects. Despite reduced power from the relatively small sample of treatment and placebo firms, the coefficient is still statistically significant.

[Insert Table 3.4 about here]

Figure 3.1 reports the differential performance between placebo and treatment firms for each year of the sample. I rewrite Equation 3.2 with each term multiplied by a vector of year dummies ϕ_t as[10]

$$\begin{aligned} Profitability_{i,j,t} = & \beta_{1,t} \times \phi_t \times Reclassifier_i \times Treatment_i \\ & + \beta_{2,t} \times \phi_t \times Reclassifier_i \\ & + \beta_{3,t} \times \phi_t \times Treatment_i \\ & + \beta_4 Profitability_{i,t-1} + \phi_j + \phi_t + u_{i,j,t} \end{aligned} \tag{3.3}$$

Thus, $\beta_{1,t}$ is now a vector of annual treatment coefficients. The figure makes two key points. First, each the pre-treatment year is not statistically different from zero, satisfying the parallel trends assumption. Second, all years after treatment begins have positive point estimates, and the 2004 estimate is statistically significant at the 10% level.

[Insert Figure 3.1 about here]

3.3.3 Robustness and falsification tests

To validate the main findings, I run several falsification tests as suggested by Roberts and Whited [2013]. The results are reported in Table 3.5. I first assume incorrect dates for the onset of treatment. The baseline specification takes treatment as beginning in 2003, the first year after the passage of the rule. Instead of that, column 1 falsely assumes treatment begins in 2001 and column 2 falsely assumes treatment begins in 2005. Both columns report no differential performance between placebo firms (treatment firms that reclassify) and treatment firms that do not reclassify.

[10] Some terms in Equation 3.2 drop out.

[Insert Table 3.5 about here]

Next, I alter the "dosage" received by placebo firms. Recall that the main test includes a key restriction: firms must be near the 50% independence cutoff in 2001. This restriction is primarily imposed to improve balance. However, treatment firms far from the cutoff that reclassify present an additional challenge for the main test: Even if these firms reclassify a director, they will not escape constraints of the mandate. As such, firms far from the cutoff will be "treated" to some degree even if they reclassify. Avoiding contaminated treatment/placebo definitions thus offered further motivation to remove these firms from the main test. Yet, this issue presents the opportunity for a new robustness test: Including these firms should attenuate the main result severely. Thus, column 3 relaxes the key restriction that firms be near the 50% independence cutoff in 2001 and includes distant treatment and placebo (but not control) firms. As expected, the coefficient is deflated, to 0.02 and is not statistically significance.

Finally, column 4 examines whether reclassifying firms perform better *within* the set of control firms. They should not, as non-reclassifying control firms, unconstrained by the mandate, have not made board changes against their will. Furthermore, reclassification among control firms is more likely to reflect natural shifts due preexisting board arrangements (such as former employees shifting to independent status) than targeted efforts to avoid the mandate. Thus, no post-2003 differential should be expected. Indeed, row three, $\mathit{Reclassifier} \times \mathit{Post}$ finds no evidence that control firms perform better when they reclassify directors.

3.4 Performance and other changes

The section presents an additional set of tests to explore possible mechanisms for the change in performance and the valuation implications of the performance differential. The tests are presented in Table 3.6, in which each regression replaces the outcome variable (and its lagged counterpart) with a new variable. For brevity, I only report the main DDD coefficient.

I begin by investigating the source of increased accounting profits. After showing that it stems from an increase in sales, I investigate competitive, investment, and risk taking changes that might explain the sales increase. I conclude by asking if the profitability results in capital gains or dividend increases.

3.4.1 Decomposing the profitability response

Profitability is defined as $OIBDP/AT$, which can be further written as $(SALE - COGS - XSGA)/AT$. Panel A of Table 3.6 examines each subcomponent of this variable.

[Insert Table 3.6 about here]

Column 1 reports that relative asset values (measured by *Log(AT)*) between the placebo and treatment firms do not change after the passage of the law. Columns 2 to 4 examine $SALE/AT$, $COGS/AT$, and $XSGA/AT$, respectively. The principle effect driving the profitability increase is a statistically significant increase in sales of 10.1 percentage points. This outpaces a negligible increase in costs of goods sold and a statistically significant increase in selling, general, and administrative expenses.

3.4.2 Investment types and risk taking

The question that follows is obvious: What activities might change for firms when their board is forced to become more independent such that costs decrease and sales decrease even more (relative to a placebo firm)? The literature on board governance focuses on two key differences between inside and outside ("independent") board members: information acquisition and agency concerns. Independent directors typically have less information about the firm but are seen as more stringent advisors (e.g. Duchin, Matsusaka, and Ozbas [2010]). Independent board members may place different values on the level and allocation of investments, as they have different information about the competitive landscape and a different relationship with projects within the firm. In particular, independent board members might place a lower value on an ongoing project than insiders that have personal and psychic capital attached to the project.

This logic suggests that treatment firms might (relatively) decrease investment. Given that placebo groups have higher SG&A expenses (of which investment is a component), this seems plausible. Furthermore, based on the same logic, treatment firms might shift investment to less risky projects where monitoring progress is easier and rents are harder for insiders to capture.

Panel B tests these predictions (which imply a *positive* DDD coefficient for investment models), but finds only weak support given the (lack of) precision of the estimates. Concretely, estimates indicate that treatment firms decrease expenditures on CAPX, acquisitions, and R&D and hold less cash following the rule change, but none of these estimates are statistically significant. However, there is some evidence that risk profiles change somewhat, which suggests that the underlying firm assets might be shifting due to changing investment allocations (at least relative to treatment firms).

3.4.3 Competition

Given that revenues and profits are higher for placebo firms after the rule changes and that investment choices do not seem to explain the increase, the obvious remaining explanation is changes in the competitive landscape. If placebo firms persist in industries where competition falls, then profits would rationally rise. This, of course, would drastically change the interpretation of the main findings. Panel C shows that this concern is unfounded: Changes in competition for placebo firms and treatment firms are similar.

3.4.4 Valuation and payout policy

Finally, I ask if the increased profitability of reclassifying firms flows through to increased equity valuations and/or dividends. Panel D reports the results. Again, while the point estimate indicates that placebo firms increased in value and paid more dividends than treatment firms as a result of the rule changes, the estimates are imprecise. As such, no definitive conclusions can be drawn about how valuations change.

3.5 Conclusion

Corporate governance reform has long focused on increased monitoring to prevent excess private rent accruing to executives, reduce corporate malfeasance, and avoid scandals that can bring down firms. These efforts focus on tail risks but impose costs due to regulatory compliance. However, well identified evidence regarding the average relationship between board structure and performance is limited. Notable exceptions include Chhaochharia and Grinstein [2007] and Duchin, Matsusaka, and Ozbas [2010] which find that independence increases firm performance and value in some cases. Yet, Duchin, Matsusaka, and Ozbas [2010] also finds that increasing board independence can harm performance when information is costly for outsiders to obtain. Additionally, Ahern and Dittmar [2012] provides evidence from Norway that a different kind of mandated board changes reduce firm value.

This paper takes a completely novel approach to examining this key relationship. By exploiting the difference between legal and economic changes in board structure, I estimate the impact of board independence on performance and find evidence that the independent board mandate reduced firm performance. In other words, real outcomes are better for firms that were able to avoid the constraint imposed by the exchanges in the aftermath of SOX. This result is consistent with the view that boards are optimal contracting outcomes.

Figure 3.1: Yearly DDD coefficient estimation

This figure decomposes the main DDD coefficient reported in Column 1 of Table 3.4 to show, for each sample year, the relationship between *Profitability* (the independent variable) and the interacted term *Reclassifier*Treatment*. Specifically, all terms in Equation 3.2 used for Table 3.4 are multiplied by a vector of year dummies, resulting in Equation 3.3. The coefficient $\beta_{1,t}$ on each $\phi_t \times Reclassifier_i \times Treatment_i$ term is reported below by the solid line, along with the 90% confidence interval in the shaded region. Thus, this figure contains a visual test for parallel trends in the pre-treatment period (before 2002) and illustrates the timing of the positive loading on the DDD coefficient. See Table 3.2 for more details on the sample and variable definitions. Industry and year fixed effects are included. Standard errors are clustered by firm.

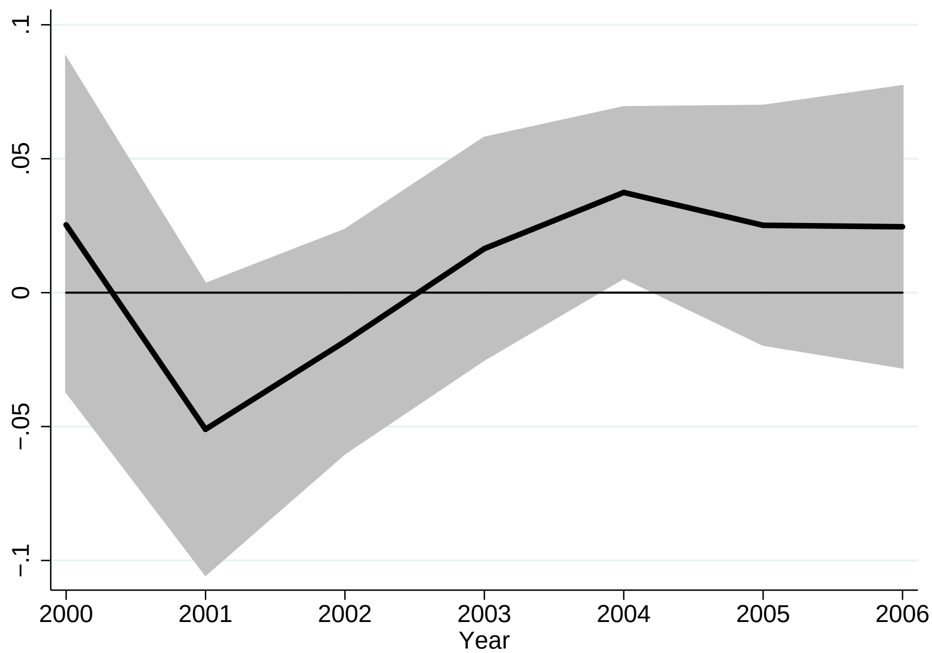


Table 3.1: Board structure dynamics

This table details how the average board restructures, in terms of legal independence, throughout the sample period of 1999 to 2006. Board data are from IRRC and the sample is a result of a merge to the CRSP-Compustat Merged database and Execucomp. N refers to directors legally classified as non-independent. I refers to directors legally classified as independent. Directors are classified as I or N in a way that matches the rules of the exchanges as closely as possible, given exceptions the exchanges provided firms in classifying directors. Section 3.2.2 provides more details on the classification of directors. Panel A reports the average board structure in terms of size and independence. Panels B and C report how the structure of the average board is modified via the hiring, reclassification, and departure of directors. Panel D reports the likelihood of a director being legally reclassified, conditional on being retained on the board from the prior year. Panel E reports the fraction of firms, in a given year, for which at least one director is reclassified.

Year:	1999	2000	2001	2002	2003	2004	2005	2006
Firms	1,182	1,194	1,248	1,073	1,077	1,077	1,062	991
Panel A: Board composition								
Size	9.05	8.88	8.71	8.70	8.79	8.81	8.85	9.07
Independent (I)	6.62	6.54	6.52	6.60	6.84	6.92	7.03	7.26
Not independent (N)	2.43	2.33	2.19	2.09	1.95	1.88	1.81	1.80
Panel B: Turnover increasing board independence								
Hire I	0.61	0.57	0.61	0.64	0.73	0.72	1.00	0.67
Reclassify N to I	0.13	0.12	0.11	0.09	0.10	0.08	0.07	0.06
N leaves	0.22	0.29	0.26	0.25	0.21	0.18	0.26	0.18
Panel C: Turnover decreasing board independence								
Hire N	0.17	0.19	0.18	0.16	0.10	0.11	0.19	0.14
Reclassify I to N	0.07	0.06	0.06	0.05	0.04	0.06	0.07	0.06
I leaves	0.60	0.69	0.63	0.65	0.61	0.65	0.92	0.51
Panel D: Fraction of returning directors that reclassify								
Director reclassifies N to I		0.06	0.05	0.05	0.05	0.05	0.04	0.03
Director reclassifies I to N		0.01	0.01	0.01	0.01	0.01	0.01	0.01
Panel E: Fraction of firms with reclassifying directors								
Firm reclassifies N to I	0.12	0.11	0.10	0.08	0.09	0.07	0.06	0.05
Firm reclassifies I to N	0.06	0.06	0.05	0.05	0.04	0.05	0.06	0.06
Firm reclassifies any	0.17	0.15	0.15	0.12	0.12	0.12	0.12	0.11

Table 3.2: Summary statistics for main regression sample

This table reports summary statistics for the main regression sample in column 1 of Table 3.4. Data are yearly observations (1999-2006) from IRRC and result from a merge to the CRSP-Compustat Merged (CCM) database and Execucomp. I drop observations with SIC codes 4900-4949 and 6000-6999, total assets less than five million, and observations where DLC plus DLTT exceed AT. The firm must be listed on the NYSE or NASDAQ exchange. To improve ex-ante balance and reduce the bias induced by heterogeneity, the sample is restricted to firms that were close to the 50% independent threshold in the year before the first rule proposals were made. Specifically, all firms in the main regression are within two members of the 50% independent threshold (i.e. $|I - N| \leq 2$) in 2001. Section 3.2.2 provides more details on the classification of directors. Column 1 names the variable, column 2 defines the variable, and column 3 lists the data source. Variables based on CCM and CRSP are winsorized at the 1% tails. Variables from *HP* are from Hoberg and Phillips [2016]. Firms are defined as *Treatment* if its board was not composed of a majority of independent directors in 2001. Firms are defined as *Reclassifiers* if they reclassify directors as independent during the 2003-2006 period. Section 3.2.4 provides more details on the definition of *Treatment* and *Reclassifier*.

Variable	Definition	Source	N	Mean	SD	5th perc.	Median	95th perc.
Profitability	OIBDP/AT	CCM	1,869	0.14	0.10	0.00	0.14	0.30
FCF	(OIBDP-DV)/AT	CCM	1,858	0.14	0.11	0.00	0.13	0.29
Firm size	Log(AT)	CCM	1,869	6.72	1.22	4.92	6.60	9.02
Treatment	See Section 3.2.4	New	1,869	0.10	0.30	0.00	0.00	1.00
Reclassifier	See Section 3.2.4	New	1,869	0.20	0.40	0.00	0.00	1.00
Sales	SALE/AT	CCM	1,869	1.14	0.72	0.32	0.96	2.61
COGS	COGS/AT	CCM	1,869	0.73	0.63	0.08	0.56	2.09
SGA	XSGA/AT	CCM	1,869	0.26	0.23	0.00	0.21	0.72
CAPX	CAPX/AT	CCM	1,869	0.06	0.06	0.01	0.04	0.19
Acquisitions	AQC/AT	CCM	1,869	0.03	0.07	0.00	0.00	0.19
R&D	XRD/AT	CCM	1,869	0.04	0.06	0.00	0.00	0.15
Cash	CHE/AT	CCM	1,869	0.20	0.19	0.01	0.14	0.59
Return SD	SD(returns)	CRSP	1,869	0.32	0.57	0.09	0.22	0.81
HHI based on FIC 300	"FIC300HHI"	HP	1,842	0.21	0.18	0.06	0.16	0.63
HHI based on TNIC peers	"TNIC3HHI"	HP	1,857	0.22	0.19	0.05	0.15	0.62
Total Simularity of TNIC peers	"TNIC3TSIMM"	HP	1,857	3.30	3.34	1.04	2.02	9.98
Market value	Log(CSHO*PRCC_F)	CCM	1,869	7.09	1.28	5.19	6.97	9.34
Tobin's Q	(AT+CSHO*PRCC_F -CEQ-TXDB)/AT	CCM	1,765	2.30	1.55	0.96	1.83	5.27
Dividends	DV/AT	CCM	1,869	0.01	0.02	0.00	0.00	0.04

Table 3.3: Ex-ante balance

This table reports tests for balance among key covariates in the pre-treatment period in two samples. Specifically, all tests examine the equality of means, adjusting for industry fixed effects based on SIC3 codes. Columns 2 and 3 compare the treatment and control groups. Columns 4 and 5 compare the reclassifying firms to non-reclassifying firms. Section 3.2.4 provides more details on the definition of *Treatment* and *Reclassifier*. For each pair of comparison groups, two comparisons are made: First, the "Level" columns compare the 2001 value for a given variable. Second, the "Growth" columns compare the average annual pre-treatment growth for a given variable. Panel A examines the main regression sample described in Table 3.2, which focuses on firms near the 50% independent threshold (i.e. $|I - N| \leq 2$) in 2001. Panel B examines the firms that are dropped by this restriction: those far from the 50% independent threshold (i.e. $|I - N| > 2$) in 2001. All variables are defined in Table 3.2. Standard errors are robust to heteroskedasticity. T-statistics are reported in parentheses. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Ex-ante difference between:	Treatment and control		Reclassifiers and non-reclassifiers	
Ex-ante difference in:	Level	Growth	Level	Growth
Panel A: Firms near independence threshold				
Profitability	-0.017	0.144	0.042	0.150
	(-0.48)	(0.74)	(1.37)	(0.76)
FCF	-0.021	0.114	0.045	0.150
	(-0.58)	(0.56)	(1.40)	(0.71)
Sales	-0.073	0.035	0.050	-0.013
	(-0.37)	(0.56)	(0.36)	(-0.47)
CAPX	-0.002	0.227	-0.017*	0.132
	(-0.18)	(0.92)	(-1.67)	(1.13)
Return SD	-0.081	0.091	-0.085	-0.158
	(-0.98)	(0.61)	(-1.10)	(-1.03)
Q	-0.460	-0.045	0.587	-0.090
	(-1.09)	(-0.59)	(1.40)	(-1.64)
Panel B: Firms far from independence threshold				
Profitability	0.005	-0.096	-0.038*	-0.131*
	(0.22)	(-0.85)	(-1.75)	(-1.83)
FCF	0.007	-0.083	-0.043*	-0.054
	(0.28)	(-0.72)	(-1.66)	(-0.67)
Sales	-0.269*	-0.090***	0.069	0.003
	(-1.75)	(-2.73)	(1.23)	(0.17)
CAPX	-0.005	-0.033	0.003	-0.020
	(-0.40)	(-0.46)	(0.52)	(-0.72)
Return SD	0.824	0.013	0.100	0.023
	(0.88)	(0.07)	(0.64)	(0.38)
Q	-0.471*	0.024	-0.324**	-0.020
	(-1.72)	(0.58)	(-2.00)	(-0.88)

Table 3.4: Main DDD results

This table reports OLS estimates of a difference-in-difference-in-difference (DDD) model. The sample comprises firm-year observations from 1999 to 2006 where director data is available from IRRC. Additionally, the sample is restricted to firms that were close to the 50% independent threshold in the year before the first rule proposals were made. See Table 3.2 for more details on the sample. Firms are defined as a *Reclassifier* if they reclassify directors as independent during the 2003-2006 period. Section 3.2.4 provides more details on the definition of Reclassifier. Firms are defined as *Treatment* if its board was not composed of a majority of independent directors in 2001. *Post* is defined as one if the year is 2002 or later and zero otherwise. All other variables are defined in Table 3.2. Column 4 uses a different measure of *Reclassifier* based on reclassification in the 2000-2003 period. Industry fixed effects use SIC3 codes. T-statistics, clustered by firm, are reported in parentheses. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Independent variable: Reclassification years:	Prof. 2003-2006 (1)	FCF 2003-2006 (2)	Prof. 2003-2006 (3)	Prof. 2000-2003 (4)	Prof. 2003-2006 (5)
Reclassifier×Post×Treatment	0.049***	0.046***	0.047***	0.048**	0.033*
	(3.48)	(3.47)	(3.13)	(2.21)	(1.93)
Reclassifier×Treatment	-0.021	-0.021	-0.019	-0.046**	
	(-1.44)	(-1.40)	(-1.32)	(-2.03)	
Reclassifier×Post	-0.005	-0.006	-0.006	-0.016**	-0.002
	(-0.80)	(-0.82)	(-0.89)	(-2.18)	(-0.28)
Post×Treatment	-0.012	-0.015	-0.013	-0.030**	-0.013
	(-1.19)	(-1.45)	(-1.19)	(-2.56)	(-1.01)
Reclassifier	0.005	0.005	0.005	0.020***	
	(0.72)	(0.70)	(0.68)	(2.93)	
Treatment	0.003	0.002	0.003	0.026***	
	(0.31)	(0.29)	(0.36)	(2.94)	
Y(t-1)	0.727***	0.710***	0.730***	0.738***	0.392***
	(24.23)	(22.17)	(26.00)	(19.03)	(10.79)
Book Leverage			0.027***		
			(2.65)		
Log(Assets)			-0.000		
			(-0.18)		
R&D			-0.039		
			(-0.58)		
CAPX			-0.034		
			(-0.94)		
SIC3 FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE					Yes
Observations	1,869	1,847	1,869	1,800	1,841
Adj. R^2	0.63	0.64	0.63	0.62	0.68
Avg. Y	0.14	0.14	0.14	0.13	0.14

Table 3.5: Falsification tests

This table contains falsification tests of the main results from Column 1 of Table 3.4. The independent variable is *Profitability*. The sample comprises firm-year observations from 1999 to 2006 where director data is available from IRRC. Additionally, the sample is restricted to firms that were close to the 50% independent threshold in the year before the first rule proposals were made. See Table 3.2 for more details on the sample. Column 1 redefines the *Post* indicator as one beginning in 2001, and zero before. Column 1 redefines the *Post* indicator as one beginning in 2005, and zero before. Column 3 relaxes the requirement that firms be close to the 50% independent threshold in the year before the first rule proposals were made. In particular, all treatment firms are included. Column 4 examines if reclassification is associated with performance gains within the control group. All other variables are defined in Table 3.2. Industry fixed effects use SIC3 codes. T-statistics, clustered by firm, are reported in parentheses. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Test modification:	Assume treatment begins in 2001	Assume treatment begins in 2005	Include treatment firms far from cutoff	Control firms only
	(1)	(2)	(3)	(4)
Reclassifier×Post×Treatment	-0.037	0.035	0.021	
	(-0.89)	(1.60)	(1.54)	
Reclassifier×Treatment	0.042	-0.001	-0.015	
	(1.25)	(-0.12)	(-1.20)	
Reclassifier×Post	-0.000	0.000	-0.005	-0.006
	(-0.02)	(0.02)	(-0.82)	(-0.86)
Post×Treatment	-0.003	-0.007	-0.007	
	(-0.23)	(-0.79)	(-0.83)	
Reclassifier	0.002	0.001	0.005	0.005
	(0.15)	(0.29)	(0.81)	(0.64)
Treatment	-0.003	-0.003	0.000	
	(-0.26)	(-0.46)	(0.02)	
Y(t-1)	0.727***	0.727***	0.732***	0.715***
	(24.28)	(24.32)	(25.29)	(22.39)
SIC3 FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1,869	1,869	1,996	1,680
Adj. R^2	0.63	0.63	0.63	0.61
Avg. Y	0.14	0.14	0.14	0.15

Table 3.6: Additional tests

This table repeats the OLS specifications in Column 1 of Table 3.4 with different independent variables (and their corresponding lagged value). The central DDD coefficient is reported for each model, while other controls are omitted for brevity. The sample comprises firm-year observations from 1999 to 2006 where director data is available from IRRC. Additionally, the sample is restricted to firms that were close to the 50% independent threshold in the year before the first rule proposals were made. See Table 3.2 for more details on the sample. Panel A decomposes *Profitability*, which is computed in Compustat as *OIBDP/AT*, into the separate components of the denominator (*Firm size = Log(Assets)*) and numerator (*Sales, COGS, and SGA*). Panel B evaluates measures of investment and realized risk proxies. Panel C evaluates measures of competition. Panel D evaluates valuation and payout ratios. All variables are defined in Table 3.2. Each model includes industry fixed effects (SIC3) and year fixed effects. To facilitate interpretation, for each model I report the mean of the dependent variable as *Avg. Y*. T-statistics, clustered by firm, are reported in parentheses. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Decomposing Profitability Effect

Independent variable:	Firm size	Sales	COGS	SGA
Reclassifier×Post×Treatment	0.030	0.101***	0.019	0.039**
	(0.59)	(3.03)	(0.66)	(2.45)
Avg. Y	6.72	1.14	0.73	0.26

Panel B: Investment and Realized Risk

Independent variable:	CAPX	Acquisitions	R&D	Cash	Return SD
Reclassifier×Post×Treatment	0.011	0.045	0.000	0.001	0.373*
	(0.72)	(1.25)	(0.04)	(0.02)	(1.68)
Avg. Y	0.06	0.03	0.04	0.20	0.32

Panel C: Competition

Independent variable:	HHI based on FIC300	HHI based on TNIC peers	Total Simularity of TNIC peers
Reclassifier×Post×Treatment	0.005	0.025	0.493
	(0.29)	(0.79)	(1.50)
Avg. Y	0.21	0.22	3.31

Panel D: Valuation and Payout

Independent variable:	Market value	Tobin's Q	Dividends
Reclassifier×Post×Treatment	0.100	0.435	0.005
	(0.60)	(1.09)	(1.36)
Avg. Y	7.09	2.30	0.01

Bibliography

Kenneth R. Ahern and Amy K. Dittmar. The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, 127(1):137–197, 2012.

Lucian A. Bebchuk and Yaniv Grinstein. The growth of executive pay. *Oxford Review of Economic Policy*, 21(2):283–303, 2005.

Vidhi Chhaochharia and Yaniv Grinstein. Corporate governance and firm value: The impact of the 2002 governance rules. *Journal of Finance*, LXII(4):1789–1825, 2007.

Lauren Cohen, Andrea Frazzini, and Christopher J Malloy. Hiring cheerleaders: Board appointments of independent directors. *Management Science*, 58(6):1039–1058, 2012.

Jeffrey L Coles, Naveen D Daniel, and Lalitha Naveen. Boards: Does one size fit all? *Journal of Financial Economics*, 87(2):329–356, 2008.

Jeffrey L Coles, Naveen D Daniel, and Lalitha Naveen. Co-opted boards. *Review of Financial Studies*, 27(6):1751–1796, 2014.

Harold Demsetz and Kenneth Lehn. The structure of corporate ownership: Causes and consequences. *Journal of Political Economy*, 93(6):1155–1177, 1985.

Ran Duchin, John G. Matsusaka, and Oguzhan Ozbas. When are outside directors effective? *Journal of Financial Economics*, 96:195–214, 2010.

Olubunmi Falaye, Rani Hoitash, and Udi Hoitash. The costs of intense board monitoring. *Journal of Financial Economics*, 101:160–181, 2011.

Katherine Guthrie, Jan Sokolowsky, and Kam-Ming Wan. CEO compensation and board structure revisited. *Journal of Finance*, LXVII(3):1149–1168, 2012a.

Katherine Guthrie, Jan Sokolowsky, and Kam-Ming Wan. CEO compensation and board structure rejoinder, 2012b. SSRN Electronic Journal. Available at http://www.ssrn.com/abstract=2084630.

Gerard Hoberg and Gordon Phillips. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465, 2016.

Michael R Roberts and Toni M Whited. Endogeneity in empirical corporate finance, . In George Constantinides, René Stulz, and Milton Harris, editors, *Handbook of the Economics of Finance*, volume 2, Part A, chapter 7, pages 493–572. Elsevier, Amsterdam, Netherlands, 2013.