2020

Essays on Corporate Finance and Institutional Investors

Ang Li
University of Kentucky, angli.hk@gmail.com
Author ORCID Identifier:
https://orcid.org/0000-0001-7348-8367
Digital Object Identifier: https://doi.org/10.13023/etd.2020.124

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student’s advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student’s thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Ang Li, Student

Dr. Mark H. Liu, Major Professor

Dr. Paul Childs, Director of Graduate Studies

ESSAYS ON CORPORATE FINANCE AND INSTITUTIONAL INVESTORS

DISSERTATION

A dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of
Philosophy in the College of Business and
Economics at the University of Kentucky

By

Ang Li

Lexington, Kentucky

Director: Dr. Mark H. Liu, Professor of Finance

Lexington, Kentucky

2020

https://orcid.org/0000-0001-7348-8367

ABSTRACT OF DISSERTATION

ESSAYS ON CORPORATE FINANCE AND INSTITUTIONAL INVESTORS

The dissertation consists of three chapters spanning two areas of finance: corporate finance and institutional investors. In the first chapter, “Board Gender Diversity, Firm Culture, and Female Inventors”, I study how employee's Research and Development (R&D) activity is affected by firm culture. Using board gender diversity as a proxy for female-friendly culture, I find that a greater representation of women in the boardroom is associated with increased productivity and innovation output by female inventors relative to male inventors. Female inventors file more highly cited patents and less uncited patents. The effect is driven by female directors with R&D or high-tech expertise. The results support the notion that women in leadership positions cultivate a more female-friendly firm culture. In the second chapter, “Mutual Fund Preference for Pure-Play Firms” coauthored with Bradford Jordan and Mark Liu, we examine how a firm’s organization form affects mutual fund investments. We show that actively managed mutual funds avoid diluting their industry expertise by holding more pure-play firms. Specifically, mutual funds prefer firms that operate in fewer industries and firms with higher industry beta; this preference is stronger among mutual funds with greater industry expertise. We propose a measure, *Pureplayness*, as the fraction of a fund's equity invested in pure-play firms. Our results suggest that funds with higher *Pureplayness* have better risk-adjusted performance. In the third chapter, “Can Mergers and Acquisitions Internalize Positive Externalities in Funding Innovation” coauthored with Thomas Chemmanur and Mark Liu, we find that mergers and acquisitions between innovation users and innovation producers enhance innovation output. The combined firm can share the upfront costs and better appropriate the benefits associated with the innovation.

KEYWORDS: Corporate Innovation, Gender Diversity, Governance, Mergers and Acquisitions, Mutual Funds, Pure-Play Firms

Author’s signature: Ang Li

Date: March 15, 2020

ESSAYS ON CORPORATE FINANCE AND INSTITUTIONAL INVESTORS

By
Ang Li

Director of Dissertation: Dr. Mark H. Liu

Director of Graduate Studies: Dr. Paul Childs

Date: March 18, 2020

I dedicate this dissertation to my parents, Sen Li and Hua Zhao, who supported and encouraged me every step of the way.

ACKNOWLEDGEMENTS

First, I wish to thank my dissertation committee members, Mark Liu, Grant Clayton, Chris Clifford, Russell Jame, Adib Bagh, and Steven Lugauer, for their invaluable guidance and support. I would like to express my sincere appreciation to my advisor, Mark Liu, for his unwavering support and mentorship throughout my doctoral studies at the University of Kentucky. I am incredibly thankful for Mark's encouragement and patience with me during the challenging process of writing this dissertation. I must also express my gratitude to Kristine Hankins and Russell Jame for their precious insights and incisive explanations of research methods in the finance seminars over the years. I would also like to thank Bradford Jordan, Thomas Chemmanur, William Gerken, and Jay Li for their helpful advice and comments during the writing of the dissertation, and Wendy Liu and Jon Chait for their useful teaching advice.

Lastly, I would like to thank my parents, Sen Li and Hua Zhao, for their unconditional love, and Qifeng Chen, Ke Ke, Zhou Lyu, my fellow Ph.D. students, and countless others who gave me support. They are always there for me and have made this process a pleasant journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Board Gender Diversity, Firm Culture, and Female Inventors

1.1 Introduction

From 1996 to 2007, women occupied 9% of corporate board seats, and 9% of inventors filed patents.[1] Academics, media, and policymakers have widely discussed the underrepresentation of women in management positions. European countries and the state of California have adopted binding quotas to promote board gender diversity. Female directors can cultivate female-friendly firm culture (Matsa and Miller, 2011; Tate and Yang, 2015), and a good firm culture enhances employee satisfaction and firm values (Edmans, 2011; Bargeron, Lehn, and Smith, 2015). In contrast, fewer papers studied the underrepresentation of women in Research and Development (R&D) activity. This paper asks whether female directors, associated with female-friendly culture, improve the output and productivity of female inventors in innovative firms.

Using 7,500 firm-year observations from 1996 to 2007, I show a positive and significant association between the fraction of female directors and the innovation output by female inventors[2]. Female inventors' productivity, as measured by output per person, is also higher in firms with a higher fraction of female directors. The paper focuses on innovation because it is a major driving force of economic growth (e.g., Aghion and Howitt, 1992; Kogan et al., 2017). It is important to understand what fosters innovation at the individual level. Innovation output, as measured by the number of patents and the number of citations, is also a clean measure of employees' output and productivity.

The higher innovation output by females is driven by female-friendly culture and the "women helping women" channel. The focus on female-friendly firm culture stems

[1] Following previous literature on innovation, I define an inventor as an individual who holds a patent (Bell, et al., 2018; Brav, et al., 2018).

[2] This finding is consistent with previous literature that board diversity leads to higher firm-level innovation output (Bernile, Bhagwat, and Yonker, 2018; Chen, Leung, and Evans, 2018; Mayer, Warr, and Zhao, 2018; Griffin, Li, and Xu, 2019). The literature explains that female directors are better monitors and board diversity leads to robust firm policies and long-term focus. This paper offers an additional explanation.

from economics and social psychology studies. Literature shows that women at top management levels redirect firm culture, promote family-friendly policies such as onsite childcare centers, and improve the productivity of existing female employees. A greater representation of women in leadership roles also sends positive signals to women seeking senior positions, increased their likelihood of applying for the jobs, and has a positive effect on female workers' career outcomes (Matsa and Miller, 2011; Kurtulus and Tomaskovic-Devey, 2012; Tate and Yang, 2015).[3] In addition, women in female-friendly firms may face less discrimination within firms and face fewer hurdles in patenting (Ding, Murray, and Stuart, 2006; Jensen, Kovács, and Sorenson, 2018).

To identify this channel, I show that the effect of female directors is more significant on female inventors than male inventors in the same firm. I control for firm fixed effects to mitigate the concern that the results are driven by time-invariant firm characteristics, such as products marketed to women by a firm.

Firms promoting gender equity may affect the quality of innovation output in either direction. A diversified board is less tolerating of managers' poor performance; therefore, managers tend to pursue less explorative projects that generate uncertain outcomes (Adams and Ferreira, 2009; Manso, 2011). It is also easier to increase the quantity relative to the quality of innovation output. Therefore, female inventors may file more patents with the same quality or even worse quality in female-friendly firms. Alternatively, a diversified board can provide a broader range of opinions and expertise to help the firm conduct more original and explorative research. Female-friendly firms may also be more tolerant of female employees regarding failed projects. Female inventors can then pursue riskier research projects that generate more influential inventions, and it is discouraged in female-unfriendly firms with less tolerance for failure. To test the effect of the diversified board on the quality of innovation output, I follow Balsmeier, Fleming, and Manso (2017) to examine the distribution of citations.

[3] A competing theory is the "queen bee" syndrome that incompatibility exists among female workers. If women are used as tokens, and only one woman is "allowed" to ascend to the senior ranks, women will view other women more critically (Staines, Tavris, and Jayaratne, 1974). This does not apply to my paper because directors and regular employees are not in direct competition, which makes the "queen bee" syndrome less of a concern.

I find that in firms with a higher fraction of female directors, female inventors file more patents in the right tail of the distribution (receive more citations) and less in the left tail of the distribution (receive no citations).[4]

The OLS results do not imply causation due to the endogenous nature of corporate boards. Firms spend a long time choosing directors that are experienced, skillful, and complement targeted board structure (Davis, Yoo, and Baker. 2003; Nguyen and Nielsen, 2011). The demand for female directors is affected by unobservable omitted variables, such as implicit bias regarding gender (Giannetti and Wang, 2019). Therefore, firms with unobservable incentives to promote a gender diversity policy may appoint female directors as a first step. Additionally, female directors are scarce human capital, and thus they can pick the best-suited firm to serve as a director. This matching process leads to a very close association between female directors and firms' pro-gender policy. The OLS results imply that a higher fraction of female directors on the board is associated with better performance of female inventors.

To better understand the cause of female directors' impact on female inventors, I explore the heterogeneity of directors. Compared to inside directors and affiliated directors, independent directors have the responsibility to monitor and advise firm management.[5] Directors hold multiple board positions to gain experience, increase exposure, and advance their careers (Beckman and Haunschild, 2002; He and Huang, 2011). Therefore, independent directors have the incentive and ability to impact firm policy. I find evidence that a higher fraction of female independent directors correlates with more innovation output by female inventors. Next, I show that female independent directors with R&D and high-tech expertise have a larger impact on female inventors,

---

[4] Previous literature uses various measures of innovation output. Besides number of patents, number of citations is used to represent the importance and value of patents (Hall, Jaffe, and Trajtenberg, 2005). Bena and Li (2014) adjust for heterogeneity of technology classes and create a measure of patent index. Hall, Jaffe, and Trajtenberg (2001) measure originality and generality of patents based on the U.S. Patent and Trademark Office's classification of technologies. These measures serve as a proxy for the average innovation output. The distribution of patents offers a better understanding of the effect of interest.

[5] Inside directors are defined by Institutional Shareholder Services (ISS) as directors that are employees, highly paid, or beneficial owners of the company. The majority of affiliated directors are former CEO, non-CEO executives, family members, and directors that are linked to the firm through other relationships.

compared to directors without R&D expertise. The effect is significant after controlling for industry fixed effects. This finding suggests a link between female directors and female inventors. Lastly, if female directors are merely a signal of female-friendly culture, the effect on female inventors should be significant with just one female director on board. I find that it is not the case. Female inventors produce more patents in firms with two or more female directors. It suggests that female directors influence firm management after reaching a critical mass (Schwartz-Ziv, 2017).

Next, I test whether female directors proactively cultivate female-friendly firm culture through the "women helping women" channel and enhance the output of female inventors. Directors have been shown to reshape CEO compensation, investments, and redirect firm culture (Güner, Malmendier, and Tate, 2008; Chhaochharia and Grinstein, 2009; Matsa and Miller, 2011; Tate and Yang, 2015). To examine the causal effect, I use instrumental variable estimation by exploiting the role of social connection in supplying female directors to firms. Adams and Ferreira (2009) propose that the fraction of male directors with board connections to female directors in other firms is a valid instrumental variable for board gender diversity. The ratio of female directors is under 10% in my sample, indicating a limited supply of female directors. Holding the demand constant, more connections to female directors relieves the supply constraint, and firms are more likely to have female directors on board. However, one should be cautious when interpreting the results. As discussed by Jiang (2017), IV regressions estimate local average treatment effects (LATE), where firms choose to hire a female director only when the benefits of diversity outweigh its costs. The positive impact of gender diversified board does not apply to all firms and does not suggest the imposition of diversity rules on all firms. I discuss in detail the economic logic and limitations of this instrumental variable in section 1.3.

The positive impact of female directors on female inventors predicts that the stock market should react positively to female director appointments. If female directors help female inventors more than male inventors, the market reaction should be stronger when female inventors are more critical. I measure the importance of female inventors of each industry by the share of patents that have at least one female inventor. I find that in industries where female inventors are more important, firms have significant positive announcement abnormal returns around new female director appointments. In industries where female inventors are less important, firms have insignificant negative abnormal returns. The evidence is consistent with the result that gender diversity creates value for the firm.

This paper contributes to several strands of literature. Studies show that corporate culture affects firm value (Edmans, 2011; Bargeron, Lehn, and Smith, 2015), and managers shape firm culture (Tate and Yang, 2015). I find consistent evidence that female-friendly culture is associated with more productive female inventors. This paper also contributes to the literature on the board of directors. Independent directors and female directors affect firms in various meaningful ways (Adams and Ferreira, 2009; Matsa and Miller, 2011; Levi, Li and Zhang, 2013; Fahlenbrach, Low, and Stulz, 2017; Bernile, Bhagwat, and Yonker, 2018). I provide evidence to support the "women helping women" channel. Lastly, this paper is related to the literature on corporate innovation. Innovation is a major driving force of economic growth (Aghion and Howitt, 1992). Earlier studies found that innovation is affected by IPO activity, hedge fund activism, governance, and employee treatment. (Bernstein, 2015; Chen, et al., 2016; Balsmeier, Fleming, and Manso, 2017; Brav, et al., 2018). This paper confirms the corporate culture's impact on innovation.

The remainder of the chapter is organized as follows: Data and sample construction are described in Section 1.2. Section 1.3 reports the cross-sectional, within firm, and within inventor results, and discusses the heterogeneity of directors and the instrumental variable estimation. Section 1.4 concludes the chapter.

1.2 Data and Sample Construction

1.2.1 Sample Selection

The sample construction begins with all firms covered by the Institutional Shareholder Services Directors database (ISS; formerly RiskMetrics). ISS provides firm-director-year level data for all S&P 1500 firms since 1996. The data include directors' gender, classification (inside, affiliated, or independent), employment title (CEO, CFO, chairman, etc.), and employment category (executives, retired employees, consultants, etc.). The firm and director identifiers are inconsistent in the dataset, so I follow Coles, Daniel, and Naveen (2014) to create unique director identifiers. [6] I exclude observations where firms report segments in the financial sector (SIC 6000-6999) or utility sector (SIC 4900-4999) because of they face different regulations. The director dataset is then merged with Compustat to obtain the firm characteristics.

The innovation output dataset of patents and citations is constructed by NBER (Hall, Jaffe, and Trajtenberg, 2001) and extended by Kogan et al. (2017). The dataset includes the entire history of US patent documents from Google Patents. The United States Patent and Trademark Office (USPTO) allows only individuals to be the inventor, but an individual can assign granted patent to another person or a corporation. Therefore, patents always have an inventor, and sometimes they have been assigned to one or more corporations. Kogan et al. then matched the corporation names to firms in the Center for Research in Security Prices (CRSP) stock return database. Relative to the NBER patent project, this dataset provides 1.9 million patents that can be matched to companies, 27 percent of which are not included in the NBER data. The dataset covers patents granted from 1926 to 2010 that are assigned to firms in the CRSP database. It also keeps track of all citations for patents granted from 1976 to 2010.[7] The processing time of a patent is usually a few years between the filing date and the granting date. Many patents that were filed after 2007 were not granted in 2010, and thus they are not included in the dataset. For example, the number of patents in 2008 is 10% of that in 2007. To avoid bias caused by the variation of processing time of

---

[6] ISS provides two datasets before and after 2006, Directors and Directors Legacy, and uses different director identifiers in the two datasets. Using either identifier will omit director-year observations and lead to incorrect measures of board size and fraction of female directors. The unique identifier dataset is kindly provided by Lalitha Naveen on her website. The dataset also links the ISS firm identifier (CUSIP) to Compustat/CRSP firm identifier (GVKEY/PERMNO). (https://sites.temple.edu/lnaveen/data/)

[7] The dataset is provided by Noah Stoffman on his website. This dataset covers more patents and corresponding firms than NBER patents data mainly because the patent text files provided by Google have better quality than the files provided by USPTO, so more patent assignees can be identified. More details of the patent data construction can be found in the paper and the online appendix of Kogan et al. (2017). (https://iu.app.box.com/v/patents)

different patents, I drop innovation data after 2007.

Harvard Business School (HBS) Patent and Inventor database (Li et al., 2014) provides inventors' full names for each patent.[8] To identify the gender of inventors by the first name, I utilize the new-born baby names data from the Social Security Administration. It includes the frequency of first names that are given to new-born babies of each gender in each year. An inventor is coded as female (male) if more than 90% of the babies born between 22 to 60 years ago are female (male). This procedure drops inventors with rare or gender-neutral first names and allows me to identify the gender of inventors that filed more than 80% of the patents.

1.2.2 Variables

I employ various measures of innovation to capture different aspects of a firm's innovation performance, including the number of patents filed in each year and the number of citations received by patents in the subsequent years. I count patents at the time when they are filed with the USPTO because inventors have the incentives to file the patent as soon as it is finished, and thus the filing date is the closest to the actual time of innovation. I count citations after the grant date of a patent because this is when a patent is revealed to the public and starts to be cited. Because the distribution is positively skewed, I use the natural logarithm of these innovation measures.

The number of patents and citations are intuitive measures but have their limitations. They are meaningful only when used comparatively, a firm with 10 patents in one research area and another firm with 100 patents in a different area do not say which firm is more innovative. Therefore, I evaluate the patent intensity with references to some "benchmark" intensity. I follow Bena and Li (2014) to calculate the patent index of each firm. The patent index is the patent number adjusted by the median value of each technology class. In addition, I follow Hall, Jaffe, and Trajtenberg (2001) to measure the originality and generality of patents. Generality measures the level of the widespread impact of patents. Originality measures the range of fields of citations a

[8] Available at: http://dvn.iq.harvard.edu/dvn/dv/patent.

patent made.[9] In the regression analyses, I use these measures of innovation output as the dependent variables. I then merge the patent and citation data with the inventor data and calculate the measures of innovation output by gender in each firm-year.

Another commonly used measure of innovation is R&D expenditures, but 65 percent of firm-year observations in Compustat have missing values. Missing R&D expenditures in financial statements do not necessarily mean that the firm is not innovative (Koh and Reeb, 2015). Therefore, compared to R&D expenditures, patent-based metrics better reflect the productivity of R&D and more realistically reflect a firm's innovation performance.

The key independent variable is board gender diversity, measured as the fraction of directors that are female. In the multivariate tests, I include control variables that may affect firm innovation and board gender diversity at the same time. *Log(Sales)* is the natural logarithm of total sales to measure firm size. *ROA* is the return on assets, defined as the operating income before depreciation divided by total assets. *Leverage* is the total debt divided by total assets. *Tangibility* is the total gross property, plant, and equipment divided by total assets. *M/B* is the ratio of market value to book value of assets.

For firm-year observations that exist in the director dataset but are missing in the innovation dataset, I replace the missing innovation output measures with zeros. Because the study is to examine the effect on innovation, I drop firms that never filed patents during the sample period. Since boards of directors for highly regulated firms are systematically different from boards of directors of other firms (e.g., Hermalin and Weisbach, 1988), I drop all firms in the financial industry (SIC codes 6000-6999) and utilities industry (SIC codes 4900). Lastly, I drop observations where control variables have missing values. The final sample period is from 1996 to 2007, with 7,500 firm-year observations and 1,311 unique firms.

Table 1.1 reports the summary statistics of firm innovations, the board of directors, and control variables. An average firm files 43 patents a year, has 9 directors on board, and less than one of them is a female director.

---

[9] Detailed definitions are in Appendix A.

1.3 Empirical Results

1.3.1 The Determinants of Board Diversity

If board composition is endogenously determined, we would expect it to vary across firms with different characteristics. Table 1.2, Panel A reports the correlation between board gender diversity and firm characteristics to see what firms tend to have more female directors on board. The dependent variable is the fraction of directors that are female in columns 1 and 2, and the fraction of directors that are female independent directors in columns 3 and 4. The independent variables include firm characteristics (size, ROA, market-to-book ratio, etc.), board characteristics (board size and CEO's position on the board), and firm location characteristics. I control for location characteristics because board gender diversity varies by location. For example, boards of firms headquartered in New York and Florida are more diversified than firms in Texas and Oregon (Bernile, Bhagwat, and Yonker, 2018). Gender diversity also varies by industry and by year. In my sample, the Female Directors Ratio is 15.7% in the retail industry and 6.8% in the electrical equipment industry; the ratio increased from 6.9% in 1996 to 11.2% in 2007. To control for these variations in the three dimensions, I include year fixed effects and industry (Fama-French 48) fixed effects in all specifications. I add headquarter county fixed effects in columns 2 and 4.

Panel A shows the correlation between board diversity and firm characteristics rather than presenting causal relationships. The results in Table 1.1.2 show that Female Directors Ratio is higher in firms that are larger, older, pay more dividends, and have a higher market-to-book ratio. It indicates that mature firms care about board gender diversity. Larger board size correlates with more female directors. Firm local conditions do not have much of an effect on board gender diversity after controlling for firm characteristics.

In Panel B, I use a duration model to test the determinants of appointing a female director. It addresses the concern of reverse causality that firms may appoint female directors after female inventors file more patents. The "failure event" is firms adding a female director to the board for the first time during the sample period. Firms are dropped from the sample once they have a female director on the board. I also exclude firms that have female directors when they first appear in the sample. The explanatory variables include innovation output, firm characteristics, and county characteristics. I measure the innovation activity by the natural logarithm of the number of patents filed by the firm and by female inventors in the firm, and the number of citations received by those patents.

The results in Panel B show that the coefficients on the measures of innovation output are small and insignificant. It indicates that the timing of the appointment of the first female director is not related to previous innovation activity in the firm. The coefficients on *Log(Sales)* and *Log(Board Size)* are positive and significant, meaning that larges firm size and having more directors on the board increase the "hazard" of adding a female director. The coefficient on firm age is negative, and this is not surprising because the test is conditioning on firms that do not have female directors when they join the sample. Older firms are less likely to add a female director soon if they have not had one already. Other variables have insignificant coefficients, indicating that those variables (e.g., ROA, market to book ratio, cash, personal income per capita in the county) do not determine the timing of adding a female director for the first time.

1.3.2 Female Directors and Female Inventors

The first conjecture is that the higher fraction of female directors signals a female-friendly firm culture. Therefore, the innovation output of female inventors should be higher in firms with a more diversified board, and the increase is larger relative to male inventors. To test this prediction, I run the following OLS regressions with an interaction term:

$$\begin{aligned} Innovation_{i,g,t} = \beta_0 + \beta_1 Female\ Director\ Ratio_{i,t} + \beta_2 Female\ Dummy_g \\ + \beta_3 Female\ Director\ Ratio_{i,t} \times Female\ Dummy_g \\ + \beta_4 Controls_{i,t} + FE + \epsilon_{i,t} \end{aligned} \tag{1.1}$$

where *Female Director Ratio* is the number of female directors divided by the total number of directors on the board. *Female* is a dummy that equals one for innovation output by women and equals zero for innovation output by men. The sample is at the firm-year-gender level. Table 1.3 presents the baseline results. Regressions in Panel A include all control variables, industry (Fama-French 48) fixed effects, headquarters county fixed effects, and year fixed effects. Panel B replaces industry and county fixed effects with firm-gender fixed effects. Standard errors are clustered at the firm level.

The results in Table 1.3, Panel A show that board gender diversity is associated with higher innovation output by female inventors, compared to male inventors, and the association is statistically significant in four out of five specifications. In column 1, the coefficient on the Female dummy is -1.126, indicating that the number of patents filed by females is 32.43% (=exp(-1.126)) of the number of patents filed by males. It is consistent with the summary statistics that only 15% of the patents have a female inventor. The coefficient on Female Directors Ratio indicates that having one more female director is associated with 6.2% (=(1/9.093)*0.564) more patents filed by male inventors and 10% more patents filed by female inventors. The different effect of board gender diversity on female and male inventors is significant when innovation output is measured by citations, tech-class adjusted patents, generality, and originality. Panel A uses industry fixed effects to explore the cross-sectional variations. The results may be biased by omitted variables. For example, firms may have more women leaders and more female inventors if their products are marketed to women, compared to other firms in the same industry.

In Panel B, I control for firm-gender fixed effects and find similar results, indicating that the results in Panel A cannot be explained by omitted time-invariant firm characteristics. Within a firm, having more female directors on the board is associated with more innovation output by female inventors. Results in column 1 indicate that one more female director is associated with 2.9% (=(1/9.093)*(-0.222+0.482)) more patents filed by female inventors. The magnitude is smaller than Panel A because board gender diversity has smaller within-firm variations compared to cross-sectional variations. However, within-firm tests do not control for omitted time-varying factors, for example, firm culture changes to be more female-friendly, and it leads to more female directors and more female inventors. The results in Table 1.3 should be interpreted as an association, rather than a causal relation, between female directors and female inventors. The next question is whether female directors affect female inventors on the extensive margin or intensive margin, i.e., whether the higher output by females is driven by more female inventors or more productive female inventors. Table 1.4 presents results to answer this question. In Panel A, column 1, the dependent variable is the logarithm of net hiring of inventors by gender, and the regression controls for industry fixed effects. An inventor is counted as a new hire if she files a patent for a firm for the first time during the sample period. Results show that higher Female Directors Ratio is associated with more hiring of female inventors, compared to make inventors. In column 2, the dependent variable is the logarithm of the total number of inventors, and the regression controls for firm-gender fixed effects. An inventor works for a firm between the first year and last year she files a patent for the firm. There is a significant association between female directors and the number of female inventors in the firm.

In Panel B, the outcome variable is the productivity of inventors by gender. Regressions in columns 1 and 2 use firm-level data, where I measure productivity by the logarithm of the number of patents (citations received by patents) divided by the number of inventors. Results in column 1 show that one more female director is associated with 6.8% (=(1/9.093)*0.622) more patents filed by female inventors compared to male inventors. The difference is statistically significant. Regressions in columns 3 and 4 use inventor level data, where I measure productivity by the logarithm of the number of patents (citations received by patents) filed by each inventor. The regressions include inventor fixed effects to adjust for omitted time-invariant factors at the inventor level. The result in column 3 shows that an average female inventor files 0.9% (=(1/9.093)*(0.0805)) more patents when a firm has an additional female director, compared to male inventors. The magnitude of the inventor-level result is smaller than the firm-level result because a large sample, while it increases the power of regressions, also introduces more noise. For example, inactive innovators that filed one patent during the sample period will attenuate the estimate. Nevertheless, the coefficient on the interaction term is statistically significant at the 10% level in column 3. Results in

Table 1.4 indicate that female directors are positively associated with both the number and the productivity of female inventors.

1.3.3 The Distribution of Citations

The previous subsection shows female-friendly firm culture's effect on the total number of patents filed and the total number of citations received by female inventors. Still, the effect on the distribution of citations is unclear. In other words, female inventors may generate more breakthrough inventions or inventions with no contribution. On the one hand, female inventors may focus on quantity over quality of patents because the number of patents is a simple countable measure of their performance. On the other hand, female-friendly firms may have a higher tolerance for failed projects conducted by female employees. Therefore, female inventors may pursue risky research and generate more important patents and fewer failed inventions.

To test these two possible outcomes, I follow Balsmeier, Fleming, and Manso (2017) to model the number of patents of different levels of importance. I categorize patents into four groups: 1) Breakthrough patents that receive citations within the top 1% among all patents in the same technology class and application year; 2) Important patents that receive citations within the top 2% - 10% among all patents in the same technology class and application year; 3) Incremental patents that receive citations outside the top 10% among all patents in the same technology class and application year, and; 4) Failed inventions that receive zero citations. I calculate the number of patents in each of the four groups that are filed by male inventors and female inventors, respectively. Table 1.5 reports the corresponding results.

The regressions control for firm-gender fixed effects. The estimates in column 1 show that female inventors file more breakthrough patents in firms with more female directors on board. Female inventors also file more important patents, but the effect is statistically insignificant, as shown in column 2. Results in columns 3 and 4 find that female inventors file significantly fewer patents that receive zero citations and file more patents that have a marginal contribution. The results are consistent with the hypothesis that female inventors conduct riskier research, create fewer failed patents and more important inventions.

1.3.4 Directors Heterogeneity

In this subsection, I examine the effect of different types of female directors on female inventors. Directors are categorized by their affiliation with the firm into inside directors, affiliated directors, and independent directors. Independent directors are responsible for monitoring and advising top managers, so they have the channel to redirect firm culture and impact female inventors. I then calculate the fraction of directors that are female independent directors and the fraction of directors that are female non-independent directors. The test results are reported in Table 1.6, Panel A. The dependent variable is the innovation output by female inventors scaled by that of male inventors. The result in column 1 shows that more female independent directors are associated with more patents filed by female inventors compared to male inventors. The results are robust to different measures of innovation output in columns 2 to 5.

Independent directors contribute to a firm not only with managerial experience but also specialized expertise. The homophily theory predicts that independent female directors with R&D experience or experience in high-tech firms have a larger impact on female inventors. I follow Knyazeva, Knyazeva, and Masulis (2013) to define R&D/Tech experts as independent directors whose primary firm has positive R&D expenses or is in high-tech industries[10]. I then calculate the fraction of directors that are female independent with (without) R&D/Tech experience. Results in Panel B shows that the effect is driven by R&D/Tech experts.

In Panel C, I estimate the marginal effect of an additional female director and find that the results are driven by firms with two or more female directors on board. The results are consistent with the idea that the number of women needs to reach a critical mass to

[10] Baginski, Hassell, and Kimbrough (2004) define high-tech firms as firms in industries with SIC codes of 2833-2836, 3570-3577, 3600-3674, 7371-7379, or 8731-8734. A few examples of high-tech industries are medicinal chemicals, semiconductors, and commercial biological research.

affect corporate culture. Given that finding a qualified director is time-consuming (185 days according to Nguyen and Nielsen (2011)), it is costly for a firm to hire two female directors just for the signaling effect. This finding supports the view that female directors proactively redirect corporate culture and impact female inventors.

1.3.5 IV Regressions

OLS regressions do not control for time-varying omitted firm characteristics. This endogeneity may bias the OLS estimations in either direction. On the one hand, firms finding it costly to promote gender equality at the employee level may use female directors merely as a signal. On the other hand, firms seeing the benefits of diversity may promote female-friendly culture and hire female directors and female inventors at the same time. OLS estimates are biased upwards in the latter case and may explain the baseline results. Reverse causality is also a concern. Firms with better performing female employees may have the incentive to promote female workers to management positions.

In this subsection, I use an instrumental variable to estimate the relation between female directors and female inventors. A valid instrumental variable needs to satisfy the relevance requirement and the exclusion restriction. Table A.2 of the appendix offers some insight on how to find a relevant instrument. ISS provides the primary employer of directors that sit on two or more boards. I then match it to firm characteristics to compare directors' primary firm and nonprimary firms.

Table A.2, Panel A shows that the size of a female director's primary firm is $9,010 million, and the size doubled for her nonprimary firm. For male directors, nonprimary firms are also larger than primary firms, but by a much small amount. The difference between female and male directors is the same for other firm characteristics such as market-to-book ratio, ROA, firm age, and dividend yield. The summary statistics are consistent with the fact that female directors are a scarce human resource, and the supply is limited. Female directors can pick a suitable firm to join. Panel B reports the distance between directors' primary and nonprimary firms. The average distance is over

600 miles, and 31.4% of firm pairs are under 60 miles. This finding is consistent with Knyazeva, Knyazeva, and Masulis (2013) who find that director supply is strongly affected by the local director labor market.[11] Male and female directors do not consider distance differently. Panel C examines the probability of a director joins another firm in the same industry. Only 7% of the primary-nonprimary firm pairs are in the same 3-digit SIC industry; 12% of firm pairs are in the same 2-digit SIC industry or supplier-customer related industries[12]. This result indicates that directors and executives in the same industry are not the main supply of independent directors. It is not surprising because the director's managerial or specialized experience can be applied across industries.

Panel D examines whether the primary and nonprimary firm of a director is connected by another male director. Medland (2004) and Adams and Ferreira (2009) point out the lack of social connection of female directors, and a female director can be observed by another firm through the connection of another director. Results in Panel D support this argument. For female directors, 86% of her primary firm and nonprimary firm are connected by another male director. This number is 9% for male directors. It suggests that social connection is critical for firms to locate potential directors.

Therefore, I follow Adams and Ferreira (2009) to define the instrumental variable as the fraction of male directors that have a connection to female directors on boards of other nonlocal firms. I require other firms to be more than 60 miles away to ensure that the supply pool does not reflect the local economics condition and the local supply of female inventors. The instrumental variable, Connected Male Directors Ratio, may correlate with local directors' gender diversity and therefore violates the exclusion restriction. I provide the following evidence to mitigate this concern. First, the correlation between connected male directors and local diversity is 0.04. Second, I

[11] Knyazeva, Knyazeva, and Masulis (2013) include local law firms and local financial institutions as potential supply pool for independent director. Their measure of local supply should be higher than 31.4%.

[12] I follow Ellis, Fee, and Thomas (2018) to define supplier-customer related industries. I use industry input/output data provided by the U.S. Bureau of Economic Analysis (BEA) to find the major supplier industry and major customer industry.

include local directors' gender diversity in the first stage and the second stage of the 2SLS regression. After controlling for local diversity, Connected Male Directors Ratio still positively affects board gender diversity. Third, I include county fixed effects and find similar results. Another possibility is that male directors' connection to female directors correlates with industry effects, e.g., the pharmaceutical industry has more productive female inventors, and its male directors are connected to more female directors at the same time. Therefore, I control for industry fixed effects.

Table 1.7. Panel A reports the IV regression results. Column 1 presents the results of the first-stage regression. *Connected Male Directors Ratio* is significantly correlated with more female independent directors on board. In columns 2 to 6, the results show that innovation output by female inventors is significantly higher in firms with a higher fraction of female independent directors. In Panel B, the dependent variable is the innovation output by female inventors scaled by male inventors. The results suggest that a higher female independent director ratio leads to more innovation output by female inventors compared to male inventors.

1.3.6 Announcement returns of female director appointments

Board gender diversity improves female inventors' performance, but does it create value for shareholders? I employ the event study methodology to analyze stock returns around announcements of female director appointments. Previous studies find significant positive announcement returns around appointments of independent outside directors (Masulis and Mobbs, 2011) and negative returns around sudden loss of independent directors (Nguyen and Nielsen, 2010). Farrell and Hersch (2005) document an insignificant announcement return of a woman added to the board. If female directors have a larger impact on female inventors than male inventors, we should expect that the market reacts stronger for firms where female inventors are more important. I measure the importance of female inventors by the share of patents that are filed by women in an industry-year. I then divide firms into five groups based on the measure of importance. Using ISS data, I have 1,849 events of a female director joining a firm for the first time, 284 events of firms in the top quintile, and 289 events of firms in the bottom quintile.

To investigate the market reaction to new female directors, I use daily returns of a 7-day window from CRSP for each event (from day -3 to day +3). Day 0 is the event day of annual meetings. I use the Capital Asset Pricing Model (CAPM) and the Fama-French-Carhart four-factor model (Fama and French, 1993; Carhart, 1997) to calculate the abnormal returns. I estimate the factor model using a 60-day estimation window. I require firms to have at least 45 days of data available in the estimation window and apply a 30-day gap between the estimation window and event window. I then calculate the daily risk-adjusted abnormal returns and cumulative abnormal returns (CARs) in the 7-day window. Consistent with the prediction, I find that firms in the top quintile have an average CAR of about 1%, and firms in the bottom quintile have a negative average CAR. Figure 1.1 shows this result. Positive market responses suggest that board gender diversity is value-enhancing for firms where female inventors are important assets.

1.4 Conclusion

Firm policy regarding gender diversity is a particularly important topic. In the past 30 years, more and more organizations start to promote diversity. European countries, including Norway, Spain, and France, have adopted quota rules to promote gender diversity in corporate leadership positions. In 2018, California became the first state in the U.S. to adopt gender quota rules. The literature finds mixed results regarding the effect of gender diversity on firm performance (e.g., Ahern and Dittmar, 2012; Eckbo, Nygaard, Thorburn, 2019). This paper contributes to the discussion by examining the channels through which female directors could affect firm value. I hypothesize and provide evidence that female directors contribute to firms by cultivating a more female-friendly culture that enhances the productivity and output of female employees.

This paper examines the effect of female-friendly corporate culture, associated with female directors, on female employee’s Research and Development (R&D) activity. I measure R&D activity by patents and citations to quantify to effect. This paper finds that a larger fraction of female directors is associated with higher productivity and innovation output of female inventors, compared to male inventors. The results are significant after controlling for industry fixed effects, firm fixed effects, or inventor fixed effects. The effect is driven by female directors with R&D expertise or experience from high-tech industries and exists in firms with two or more female directors. The finding is supported by the critical mass theory that the influence of a subgroup grows when a certain threshold is reached (Torchia, Calabrò, and Huse, 2011). It is consistent with the conjecture that female directors proactively redirect firm culture.

Following Adams and Ferreira (2009), I use instrumental variable regressions and find a causal relationship between female directors and female inventors. The results offer another piece of evidence to support the homophily theory that women in leadership positions cultivate more female-friendly culture. However, we need to be cautious when interpreting the results. The IV regression estimates a local average treatment effect (LATE), where firms choose to hire a female director only when the benefits of diversity outweigh its costs (Jiang, 2017). The positive effect of gender diversified board does not suggest impositions of diversity on all firms.

1.5 Tables

Table 1.1: Summary statistics

This table reports the summary statistics of key variables from 1996 to 2007. The sample includes S&P 1500 firms that are covered by the Institutional Shareholder Services Directors database (formerly RiskMetrics). The corporate innovation output data are from NBER and Kogan et al. (2017). Firm characteristics and county-level control variables are from Compustat and the Census Bureau. Innovative firms that filed at least one patent during the sample period comprise the sample. The final sample consists of 7,500 firm-year observations with 1,311 unique firms. The variable definitions are in Table A.1.

	Mean	Median	St. Dev.
Innovation Measures			
#Patents	42.187	3.000	184.853
#Citations	391.598	10.000	2500.147
#Patents - Citation weighted	86.883	5.008	400.225
Patent Index	27.611	2.500	113.678
Generality	0.357	0.180	0.377
Originality	0.487	0.637	0.379
Board of Director Measures			
Board Size	9.093	9.000	2.560
#Female Directors	0.868	1.000	0.872
#Female Independent Directors	0.775	1.000	0.823
Firm Controls			
Log(Sales)	7.270	7.186	1.609
ROA	0.134	0.137	0.099
Leverage	0.214	0.209	0.167
R&D/Assets	0.046	0.023	0.059
M/B	2.152	1.694	1.403
Tangibility	0.517	0.436	0.337
Dividend Yield	0.012	0.003	0.026
Cash/Assets	0.158	0.079	0.182
Log(Age)	2.992	3.045	0.837
County Controls			
Population	13.701	13.726	1.021
Personal Income Per Capita	10.389	10.399	0.180
Personal Income Per Capita Growth	4.502	4.700	2.024

Table 1.2: Determinants of the board gender diversity

The table tests the determinant of the board gender diversity. Panel A reports the panel regression of the female director ratio on firm and headquarters county characteristics. The dependent variable is the number of female (independent) directors divided by the total number of directors on the board. All regressions include industry fixed effects and year fixed effects. Headquarter county fixed effects are included in columns 2 and 4. Panel B reports the duration model where the "failure event" is the appointment of a female director in a firm for the first time. Firms are dropped from the sample once they have a female director on the board. I also exclude firms that have female directors when they first appear in the sample. The variable definitions are in Table A.1. The t-statistics in the parentheses are adjusted for firm-clustering effect. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Panel A: Determinants of board gender diversity: the OLS model

	(1)	(2)	(3)	(4)
	Female Directors Ratio		*Female Independent Directors Ratio*	
Log(Sales)	0.0176***	0.0148***	0.0182***	0.0158***
	(11.34)	(8.711)	(12.61)	(9.745)
ROA	-0.0107	-0.00627	-0.00408	0.00183
	(-0.497)	(-0.325)	(-0.230)	(0.111)
Leverage	0.0200*	0.0157	0.0240**	0.0159
	(1.707)	(1.455)	(2.222)	(1.567)
R&D/Assets	0.0317	0.0224	0.0865**	0.0735
	(0.587)	(0.449)	(1.987)	(1.634)
M/B	0.00251*	0.00300**	0.00270**	0.00239*
	(1.697)	(2.101)	(2.149)	(1.913)
Tangibility	0.0157**	0.00842	0.0132*	0.00467
	(2.084)	(1.123)	(1.939)	(0.706)
Dividend Yield	0.198***	0.127**	0.163***	0.107**
	(3.010)	(2.433)	(2.827)	(2.247)
Cash/Assets	-0.000979	-0.00189	-0.00855	-0.00321
	(-0.0684)	(-0.141)	(-0.642)	(-0.262)
Log(Age)	0.00862***	0.00487*	0.00869***	0.00670**
	(3.182)	(1.768)	(3.305)	(2.557)
Log(Board Size)		0.0298***		0.0181***
		(4.211)		(2.675)
CEO is Chair or President		0.00917**		0.00881**
		(1.964)		(2.076)
Population		0.0251		0.0323
		(0.842)		(1.201)
Personal Income Per Capita		-0.0225		-0.00995
		(-0.350)		(-0.164)

Table 1.2 (continued)

Personal Income Per Capita Growth		0.000252		0.000707
		(0.301)		(0.902)
Observations	7,500	7,482	7,500	7,482
Adjusted R^2	0.255	0.395	0.264	0.400
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
County FE	-	Yes	-	Yes

Table 1.2 (continued)

Panel B: Timing of adding the first female director: the duration model

	(1)	(2)	(3)	(4)
Log(1+Patents)	-0.0132			
	(-0.273)			
Log(1+Patents by Female)		-0.00768		
		(-0.0800)		
Log(1+Citations)			0.0168	
			(0.561)	
Log(1+Citations by Female)				0.0171
				(0.362)
Log(Sales)	0.200***	0.197***	0.188***	0.190***
	(3.442)	(3.374)	(3.232)	(3.250)
ROA	-0.229	-0.228	-0.212	-0.202
	(-0.341)	(-0.339)	(-0.314)	(-0.299)
Leverage	-0.661*	-0.658*	-0.652*	-0.655*
	(-1.879)	(-1.870)	(-1.850)	(-1.858)
R&D/Assets	-1.357	-1.511	-1.926	-1.747
	(-0.921)	(-1.043)	(-1.304)	(-1.225)
M/B	0.0581	0.0576	0.0536	0.0537
	(1.253)	(1.238)	(1.146)	(1.145)
Tangibility	0.0181	0.0162	0.0129	0.0108
	(0.105)	(0.0939)	(0.0745)	(0.0624)
Dividend Yield	-0.635	-0.654	-0.706	-0.653
	(-0.222)	(-0.228)	(-0.242)	(-0.228)
Cash/Assets	0.165	0.164	0.159	0.159
	(0.373)	(0.370)	(0.359)	(0.360)
Log(Age)	-0.165*	-0.165*	-0.164*	-0.164*
	(-1.824)	(-1.823)	(-1.813)	(-1.809)
Log(Board Size)	2.255***	2.251***	2.242***	2.243***
	(9.605)	(9.615)	(9.591)	(9.611)
CEO is Chair or President	-0.0552	-0.0572	-0.0669	-0.0612
	(-0.318)	(-0.329)	(-0.384)	(-0.352)
Population	-0.0628	-0.0631	-0.0644	-0.0637
	(-1.370)	(-1.377)	(-1.395)	(-1.384)
Personal Income Per Capita	-0.101	-0.0985	-0.0768	-0.0906
	(-0.252)	(-0.247)	(-0.191)	(-0.227)
Personal Income Per Capita Growth	-0.0211	-0.0208	-0.0202	-0.0204
	(-0.765)	(-0.752)	(-0.732)	(-0.740)
Observations	3,954	3,954	3,954	3,954
Chi-square	159.7	159.3	157.6	158.3

Table 1.3: Effects of female directors on innovation output

The table presents panel regression results estimating the effect of female director ratio on innovation output by female inventors. I use a sample of all S&P 1500 firms from 1996 to 2007 covered by the ISS Directors database. The variable definitions are in Table A.1. The model examines the effect of female director ratio on innovation output by female inventors and male inventors, using firm-year-gender level data and an interaction term. *Female Directors Ratio* is the number of female directors divided by the total number of directors on the board. *Female* is a dummy that equals one for innovation output by females and equals zero for innovation output by males. In Panel A, all regressions include industry (Fama-French 48) fixed effects, headquarters county fixed effects, and year fixed effects. In Panel B, all regressions include year fixed effects and firm-gender fixed effects. Detailed definitions of all variables are in Table A.1. The t-statistics in the parentheses are adjusted for firm-clustering effect. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Panel A: Cross-sectional tests

	(1)	(2)	(3)	(4)	(5)
	Log(1+Patents)	*Log(1+Citations)*	*Patent Index*	*Generality*	*Originality*
Female Directors Ratio	0.564	0.157	0.535	-0.0106	-0.00811
	(1.574)	(0.316)	(1.578)	(-0.149)	(-0.103)
Female	-1.126***	-1.673***	-0.737***	-0.224***	-0.290***
	(-29.93)	(-31.36)	(-29.02)	(-25.79)	(-29.92)
Female Directors Ratio × *Female*	0.346	1.760***	0.639***	0.328***	0.462***
	(1.228)	(4.551)	(3.351)	(5.036)	(6.241)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	15,000	15,000	15,000	15,000	15,000
Adjusted R^2	0.600	0.542	0.582	0.479	0.463
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes

Table 1.3 (continued)

Panel B: Within-firm tests

	(1)	(2)	(3)	(4)	(5)
	Log(1+Patents)	*Log(1+Citations)*	*Patent Index*	*Generality*	*Originality*
Female Director Ratio	-0.222	-1.233***	-0.0805	-0.0449	-0.00553
	(-0.912)	(-2.620)	(-0.344)	(-0.565)	(-0.0762)
Female Director Ratio × *Female*	0.482***	2.053***	0.309*	0.217***	0.113
	(2.589)	(5.241)	(1.789)	(2.934)	(1.531)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	14,782	14,782	14,782	14,782	14,782
Adjusted R^2	0.891	0.792	0.878	0.695	0.710
Year FE	Yes	Yes	Yes	Yes	Yes
Firm × Gender FE	Yes	Yes	Yes	Yes	Yes

Table 1.4: Extensive margin and intensive margin

The table presents OLS regression results to examine the effect of female directors on the number of inventors and the innovation output per inventor. In Panel A, the outcome variable is the logarithm of the number of net hiring and the total number of inventors for each firm-year-gender. In Panel B, the outcome variable is the productivity of female and male inventors, as measured by the logarithm of the number of patents produced per inventor. In columns 1 and 2, the data is at the firm-year-gender level. In columns 3 and 4, the data is at the inventor-year level, and the regressions control for inventor fixed effects. *Female Directors Ratio* is the number of female directors divided by the total number of directors on the board. *Female* is a dummy that equals one for innovation output by females and equals zero for innovation output by males. Detailed definitions of all variables are in Table A.1. The t-statistics in the parentheses are adjusted for firm-clustering effect. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Panel A: Number of inventors

	(1)	(2)
	Log(1+#Net Hiring)	*Log(1+#Inventors)*
Female Directors Ratio	-1.079***	-0.271
	(-3.383)	(-1.013)
Female	0.0681***	
	(2.827)	
Female Directors Ratio × *Female*	0.991***	0.594***
	(4.491)	(2.684)
Firm and HQ county controls	Yes	Yes
Observations	12,760	14,782
Adjusted R^2	0.0597	0.889
Year FE	Yes	Yes
Industry FE	Yes	-
County FE	Yes	-
Firm × Gender FE	-	Yes

Table 1.4 (continued)

Panel B: Productivity of inventors

	Firm Level		Inventor Level	
	(1)	(2)	(3)	(4)
	Log(1+#Patents/ #Inventors)	*Log(1+#Citations/ #Inventors)*	*Log(1+#Patents)*	*Log(1+#Citation)*
Female Directors Ratio	-0.207***	-0.628***	-0.120*	-0.303*
	(-3.512)	(-3.227)	(-1.848)	(-1.733)
Female	-0.113***	-0.464***		
	(-10.81)	(-18.04)		
Female Directors Ratio × *Female*	0.622***	1.800***	0.0805*	0.321**
	(7.815)	(9.719)	(1.654)	(2.018)
Firm Controls	Yes	Yes	Yes	Yes
Observations	15,000	15,000	479,615	479,615
Adjusted R^2	0.298	0.363	0.288	0.338
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	-	-
County FE	Yes	Yes	-	-
Inventor FE	-	-	Yes	Yes

Table 1.5: Breakthrough patents or incremental patents

The dependent variables are the logarithm of one plus the number of patents that fall in different bins of the citation distribution within the patent class and application year. The bins are classified as top 1% percentile, top 2-10% percentile, cited but not in the top 10%, and without any citations. The model uses firm-year-gender level data. *Female Directors Ratio* is the number of female directors divided by the total number of directors on the board. *Female* is a dummy that equals one for innovation output by females and equals zero for innovation output by males. All regressions include year fixed effects and firm-gender fixed effects. Detailed definitions of all variables are in Table A.1. The t-statistics in the parentheses are adjusted for firm-clustering effect. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

	(1)	(2)	(3)	(4)
	Log(1+Patents)	*Log(1+Patents)*	*Log(1+Patents)*	*Log(1+Patents)*
	Top 1%	Top 2-10%	Cited, not in top 10%	Without citations
Female Director Ratio	-0.208*	-0.104	-0.729***	0.393*
	(-1.829)	(-0.656)	(-2.769)	(1.720)
Female Director Ratio × *Female*	0.175*	0.134	1.183***	-0.804***
	(1.864)	(1.029)	(5.954)	(-4.565)
Firm and HQ county controls	Yes	Yes	Yes	Yes
Observations	14,782	14,782	14,782	14,782
Adjusted R^2	0.676	0.826	0.840	0.790
Year FE	Yes	Yes	Yes	Yes
Firm × Gender FE	Yes	Yes	Yes	Yes

Table 1.6: Directors heterogeneity

The table presents the results of OLS regressions to examine the effect of independent directors and directors with R&D and technology expertise. The outcome variables are innovation output by females scaled by innovation output by males. In Panel A, *Female Independent (Non- Independent) Directors Ratio* is the fraction of directors that are independent (non-independent) females. In Panel B, *R&D/Tech Experts Ratio* is the fraction of directors that are independent females who have corporate experience at firms with positive R&D or high-tech firms. In Panel C, the key independent variables are three dummies that are equal to one if a firm has one female director, two female directors, or three or more female directors, respectively. All regressions include industry (Fama-French 48) fixed effects, headquarter fixed effects, and year fixed effects. Detailed definitions of all variables are in Table A.1. The t-statistics in the parentheses are adjusted for firm-clustering effect. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Panel A: Independent female directors

	(1)	(2)	(3)	(4)	(5)
	Log(1+Patents)	*Log(1+Citations)*	*Patent Index*	*Generality*	*Originality*
	F/M	F/M	F/M	F/M	F/M
Female Independent Directors Ratio	0.159**	0.205***	0.191**	0.313***	0.200**
	(2.288)	(2.801)	(2.023)	(2.902)	(2.348)
Female Non-Independent Directors Ratio	-0.158	-0.181	-0.101	0.0747	0.00722
	(-1.225)	(-1.533)	(-0.641)	(0.453)	(0.0508)
Firm and HQ county controls	Yes	Yes	Yes	Yes	Yes
Observations	7,483	7,485	7,483	7,489	7,491
R-squared	0.306	0.304	0.329	0.315	0.321
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes

Table 1.6 (continued)

Panel B: Directors with R&D and high-tech expertise

	(1)	(2)	(3)	(4)	(5)
	Log(1+Patents)	*Log(1+Citations)*	*Patent Index*	*Generality*	*Originality*
	F/M	F/M	F/M	F/M	F/M
R&D/Tech Experts Ratio	0.335***	0.293**	0.463***	0.423***	0.358**
	(2.979)	(2.385)	(3.041)	(2.669)	(2.269)
Non-Experts Ratio	0.0735	0.123	0.0656	0.253**	0.0856
	(0.957)	(1.533)	(0.639)	(2.012)	(0.935)
Firm and HQ county controls	Yes	Yes	Yes	Yes	Yes
Observations	7,465	7,467	7,465	7,471	7,473
R-squared	0.359	0.347	0.385	0.372	0.370
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes

Table 1.6 (continued)

Panel C: Marginal effect of an additional female director

	(1)	(2)	(3)	(4)	(5)
	Log(1+Patents)	*Log(1+Citations)*	*Patent Index*	*Generality*	*Originality*
	F/M	F/M	F/M	F/M	F/M
#Female Directors = 1	0.00791	0.0170	0.0123	0.0262	0.0269*
	(0.736)	(1.375)	(0.801)	(1.602)	(1.817)
#Female Directors = 2	0.0471***	0.0427**	0.0580**	0.0910***	0.0383*
	(2.785)	(2.343)	(2.536)	(3.419)	(1.799)
#Female Directors >= 3	0.0498*	0.0636**	0.0486	0.104***	0.0827**
	(1.873)	(2.153)	(1.495)	(3.023)	(2.514)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	7,465	7,467	7,465	7,471	7,473
R-squared	0.359	0.347	0.384	0.374	0.370
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes

Table 1.7: Instrumental variable regressions

The table presents the results of two-stage least squares regressions to examine the causal relationship between female independent director ratio and innovation output by female inventors. The instrumental variable, *Connected Male Directors Ratio*, is the fraction of male directors who have connections to female directors in nonlocal firms that are more than 60 miles away. In Panel A, column 1 reports the first-stage result, where I regress *Female Independent Directors Ratio* on the instrumental variable and control variables. In the second stage, the dependent variables are the measures of innovation output by female inventors. I regress these measures on the predicted *Female Independent Directors Ratio*, and the results are reported in columns 2 to 7. In Panel B, the dependent variable in the second stage is the ratio of female innovation output divided by male innovation output. I regress this ratio on the predicted *Female Independent Director Ratio*, and the results are reported in columns 1 to 6. All regressions include industry (Fama-French 48) fixed effects and year fixed effects. I include local director diversity instead of headquartering county fixed effects. Detailed definitions of all variables are in Table A.1. The t-statistics in the parentheses are adjusted for firm-clustering effect. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Panel A: Innovation output by female inventors

	(1)	(2)	(3)	(4)	(5)	(6)
	Female Independent Directors Ratio	*Log(1+Patents)*	*Log(1+Citations)*	*Patent Index*	*Generality*	*Originality*
Connected Male Directors Ratio	0.0293***					
	(2.780)					
Female Independent Directors Ratio		17.31**	22.28**	20.53**	3.957**	6.171**
		(2.282)	(2.117)	(2.202)	(2.214)	(2.536)
Local Female Directors Ratio	0.116**	-1.454	-1.736	-1.995	-0.340	-0.681
	(2.083)	(-0.845)	(-0.765)	(-0.908)	(-0.968)	(-1.344)
Firm and HQ county controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,360	7,343	7,345	7,343	7,352	7,351

Table 1.7 (continued)

Adjusted/Centered R^2	0.279	-0.462	-0.170	-0.322	-0.375	-0.984
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
IV F-stat	29.902					
Underidentification test (P-value)	0.006					

Panel B: Innovation output by female inventors, scaled by male inventors

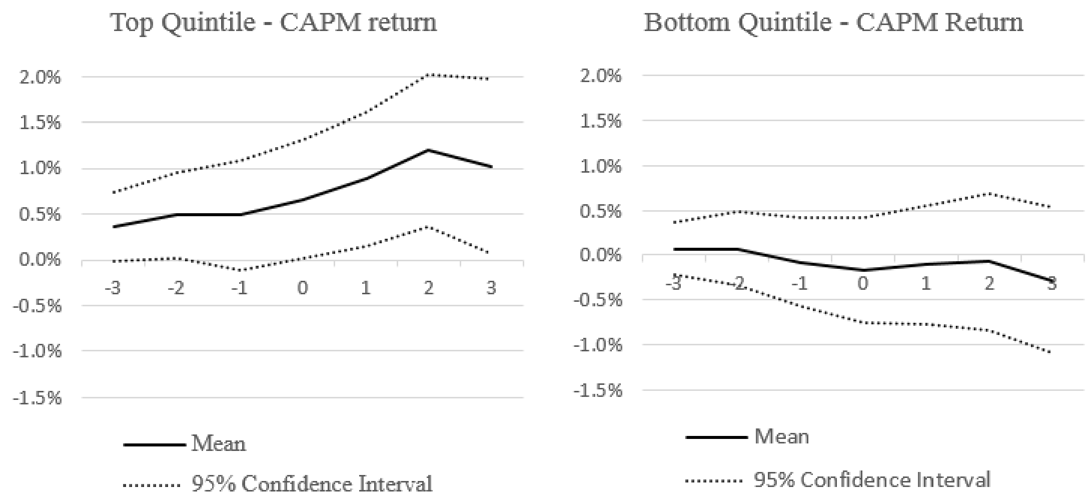
	(1)	(2)	(3)	(4)	(5)
	Log(1+Patents)	*Log(1+Citations)*	*Patent Index*	*Generality*	*Originality*
	F/M	F/M	F/M	F/M	F/M
Female Independent Directors Ratio	2.619**	2.195**	4.747**	4.696**	6.830**
	(2.351)	(2.161)	(2.396)	(2.165)	(2.464)
Local Female Directors Ratio	-0.369*	-0.201	-0.686	-0.382	-0.851
	(-1.750)	(-1.111)	(-1.598)	(-0.890)	(-1.470)
Firm and HQ county controls	Yes	Yes	Yes	Yes	Yes
Observations	7,343	7,345	7,343	7,352	7,351
Adjusted / Centered R^2	-0.642	-0.298	-0.652	-0.312	-0.728
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes

1.6 Figures

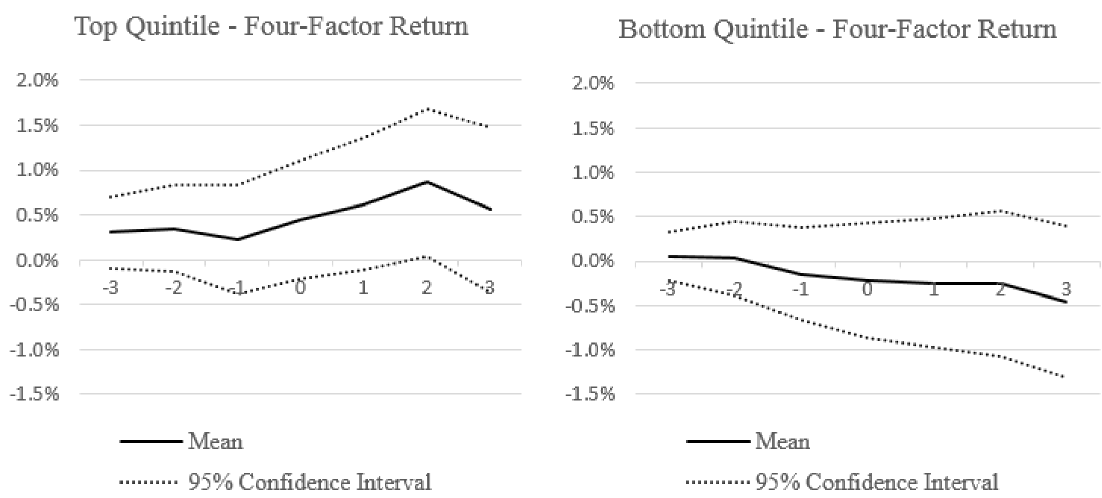
Figure 1.1: Announcement returns of female director appointments

This figure shows the cumulative abnormal returns (CARs) over a 7-day window [-3, 3] around appointments of a new female director, where announcements occur on day 0. The sample contains 1,849 events of a female director joining a firm for the first time between 1997 and 2007. I sort firms into quintiles based on the importance of female inventors of an industry to which a firm belongs. The importance is measured by the share of patents that are filed by women in that industry. The graphs on the left use firms in the top quintile (284 events), and the graphs on the right use firms in the bottom quintile (289 events). Panel A reports the CAPM cumulative returns, and Panel B reports the Fama-French-Carhart four-factor cumulative abnormal returns.

Panel A: CAPM returns, top quintile versus bottom quintile



Panel B: Four-factor returns, top quintile versus bottom quintile



Chapter 2 Mutual Fund Preference for Pure-Play Firms

2.1 Introduction

The literature has found that corporate organizational form impacts a firm in many ways such as firm valuation, systematic risk, cash holdings, and payout policy (e.g., Villalonga, 2004; Duchin, 2010; Hann, Ogneva, and Ozbas, 2013; Jordan, Liu, Wu, 2018). To date, however, no studies examine how organizational form affects mutual fund investment decisions. Are corporations with only one segment preferred by mutual funds compared to conglomerates that have diversified business? If yes, why? And do mutual fund characteristics contribute to such preference related to organizational forms?

This paper attempts to answer these questions. We find that actively managed mutual funds have a preference for pure-play firms over conglomerates.[1] The magnitude is large: mutual fund ownership (*MFO*) of conglomerates is 1.029% lower than that of a portfolio of pure-play firms in the same segments and of similar size. We measure *MFO* by the fraction of shares outstanding held by actively managed mutual funds, and in our sample, the average *MFO* is 7.657%. The 1.029% difference in *MFO* between conglomerates and a portfolio of pure-play firms represents a 13% deviation from the mean.

To find out the reason behind mutual funds' preference for pure-play firms over conglomerates, we test the "industry expertise hypothesis." Actively managed mutual funds may have expertise in specific industries. In fact, an average mutual fund in our sample invests 60% of its equity into 14 out of 360 industries classified by the 3-digit SIC code. Therefore, by holding pure-play firms in industries in which they have the expertise, mutual funds can invest 100% of their money in these industries. In contrast, if they invest in conglomerate firms, their investment will be diluted by industries in

---

[1] Throughout the paper, we use conglomerates and multi-segment firms interchangeably. We do the same with pure-play firms, pureplays, and single-segment firms.

which they may not have expertise, unless the conglomerate operates in the desired industries by chance. As a result, we should observe mutual funds prefer firms that help them exploit the industry expertise, and the preference should be more pronounced among funds with a higher level of industry expertise. Funds utilizing such strategies should have better performance and exhibit industry selecting and industry timing ability.

Consistent with the predictions of the industry expertise hypothesis, we find that firms that operate in more segments and have more diversified sales have lower mutual fund ownership. We also find that mutual funds prefer stocks with greater industry beta and industry R-squared. Further, we divide actively managed mutual funds into five groups based on the industry concentration of their holdings, following Kacperczyk, Sialm, and Zheng (2005). We find that only the top two quintiles of mutual funds show a preference for pure-play firms over conglomerates, while the other three quintiles do not. Both firm-level and fund-level cross-sectional variations show strong support that mutual funds' industry expertise drives the preference for pureplays.

To further test that the preference is caused by mutual funds' expertise other than fundamental differences between the two organizational forms, we perform a falsification test and a difference-in-difference test. First, we construct a "pseudo conglomerate" for each actual conglomerate in our sample using a portfolio of pure-play firms in the same segment and of similar size. These pseudo conglomerates serve as control firms. Our falsification test shows that actual conglomerates have significantly lower *MFO* than pseudo conglomerates, indicating that controlling for firm characteristics that are associated with industry and size does not alter our results. Second, we consider a sample of mergers and acquisitions (M&As) and find a decrease in *MFO* around conglomerating M&As. Because we examine the change in *MFO* of the same firm around an M&A, we capture the within-firm change and control for time-invariant firm characteristics. Further, the decrease in *MFO* is more pronounced when the degree of segment sales concentration has a larger drop after the M&A, consistent with the industry expertise hypothesis.

If the preference for pure-play firms is a proxy for industry expertise, we should expect mutual funds in our sample to benefit from this preference. Using mutual fund quarterly holdings data, we create an index, *Pureplayness*, that measures a mutual fund's fraction of assets invested in pure-play firms. Our results show that this measure is very persistent within funds over time. This finding suggests that the preference for pure-play firms is driven by systematic fund characteristics, such as mutual fund managers' informational advantage in specific industries that tend to persist over time. We then test what fund characteristics determine the *Pureplayness*. For example, we find that higher *Pureplayness* is associated with a higher turnover ratio and higher expenses. Furthermore, we construct five portfolios of funds based on their *Pureplayness* level and find that funds in the top quintile of *Pureplayness* outperform funds in the bottom quintile by 1.7% per annum after adjusted for risks represented by the six-factor model of Fama and French (2014) and Carhart (1997). Since *Pureplayness* and Industry Concentration Index (Kacperczyk, Sialm, and Zheng, 2005) both measure industry expertise, we control for Industry Concentration Index and find similar results. Lastly, we follow Daniel, et al., (1997) and measure the characteristic-based returns. The results indicate that the superior performance of high-*Pureplayness* funds is mainly driven by their industry picking and industry timing ability, which is consistent with our industry expertise hypothesis.

Actively managed mutual funds having industry expertise and higher demand for pure-play firms are the result of comparing mutual funds to other investors. Is it true that mutual funds have industry expertise even when compared to hedge funds and other institutional investors? Results in Table B.2 of the appendix show that hedge funds and other institutional investors have a weaker and non-robust preference towards pure-play firms. The finding suggests that the preference for pure-play firms are especially strong for mutual funds relative to other institutional investors.

Our study contributes to several different branches of finance literature. First, we document how the organizational form of a firm affects mutual fund holdings. Our results suggest that a firm's organizational form affects the degree of industry information reflected in stock price. This, in turn, impacts how actively managed mutual funds invest in the firm's stock because this affects how much of their investments will be allocated to the industry in which they have expertise.

Second, we contribute to studies on mutual funds by documenting their preference for pure-play firms, and by offering another piece of evidence of their industry expertise. The existing literature documents that institutional investors favor stock characteristics that fit their needs. For example, liquidity and voting rights are valued by institutional investors (Falkenstein, 1996; Bushee and Noe, 2000; Li, Ortiz-Molina, and Zhao, 2008) because stock liquidity lowers the cost of fire sales, and voting rights offer monitoring power to the institutional investors. In the same spirit, we examine whether actively managed mutual funds prefer pure-play firms to conglomerates, and what characteristics cause this preference. Furthermore, Kacperczyk, Sialm, and Zheng (2005) show that mutual funds generate higher returns if they concentrate their holdings in a few industries. Our results support the notion that some mutual funds exhibit industry expertise. They exploit this expertise by investing more in single-segment firms and earn higher returns.

The remainder of the chapter is organized as follows: We discuss the related literature in Section 2.2. Data and sample construction are described in Section 2.3. Our firm-level results are reported in Section 2.4. After controlling for industry fixed effects, time fixed effects, and firm characteristics, we find higher mutual fund ownership in pure-play firms. This finding survives the falsification test and tests in an M&A setting. Section 2.5 reports the fund-level results. Higher pureplay holding predicts higher risk-adjusted returns. Section 2.6 concludes the chapter.

2.2 Literature Review

This paper is related to several strands of literature. It complements the research on firm boundaries and corporate diversification. Since Coase (1937) proposed that firm boundaries affect resource allocation, there have been numerous studies related to firm boundaries. The earlier literature documents a “diversification discount,” suggesting that the internal capital market is inefficient (Lang and Stulz, 1994; Berger and Ofek, 1996; Rajan, Servaes, and Zingales, 2000). The over-active internal capital market and information asymmetries between central management and divisional managers lead to strategic information transmission and agency problems in diversified firms (Harris, Kriebel, and Raviv, 1982; Denis, Denis, and Sarin, 1997; Seru, 2014). Other studies show a "bright side" of diversification. After considering the endogeneity problem, the literature finds a value premium associated with firm diversification (Campa and Kedia, 2002; Villalonga, 2004; Gasper and Massa, 2011; Maksimovic and Philips, 2013). The internal capital market offers coinsurance among segments. It allows conglomerates to hold significantly less cash, pay out more dividends than pureplays, have lower systematic risk, lower costs of external financing, and lower skewness exposure (Duchin, 2010; Mitton and Vorkink, 2010; Hann, Ogneva, and Ozbas, 2013; Jordan, Liu, and Wu, 2018). To our knowledge, this paper is the first one to link mutual fund investments to corporate organizational form.

This paper is also related to the literature on institutional investors. The existing literature shows that mutual funds prefer more liquid stocks, larger stocks, stocks with higher price levels, and stocks that have a single share class (Falkenstein, 1996; Li, Ortiz-Molina, and Zhao, 2008). Our results show that organizational forms of firms can also affect mutual funds' investment decisions. Our paper focuses on actively managed mutual funds to alleviate the concerns that other institutional investors might face investment constraints regarding organizational forms.

Finally, our research contributes to the literature that explores mutual fund skills. Utilizing the quarterly holding filings, earlier literature has developed various measures of manager skill, including characteristic-adjusted returns, stock selectivity, industry concentration index, tracking errors, active shares, etc. (Daniel, et al., 1997; Wermers, 2000; Pinnuck, 2003; Kacperczyk, Sialm, and Zheng, 2005; Cremers and Petajisto, 2009). More recent papers use institutional investors' daily trading data to measure manager styles (Chakrabarty, Moulton, and Trzcinka, 2017; Russell, 2017; Farrell, 2018). We find evidence that mutual funds' holding of pure-play firms is a proxy of industry expertise, and it is positively related to future performance.

2.3 Data and Sample Construction

The sample construction begins with all firms with data reported to the Compustat Industry Segment database. A firm is required by the U.S. Securities and Exchange Commission (SEC) to report information for segments that represent 10% or more of the consolidated firm's sales. We focus on business segments and use the latest source year of each segment-year observation. We follow the Jordan, Liu, and Wu (2018) sample selection criteria to exclude observations where firms report segments in the financial sector (SIC 6000-6999), utility sector (SIC 4900-4999), or firms with a market capitalization less than $10 million. We only include stocks whose 8-digit CUSIP identifier ends with either 10 or 11, to make sure they are common stocks. Following Hann, Ogneva, and Ozbas (2013), we define a firm as a pure-play firm if it has only one 3-digit SIC segment; otherwise, we define it as a conglomerate.

The mutual fund data come from the CRSP Survivor-Bias-Free Mutual Fund Database. The database includes information on total net assets, investment objectives, investment styles, fund holdings, and other fund characteristics. We follow Jordan and Riley (2015) and Pastor, Stambaugh, and Taylor (2015) to build our sample of actively managed U.S. equity funds. We drop funds that 1) are identified by CRSP as index funds, EFTs, or annuities; 2) have a fixed income Lipper asset code; 3) are sector funds; 4) have terms in their name not associated with unleveraged, active, or equity investment.[2] We require the funds to have more than 80% of assets invested in equity and report expense ratios above 0.1% per year, since it is improbable that any actively managed funds would charge such low fees. We also require the funds to be identified as domestic equity funds by CRSP because we are analyzing their behavior concerning U.S. firms with different organizational forms.

The stockholdings of mutual funds are collected both from reports filed by mutual funds

[2] The list of terms we use in our name search is available on request. For example, we drop mutual funds that have the word "Index", "S&P", or "Sector" in their fund names. The final dataset for the main analysis does not include any equity domestic sector funds (CRSP objective code begins with "EDS" or Lipper classification code indicates that the fund is a sector fund). In untabulated tests, sector funds show a significant preference towards pure-play firms.

with the SEC and from voluntary reports generated by the funds. Most of the holdings of companies are listed on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the NASDAQ. Mutual funds are required to report their holdings to the SEC at the end of each quarter. Some funds choose to report their holdings even more frequently, on a monthly basis. To be consistent with the majority of funds, we construct our sample on a quarterly basis. For mutual funds that report holdings more frequently than every quarter, we use the end-of-quarter values. To measure the mutual fund ownership (*MFO*), we calculate the percentage as the number of shares held by mutual funds divided by the total number of shares outstanding.

We obtain accounting information such as the book value of equity and earnings from Compustat. Firm age and stock return information are from CRSP. For firms whose data exist in Compustat and CRSP but are missing in the mutual fund holding dataset, we replace the missing *MFO* values with zeros. Our final sample period is from 2003 to 2015, with 107,501 firm-quarter observations, 4,783 unique firms, of which 1,159 are conglomerates.

In our multivariate analyses, we use mutual fund ownership (*MFO*) as the dependent variable. Our key test variable is a firm's organizational form, *PurePlay*, which equals one if the firm is a pure-play firm that operates in only one 3-digit SIC code industry, and zero otherwise. Following Falkenstein (1996) and Li, Ortiz-Molina, and Zhao (2008), we include variables that affect institutional investment decisions. Market capitalization, *Mkt cap*, is defined as the dollar value of all outstanding shares at the end of the quarter. The quarterly return on the firm's stock, *Return*, is defined as the stock return over the quarter. The quarterly share turnover ratio, *Turnover*, is defined as the ratio of the trading volume to the number of shares outstanding of the quarter. The market-to-book ratio, *M/B*, is defined as the market value of assets divided by the book value of assets. Financial leverage, *Leverage*, is computed as the ratio of total debt to the market value of assets, following Baker, Stein, and Wurgler (2003). Share price, *Price*, is defined as the stock price at the end of the quarter. Return on assets, *ROA*, is net income divided by total assets. *Spread* is the average of the monthly trading spread of the stock in the quarter. The legal environment that institutions face as fiduciaries may also lead to the preference related to organizational form. Del Guercio (1996) suggests that firm age, dividend yield, S&P membership, and stock price volatility have appeared in the prudence case law that affect institutional equity investments. The dividend yield, *Divyield*, is the ratio of total dividends to the market value of equity. *S&P500* is a dummy equal to one for S&P 500 index stocks. The volatility of stock returns, *Retvol*, is the stock return volatility using monthly stock returns over the previous 12 months. Firm age, *Age*, is the number of years since the firm first appears in CRSP.

Table 2.1, Panel A reports the summary statistics of mutual fund ownership (*MFO*) and control variables across conglomerates and pureplays. The univariate tests show that mutual fund ownership is higher in conglomerates; it is not surprising because mutual funds prefer larger and more liquid stocks, which are the characteristics of conglomerates. The majority of firm characteristics are significantly different, suggesting that firm and stock characteristics differ substantially between conglomerates and pureplays. We control for these variables that potentially affect mutual fund ownership in the multivariate regressions. Panel B reports the correlation matrix of these variables.

2.4 Firm Level Analysis

2.4.1 Mutual fund ownership in conglomerates versus pureplays

To estimate the effect of organizational form on the investment decision of actively managed mutual funds, we regress mutual fund ownership (*MFO*) on the *PurePlay* dummy and other firm characteristics as control variables. Unless otherwise specified, we include the quarter fixed effect, and the Fama-French 48 industry fixed effect in all regressions. Standard errors are clustered at the firm and quarter levels (2-way clustering). The coefficient on *PurePlay* captures the difference of mutual fund ownership between conglomerates and pureplays. Table 2.2 reports the results from the following regression.

$$MFO_{i,t} = \beta_0 + \beta_1 PurePlay_{i,t} + \beta_2 Controls_{i,t} + FE_{ind} + \delta_t + \epsilon_{i,t} \qquad (2.1)$$

In Table 2.2, we run OLS panel regressions. We do not include any control variables in the first specification. The coefficient on *PurePlay* is negative, which is consistent with the summary statistics in Table 2.1. However, after we add more control variables in specification 2 to 5, the coefficients on *PurePlay* become positive and statistically significant at the 1% level in all four models, indicating that the average mutual fund ownership in pure-play firms is higher than that of conglomerates. The coefficients converge to 0.615 in column 5, suggesting that after controlling for firm and industry characteristics, mutual funds ownership is 0.615% higher in pure-play firms than conglomerates.

2.4.2 Falsification tests

Results in Table 2.2 indicate that mutual fund ownerships are different between conglomerates and pureplays even after controlling for many firm characteristics such as size, firm performance, leverage, and book to market ratio. However, it remains the case that the difference could be due to unobservable omitted variables. For example, a lack of future investment opportunities may lead a firm into diversification while actively managed mutual funds tend to avoid such firms. We would observe a difference in mutual fund ownership between conglomerates and pureplays even though the difference is not caused by organizational form. To address these concerns, we follow previous literature to perform a falsification test (Campello, 2002; McNeil, Niehaus, and Powers, 2004; Jordan, Liu, and Wu, 2018).

For each segment of an actual conglomerate, we select a single segment firm in the segment's industry with the closest value of book assets, total sales, or operating profit, or imputed market capitalization. Because market capitalization is not available for segments, we use each segment's book asset multiplied by the conglomerate's market-to-book ratio to get the imputed market capitalization of each segment. We then construct a "pseudo-conglomerate" for each actual conglomerate in the first year the conglomerate appears in our sample, with the pseudo conglomerate having the same portfolio of segments as the actual conglomerate. We use book value weighted variables of each single-segment firm[3] to create the variables for pseudo conglomerates. Because pure-play firms have higher market-to-book ratio than conglomerates (as shown in Table 2.1), a match by book value might lead to a higher market value of pseudo-conglomerates, and higher market value correlates with higher mutual fund ownership. To address this concern, we also construct pseudo-conglomerates using imputed market capitalization. We then run the following regression using all actual conglomerates and pseudo conglomerates:

$$MFO_{i,t} = \beta_0 + \beta_1 Actual_{i,t} + \beta_2 Controls_{i,t} + \psi_i + \delta_t + \epsilon_{i,t} \tag{2.2}$$

where $Actual_{i,t}$ takes value 1 for actual conglomerates and 0 for pseudo-conglomerates. The standard errors are clustered at the firm and year-quarter levels (two-way clustering). Table 2.3 reports the results.

In column 1, we use book assets to construct the pseudo-conglomerates, and the coefficient on the *Actual* dummy is -1.029 and statistically significant at the 1% level, indicating that the mutual fund ownership for actual conglomerates is about 1.029% lower than that of the pseudo conglomerates. Since the average *MFO* in our sample is 7.657% (untabulated), the difference represents a 13% deviation from the mean value. In columns 2-4, we use sales, operating profits, and imputed market capitalization to construct the pseudo-conglomerates. Results are all statistically significant and qualitatively similar.

In summary, the falsification test in Table 2.3 shows that conglomerates have lower mutual fund ownership than portfolios of similarly sized stand-alone firms in the same segments. The result indicates that the difference in mutual fund ownership between conglomerates and pureplays cannot be explained by characteristics that are purely related to industry and firm size that happen to be correlated with mutual fund ownerships.

[3] Using sales or operating profit to calculate weights gives similar results.

2.4.3 Explanations of the preference

In this section, we ask the question of why mutual fund investors differ from other investors in their demand for single-segment firms. We propose and find evidence to support the industry expertise hypothesis.

The industry expertise hypothesis posits that actively managed mutual funds exhibit industry expertise. For example, Kacperczyk, Sialm, and Zheng (2005) find that mutual funds with greater industry concentrations in their stock holdings perform better than other mutual funds. By definition, the values of pure-play firms in a particular industry are affected by one industry factor (plus market movement and idiosyncratic movements), but not industry factors in other industries. In contrast, the values of conglomerates are affected by two or more industry factors. When active mutual funds have private information about a particular industry, they will naturally prefer pureplays to conglomerates since pureplays allow these mutual funds to invest 100% of their money in that industry, while conglomerates will dilute their investment. The industry expertise hypothesis predicts that the preference is stronger for active mutual funds who exhibit higher industry concentration. Furthermore, different pure-play firms in an industry are affected by the industry factor to varying degrees. Firms with a higher industry beta are influenced more by the industry factor and should be more valuable in utilizing funds' industry expertise. Accordingly, the industry expertise hypothesis has three testable predictions:

1) *MFO* is lower in more diversified firms.

2) The preference for pure-play firms is stronger for mutual funds with a higher level of industry expertise.

3) *MFO* is higher in firms with higher industry beta or industry R-squared.

To begin with, we create two measures of firm diversification. *#Segments* is the number of business segments a firm operates in. Holding other things constant, a firm that operates in more segments is likely to be more diversified. The second measure is the Herfindahl index of sales of various segments in a firm. A lower Herfindahl index signifies more diversification of a conglomerate. By definition, both the *#Segments* and the *Herfindahl Index* of a pureplay are equal to one. In the sample, the *Herfindahl Index* ranges from 0.22 to 1.00, with a mean value of 0.92. In Table 2.4, Panel A, columns 1 and 3 show that when the level of diversification is higher, the mutual fund ownership is lower. However, one may think that the positive coefficients are driven by the organizational form instead of the level of diversification. To address this concern, we test the effect of diversification in a subsample of conglomerates in columns 2 and 4. The results show that the incremental increase of diversification correlates with a lower level of mutual fund ownership.

Next, we propose another measure of firm diversification. Instead of using 3-digit SIC code, we use 1-digit SIC code or 2-digit SIC code to define conglomerates[4]. We define a firm as a pureplay if it operates in only one 1-digit SIC industry; otherwise, it is defined as a conglomerate. In this case, we can mitigate the limitation of 3-digit SIC codes being too narrow or firms mistakenly report their industry. Besides, under the new definition, the segments of a conglomerate are less correlated. According to the first prediction, we expect to see that mutual funds prefer pure-play firms, and the magnitude is larger than that under the definition using 3-digit SIC code. We report the results in Panel B. The coefficient increases from 0.596 in column 1 to 1.012 in column 3, confirming that mutual funds prefer pure-play firms that are more focused.

To test the second prediction, we place mutual funds into groups based on their industry concentration levels. We use the Industry Concentration Index (*ICI*) measure proposed by Kacperczyk, Sialm, and Zheng (2005) to quantify the extent of portfolio concentration in 10 broadly defined industries. We use the primary SIC code of firms to find the corresponding 10 industries. This index is a proxy of how much a mutual fund's holding deviates from the market portfolio. The Industry Concentration Index at time t for a mutual fund is defined as the sum of the squared deviations of the value weights for each of the 10 different industries held by the mutual fund, $W_{j,t}$, relative to

[4] The first digit of the SIC code indicates the division, the first two digits indicate the major group, the first three digits indicate the industry group, and the fourth digit indicates the specific industry sector.

the industry weights of the total stock market, $\overline{W}_{j,\,t}$:

$$ICI_j = \sum_{j=1}^{10}(W_{j,t} - \overline{W}_{j,t})^2 \tag{2.3}$$

This index is equal to zero if a mutual fund has exactly the same industry composition as the market portfolio and increases as a mutual fund becomes more concentrated in fewer industries. The Industry Concentration Index is related to the Herfindahl index, which is commonly used in industrial organizations literature to measure the concentration of companies in an industry. The Industry Concentration Index can be thought of as a market-adjusted Herfindahl Index.[5]

We then divide actively managed mutual funds into 5 quintiles in each quarter and calculate the ownership by mutual funds with different levels of industry concentration. For example, *MFO_ICI1* is the percentage of shares held by mutual funds in the lowest quintile of industry concentration index in each quarter. We run regressions of mutual fund ownership on the *PurePlay* dummy and control variables, and the results are reported in Table 2.4, Panel C. The results show that the coefficients on the *PurePlay* dummy are insignificant in column 1 to 3, and significantly positive in the last two columns. It indicates that more concentrated mutual funds prefer pure-play firms more than less concentrated mutual funds, and the increase is monotonic. It is consistent with our second prediction and supports the industry expertise hypothesis.

We further test Prediction 3 of the industry expertise. If a firm is affected by the industry factor to a higher degree, mutual funds should prefer that firm to take advantage of its expertise in that industry. To measure the industry betas of stocks, we follow Liu (2011) to run the following regression:

$$R_{n,t} = \alpha_n + \beta_n^M * R_t^M + \beta_n^I * \left(R_t^I - \hat{\beta}_n^{IM} * R_t^M\right) + \varepsilon_{n,t} \tag{2.4}$$

where R_t^I is the return on the industry portfolio in day t for stock n's industry. We exclude stock n when we construct the industry portfolio to prevent spurious correlations between firms and industry returns in industries with a small number of

[5] Because ICI is calculated using primary SIC codes of firms, ICI and holding of pureplays are not necessarily mechanically correlated. A mutual fund can hold only conglomerates but still have a high ICI if those conglomerates' primary segments are in the same industry.

firms. We use the market value of stocks in day *t–1* to calculate the weighted industry portfolio return. R_t^M is the return on the CRSP value-weighted market index, and $\hat{\beta}_n^{IM}$ is the market beta of stock *n*'s industry, estimated from the following regression in each calendar year:

$$R_t^I = \alpha_n^I + \beta_n^{IM} * R_t^M + \varepsilon_{n,t}^I \tag{2.5}$$

Industry R-squared measures the importance of market and industry movement relative to the idiosyncratic movement in determining stock returns. It is derived from the following regression:

$$R_{n,t} = \alpha_n + \beta_n^I * \left(R_t^I - \hat{\beta}_n^{IM} * R_t^M\right) + \varepsilon_{n,t} \tag{2.6}$$

We then test the prediction that mutual fund ownership is higher in firms that are more sensitive to the industry factor, as measured by greater industry betas and higher values of industry R-squared statistics. The result in Table 2.4, Panel D supports our prediction. The coefficients on the *Indbeta* and *IndRsq* are positive, meaning that when firms are more affected by industry factors, the mutual fund ownership is higher.

2.4.4 Other institutional investors

The fractions of a stock held by mutual funds, other institutions, and retail investors must sum up to one. If the demand for organizational forms were identical among mutual funds and other investors, mutual fund ownership should be identical across conglomerates and pureplays. Section 4.3 shows that mutual funds' advantage in industry expertise leads to their higher demand for pureplay firms. However, is it true that mutual funds have greater industry expertise even compared to other institutional investors such as hedge funds and pension funds?

Table B.2 of the appendix reports the preference of hedge funds and other institutional investors for pure-play firms. We define other institutions as institutional investors excluding actively managed mutual funds and hedge funds. The baseline regressions and falsification tests in Panel A and Panel B show an insignificant and not robust preference. It means that the higher mutual fund ownership in pure-play firms is driven by mutual funds' higher demand for pure-play firms compared to retail investors.

Results in Panel C and Panel D show that industry concentration of hedge fund holdings and other institutions holdings do not affect their demand for pure-play firms. Passive investors and indexers are possibly the reason why other institutional investors, on average, do not prefer pure-play firms.

Hedge funds not showing industry expertise are possibly due to two reasons. First, hedge fund investments may require more firm-specific knowledge than industry-level knowledge. Brav, Jiang, Partnov, and Thomas (2008) find that hedge funds increasingly engage in shareholder activism. Because hedge funds are subjected to fewer regulations than mutual funds and pension funds, they can hold highly concentrated positions in fewer companies and influence corporate management and monitoring. Moreover, the 13F filings only report long positions, which is a small part of hedge funds' 12 investment strategies[6], as defined by Joenväärä, Kosowski, and Tolonen (2012). The top 4 strategies of hedge funds are long/short, multi-strategy, emerging markets, and commodity trading advisor (CTA). The number of hedge funds in the Long Only category accounts for 2% of their sample. Even for the Long/Short hedge funds, we do not know their long positions are simple buys or short-covering positions. Given the above information, industry expertise and organizational form may be a minor factor for hedge fund investments, and hedge funds do not show a preference for pure-play firms in their long positions.

2.4.5 Mutual fund ownership around M&As

To further identify organizational form's effect on mutual fund investments, we examine the mutual fund ownerships around the change in organizational forms. We use firms in conglomerating mergers as the treated group and firms involved in non-conglomerating mergers as the control group. This way, our analysis captures the within-firm effect and mitigates the concern that our results are driven by the time-invariant firm-specific characteristics within each industry. Given that the M&As data

[6] The 12 strategies are: CTAs, Emerging markets, Event driven, Global macro, Long only, Long/short, Market Neutral, Multi-strategy, Relative value, Sector, Short bias, and others.

provide exact effective dates, and the mutual fund ownership data are at a quarterly frequency, we know in which quarter a conglomerating M&A took place. We can then introduce "event time dummies" as an indicator of the quarters relative to the treatment. In Figure 2.1, *t+4* means that the sample includes the firm-quarter observations in the 4th quarter after and before the M&A effective quarter (two observations for the M&A); *t+5* indicates the firm-quarter observations in the 5th quarter before and after the M&A effective quarter, and so on. The regression model is shown below.

$$MFO_{i,t} = \beta_0 + \beta_1 After_{i,t} + \beta_2 After_{i,t} \times Conglo\ M\&A_{i,t} + \beta_3 Controls_{i,t} + \psi_i + \delta_t + \epsilon_{i,t} \quad (2.7)$$

where *After* is a dummy that equals one if it is after the M&A effective quarter and 0 otherwise; *Conglo M&A* is a dummy that equals one if it is a conglomerating M&A. To be classified as a conglomerating M&A, we require that either the target firm or the acquiring firm to be a pure-play firm in the year before the M&A, and the merged firm is a conglomerate in the year after the M&A. Accordingly, the interaction term measures the effect of conglomerating M&As on the treated firm after the treatment. The M&A sample contains 22,518 observations, which includes 574 M&As in total and 32 conglomerating M&As.

Figure 2.1, Panel A plots the coefficients and the confidence intervals on the interaction term using various periods around the M&As. The coefficients become significantly negative when we use observations 6 and 7 quarters before and after the M&A. This finding shows that conglomerating M&As do affect the mutual fund ownership, but it takes time for mutual funds to react to the change in organizational form. The decrease of the institutional ownership occurs more than one year after the treatment.[7]

Next, we check the industry expertise hypothesis by examining how much of the mutual fund ownership change around conglomerating M&A is due to the change in diversification. We start by creating a dummy HHI drop that equals one if the change of Herfindahl index of sales around the M&A is smaller or equal to zero. Then we

[7] In untabulated robustness tests, we define conglomerates using 1-digit or 2-digit SIC codes, the decrease of *MFO* occurs in the 4th quarter after conglomerating M&As, and the magnitude is similar. It is consistent with our hypothesis that conglomerates have less mutual fund ownership when they are more diversified.

include interaction terms *After * Conglo M&A * HHI drop* and *After * Conglo M&A * (1 - HHI drop)* in the regressions. This difference-in-difference-in-differences specification and allows us to disentangle the effect coming from the change of firm diversification. As shown in Figure 2.1, Panel B, the coefficients are negative and statistically significant on *After * Conglo M&A * HHI drop* starting the 6th quarter. In contrast, the coefficients in Panel C on *After * Conglo M&A * (1 – HHI drop*) are insignificant, meaning that the change of diversification explains the change of mutual fund ownership around conglomerating M&As. The result in Figure 2.1 alleviates the concerns that our findings are driving by firm heterogeneity within industries and offer another piece of evidence that supports the industry expertise hypothesis.

2.5 Fund Level Analysis

Results in section 4 show that actively managed equity mutual funds, especially the ones with industry expertise, have a significant preference for pure-play firms. The next question is whether mutual funds benefit from the industry expertise by holding more pure-play firms. We construct a new variable, *Pureplayness*, that measures the share of total net assets invested in pure-play firms. We define it as:

$$Pureplayness = \sum_{j=1}^{N}(w_{\mathrm{j}} \times Pureplay_j) \tag{2.8}$$

where w_j is the portfolio weight of asset j in the fund, and $Pureplay_j$ is a dummy that equals one for pure-play firms.[8]

Table 2.5 reports the summary statistics on fund characteristics. The sample is from 2003 to 2015 and includes 62,062 fund-quarter observations and 1,928 unique funds. The average fund has a *Pureplayness* of 0.617, meaning that an average mutual fund invests 61.7% of its equity in pure-play firms. There is dispersion in *Pureplayness*

[8] Previous literature using Compustat's segment files usually drop firms if the sum of the segment sales deviate more than 1% from the total net sales of the firm (Berger and Ofek, 1995; Seru, 2014). This filter removes 1/3 firm-year observations, and thus many mutual fund holdings cannot be matched to the segment file, and it leads to an imprecise measure of *Pureplayness*. Therefore, we ignore this sales filter and define a firm as conglomerate as long as it reports more than 1 segment. Firms have little incentive to report a segment with no business activity in it.

among funds. The *Pureplayness* ranges from 0.345 to 0.897, and the standard deviation is 0.115.

2.5.1 Persistence of the *Pureplayness*

If *Pureplayness* is causing the deviation in a fund's returns from its peers, funds should exhibit relatively persistent *Pureplayness* over time. For example, if a fund displays high *Pureplayness* due to the manager's deep understanding of an industry or her ability to select outperforming industries, its *Pureplayness* will likely to remain high in the future, no matter the manager continues to invest in the same industry or switch to another industry.

To exam whether the *Pureplayness* is persistent, we sort all funds into quintile portfolios by a fund's average *Pureplayness* in the previous 8 quarters. We then calculate the average *Pureplayness* in each quintile in the current quarter, subsequent one quarter, two quarters, and one to three years. Since we use data in the past to construct portfolios and calculate *Pureplayness* based on all funds that exist in the future, there is no look-ahead bias. Table 2.6 reports the results. The future *Pureplayness* of the high-*Pureplayness* portfolio remains higher than that of the low-*Pureplayness* portfolio in all time horizons. The gap between the high and low portfolio shrinks over time but remains significant after three years. The results indicate a persistent *Pureplayness* index.

2.5.2 Determinants of the *Pureplayness*

To explain the cross-sectional variation of *Pureplayness*, we run a panel regression of *Pureplayness* on a variety of explanatory variables. We use turnover, expense ratio, the number of stocks, and whether a fund is team-managed as explanatory variables, as they are under managers' direct control and endogenous to the dependent variable. The control variables also include fund size, fund age, previous fund returns, and net inflows. Since the dependent variable is persistent within funds, we include time fixed effects but not fund fixed effects. Since funds with different styles might have a different preference for pure-play firms, we also include fund style fixed effects[9]. Because *Pureplayness* and many independent variables are persistent over time, we cluster standard errors by fund and time.

Table 2.7 reports the results. Turnover, expense, and the number of stocks are all positively related to *Pureplayness*, indicating that active managing style partly explains the holding of pure-play firms. Fund size and fund age are positively related to *Pureplayness*, meaning larger funds invest more in pure-play firms. Team manager does not explain *Pureplayness*. It is not surprising because team members are self-selected. Past returns and past inflows have little explanation power of *Pureplayness*. In column 4, we include the lagged Industry Concentration Index (*ICI*) as an independent variable. It is consistent with the hypothesis that *Pureplayness* and *ICI* both capture the fund's industry expertise. We will carefully control for *ICI* in the following fund performance analysis. Overall, the multivariate regressions show that *Pureplayness* is not easily explained by other fund characteristics.

2.5.3 Fund performance and the *Pureplayness*

We have provided evidence that *Pureplayness* potentially captures the fund's industry expertise, and common fund characteristics cannot fully explain it. In this section, we test whether *Pureplayness* contains valuable information and can be used to predict future returns. We then discuss its difference from the Industry Concentration Index constructed by Kacperczyk, Sialm, and Zheng (2005).

We first explore the performance of funds with low and high *Pureplayness* by sorting funds into portfolios based on their past *Pureplayness*. At the beginning of each quarter, we sort funds into quintiles based on *Pureplayness* calculated using the previous quarter's holdings. The low-*Pureplayness* portfolio includes the 20% of mutual funds in the sample with the lowest share invested in pure-play firms in the prior quarter. The high-*Pureplayness* portfolio consists of 20% of mutual funds in the sample with the

[9] We use CRSP mutual fund style and objective codes. The six major fund styles are growth, growth and income, income, mid-cap, small-cap, and micro-cap.

highest percentage invested in pure-play firms in the previous quarter. For each portfolio, we compute the value-weighted average buy-and-hold before-fee returns for the subsequent quarter. The quintile portfolios are rebalanced every quarter. Since portfolios are created using data in the previous quarter, we lose observations of the first three months. The final sample period is from April 2003 to December 2015.

Considering that the difference in returns between the low-*Pureplayness* and high-*Pureplayness* funds could be explained by well-known market anomalies, we adjust the returns and report the Fama-French-Carhart six-factor alpha and factor exposures for the low and high *Pureplayness* portfolios. Table 2.8, Panel A shows that the low-*Pureplayness* portfolio has an annualized alpha of 0.481%, and the high-*Pureplayness* portfolio has an annualized alpha of 2.18%. The difference in alpha between the two portfolios is 1.7% per year and significant at 5% level[10]. High-*Pureplayness* funds tend to hold smaller, growth stocks, stocks with weak operating profitability, and stocks with aggressive investments.

The return predicting power of *Pureplayness* may be explained by the Industry Concentration Index (*ICI*) since these two indices are positively correlated and both measure industry expertise. To rule out this concern, we perform the double-sort test. Each quarter we sort funds first into *ICI* quintiles, and then within each *ICI* quintile, funds are further sorted into *Pureplayness* quintiles. We compute the value-weighted average return within each of the twenty-five portfolios and then compute six-factor alpha of each portfolio using monthly time-series data. Result in Panel B shows that within each of the five quintiles of *ICI*, high-*Pureplayness* portfolios always outperform low-*Pureplayness* portfolios, and the difference is statistically significant in three *ICI* quintiles. *Pureplayness* improves fund performance after controlling for *ICI*.

Another concern is that the predictive power of *Pureplayness* may simply be a consequence of the positive relationship between *Pureplayness* and other fund characteristics. To isolate *Pureplayness* from other fund characteristics, we estimate the following panel regression:

---

[10] The difference of alpha between the high and low *Pureplayness* portfolios are quantitatively similar when we use after-fee mutual fund returns.

$$Performance_{i,t} = \beta_0 + \beta_1 Pureplayness_{i,t-1} + \beta_2 ICI_{i,t-1} + \beta_3 Controls_{i,t-1} + \delta_t + \psi_j + \epsilon_{i,t} \tag{2.9}$$

The independent variables include *Pureplayness*, *ICI*, total net assets in the prior quarter, turnover, expense ratio, fund age at the end of the previous year, and net inflow in the previous twelve months. We include fund style fixed effects and year-quarter fixed effects. Because returns are highly correlated across funds and within funds, we cluster the standard errors by fund and time. We use four different measures of fund performance: six-factor alpha, industry selectivity, industry timing, and industry-adjusted stock selectivity. Industry selectivity measures a fund's ability to select industries that generate positive returns; industry timing measures a fund's ability to generate additional returns by exploiting time-varying returns of industries; industry-adjusted stock selectivity measures a fund's ability to pick stocks that outperform peers within industries (Daniel, et al., 1997; Kacperczyk, Sialm, and Zheng, 2005). They are defined as:

$$Industry\ Selectivity_t = \sum_j [w_{j,t-3}\ IR_t(j, t-3)] \tag{2.10}$$

$$Industry\ Timing_t = \sum_j [w_{j,t-3} IR_t(j, t-3) - w_{j,t-6} IR_t(j, t-6)] \tag{2.11}$$

$$Stock\ Selectivity_t = \sum_j w_{j,t-3} [R_{j,t} - IR_t(j, t-3)] \tag{2.12}$$

where $IR_t(j, t\text{-}3)$ is the value-weighted industry return in month t, to which stock j was allocated at the end of the prior quarter. The variable $w_{j,\,t-3}$ is the share invested in stock j at the end of the prior quarter, $R_{j,\,t}$ is the return of stock j in month t. Since these three measures are adjusted only by industry but not by other common market anomalies, we regress them on the Fama-French-Carhart six-factor model to obtain risk-adjusted industry selectivity, industry timing, and stock selectivity.

Panel regression results are reported in Table 2.9. In column 1, where the dependent variable is six-factor alpha, the coefficient is significantly positive for *Pureplayness* after controlling for *ICI*. The coefficients on both variables are positive and significant in column 2. It suggests that *Pureplayness* and *ICI* are not perfectly correlated, and *Pureplayness* can predict future industry selectivity. Results in columns 3 and 4 show that *Pureplayness* can predict industry timing ability but not stock selectivity. Overall, the results suggest that *Pureplayness* predicts returns by picking and timing industry, and it displays little stock-picking skill. The findings are consistent with our hypothesis that *Pureplayness* captures funds industry expertise[11].

2.6 Conclusion

This paper studies the effect of corporate diversification on mutual fund investments. We find that mutual funds have a preference for pure-play firms and invest 1.029% more in them. What's more, after a pure-play firm merges into a conglomerate, the mutual fund ownership decreases afterward. We also show that firm diversification and the industry expertise of mutual funds explain the preference for pure-play firms. When a firm is more diversified, or the stock price is more affected by the industry factor, or mutual funds have higher industry concentration, mutual fund ownership is higher.

This paper also joins the long-standing debate on whether active mutual fund managers have skills and generate value for investors. Consistent with Kacperczyk, Sialm, and Zheng (2005), we find that mutual fund managers possess industry expertise and exploit the industry expertise by investing more in pure-play firms. Funds with more assets in pure-play firms perform better than other funds. The results are robust after controlling for other industry expertise measures, adjusting risks using the Fama-French-Carhart six-factor model, and dissecting returns into industry selecting ability and industry timing ability.

[11] In untabulated tests, we find that risk-adjusted returns are higher for conglomerates than pure-play firms during the sample period, but the difference is insignificant. Therefore, the better performance of high *Pureplayness* funds cannot be simply explained by higher alpha of pure-play firms.

2.7 Tables

Table 2.1: Summary statistics

The sample is from 2003 to 2015 and includes 107,501 firm-quarter observations, 4,783 unique firms and of which 1,159 are conglomerates. We calculate the mean values of the variables in each quarter and then report the quarterly average across time in Panel A and correlations in Panel B. The last column in Panel A reports the t value for the paired t-tests of the differences between conglomerates and pure-play firms. Mutual fund ownership, *MFO*, is defined as the percentage of shares outstanding that are held by mutual funds. Other variables are defined in Table B.1.

Panel A: Difference in variables between conglomerates and pureplays

	Conglo	PurePlay	Diff	*t*-value
Mutual fund ownership	8.148	7.540	0.609	7.234
Market cap	3089.651	1915.452	1174.199	8.991
Stock return	0.010	0.009	0.001	0.404
Dividend yield	0.025	0.016	0.009	13.445
Return volatility	0.119	0.145	-0.026	-16.208
M/B	1.364	1.996	-0.631	-18.357
Leverage	0.225	0.171	0.054	16.628
Price	26.407	19.072	7.335	10.528
Firm age	23.422	13.941	9.481	20.192
S&P 500	0.124	0.067	0.057	19.020
ROA	0.010	-0.070	0.080	19.080
Stock turnover	4.682	5.398	-0.716	-9.773
Spread	0.010	0.012	-0.001	-4.228

Table 2.1 (continued)

Panel B: Correlation matrix

	Mutual fund ownership	*Market cap*	*Stock return*	*Dividend yield*	*Return volatility*	*M/B*	*Leverage*	*Price*	*Firm age*	*S&P 500*	*ROA*	*Stock turnover*
Market cap	0.629											
Stock return	0.204	0.086										
Dividend yield	0.174	0.368	-0.024									
Return volatility	-0.340	-0.125	-0.013	0.619								
M/B	0.187	0.570	0.084	0.543	0.533							
Leverage	-0.158	0.111	-0.147	0.864	0.887	0.571						
Price	0.538	0.903	0.002	0.615	0.105	0.708	0.393					
Firm age	0.404	0.603	0.007	0.781	0.585	0.852	0.741	0.775				
S&P 500	0.429	0.671	0.031	0.768	0.562	0.810	0.720	0.791	0.978			
ROA	0.576	0.162	0.225	-0.585	-0.830	-0.345	-0.838	-0.083	-0.350	-0.318		
Stock turnover	0.537	0.605	0.031	0.653	0.409	0.816	0.565	0.777	0.890	0.846	-0.209	
Spread	-0.585	-0.307	-0.176	0.557	0.874	0.299	0.841	-0.045	0.339	0.303	-0.932	0.157

Table 2.2: Effects of organizational form on mutual fund ownership

The sample contains 107,501 firm-quarter observations in the period of 2003 to 2015, of which 86,215 observations are pure-play firms. This table reports regressions of the actively managed mutual fund ownership (*MFO*) on the organizational form (*PurePlay*) and control variables. Mutual fund ownership, *MFO*, is defined as the percentage of shares outstanding that are held by mutual funds. *PurePlay* is the organizational form dummy, which equals one if the firm is a pure-play firm (operates in one industry), and zero otherwise. Other variables are defined in Table B.1. This table reports the coefficients from pooled OLS regressions with year-quarter fixed effects and Fama-French 48 industry fixed effects. The standard errors are adjusted for the clustering of observations at the firm and quarter level. T-stats are given in brackets. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
	MFO	*MFO*	*MFO*	*MFO*	*MFO*
PurePlay	-0.680**	0.654***	0.652***	0.596**	0.615**
	(-2.415)	(2.733)	(2.814)	(2.599)	(2.609)
Ln(Mkt cap)		1.668***	1.057***	1.093***	1.090***
		(13.32)	(8.117)	(9.886)	(9.625)
Stock return		-0.815**	-0.826**	-0.768**	-0.566
		(-2.549)	(-2.258)	(-2.110)	(-1.612)
Dividend yield		-8.840***	-11.53***	-11.74***	-11.96***
		(-6.276)	(-7.206)	(-7.069)	(-6.954)
Return volatility		-6.619***	-2.227**	-4.933***	-5.910***
		(-6.145)	(-2.445)	(-4.236)	(-4.630)
M/B			0.00328	-0.0226	0.000669
			(0.0457)	(-0.313)	(0.00894)
Leverage			-0.406	-0.795*	-0.776*
			(-0.916)	(-1.850)	(-1.743)
Ln(Price)			2.628***	2.550***	2.576***
			(11.68)	(12.09)	(12.22)
Ln(Firm age)			0.100	0.142	0.131
			(0.736)	(1.048)	(0.940)
S&P 500			-5.468***	-5.724***	-5.810***
			(-9.342)	(-11.03)	(-10.97)
ROA			-1.297***	-1.161***	-1.252***
			(-5.391)	(-5.028)	(-5.112)
Stock turnover				0.134***	
				(6.771)	
Spread				26.27***	
				(3.002)	

Table 2.2 (continued)

Lag(Stock turnover)					0.138***
					(6.740)
Lag(Spread)					22.04**
					(2.476)
Observations	107,501	99,464	97,077	96,991	93,322
R-squared	0.001	0.360	0.416	0.423	0.420
Industry FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes

Table 2.3: Falsification tests

This table provides falsification tests of the effect of conglomeration on mutual fund ownership. The dependent variable is the mutual fund ownership (*MFO*), as defined in Table B.1. The key variable is the *Actual* dummy for the actual conglomerate. The sample consists of actual conglomerates and pseudo-conglomerates. To construct pseudo-conglomerates, for each segment of an actual conglomerate, we select a single-segment firm in the segment's industry with the closest value of book assets (sales, operating profit, or imputed market size) in the year when the conglomerate first appears in our sample. The pseudo-conglomerate spans the same portfolio of industries as an actual conglomerate. The variables for pseudo-conglomerates are based on book-value weighted variables of each single-segment firm. All the regressions control for firm fixed effects and year-quarter fixed effects. Standard errors are clustered by firm and quarter.

	(1) *Matched by* *Total Assets*	(2) *Matched by* *Sales*	(3) *Matched by* *Operating Profit*	(4) *Matched by* *Imputed Market Cap*
Actual	-1.029***	-0.942***	-0.801**	-0.648*
	(-3.265)	(-2.997)	(-2.519)	(-1.972)
Controls	Yes	Yes	Yes	Yes
Observations	26,331	27,044	26,477	26,257
R-squared	0.623	0.610	0.603	0.597
Firm FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Table 2.4: Mutual fund ownership, firm diversification, and industry expertise

This table reports the results of regressions of mutual fund ownership on various measures of firm diversification. The sample includes quarterly data from 2003 to 2015. The dependent variable is the actively managed mutual fund ownership. In Panel A, *Ln(#Segments + 1)* is the natural logarithm of one plus the number of segments of a firm. *Herfindahl index* of a firm is calculated using the sales of various segments in the firm. In Panel B, a firm is defined as a conglomerate if it operates in more than one segment with a different 3-digit SIC code, or 2-digit SIC code, or 1-digit SIC code. A *PurePaly* dummy that equals one is assigned to firms that are not conglomerates, and zero for conglomerates. In Panel C, we divide mutual funds into 5 quintiles by their Industry Concentration Index (*ICI*). E.g., *MFO_ICI1* is the mutual fund ownership held by mutual funds that are in the lowest quintile of industry concentration in quarter t. *MFO_ICI5* is the ownership held by mutual funds with the highest industry concentration (expertise). In panel D, we create industry beta and industry R-squared to measure how much a firm is affected by the industry factor. Other control variables are defined and listed in Table 2.1. All standard errors are clustered by firm and quarter.

Panel A: Mutual fund ownership and firm diversification

	(1)	(2)	(3)	(4)
	MFO	*MFO*	*MFO*	*MFO*
		Subsample of		*Subsample of*
	Full sample	*Conglomerates*	*Full sample*	*Conglomerates*
Ln(#Segments + 1)	-1.487***	-3.410***		
	(-3.603)	(-3.661)		
Herfindahl Index			1.215**	0.102
			(2.208)	(0.104)
Controls	Yes	Yes	Yes	Yes
Observations	96,790	19,661	96,790	19,661
R-squared	0.424	0.434	0.423	0.429
Industry FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Table 2.4 (continued)

Panel B: Alternative definition of conglomerates

	(1)	(2)	(3)
	MFO	*MFO*	*MFO*
PurePlay SIC3	0.596**		
	(2.599)		
PurePlay SIC2		0.687***	
		(2.801)	
PurePlay SIC1			1.021***
			(3.541)
Controls	Yes	Yes	Yes
Observations	96,991	96,309	96,386
R-squared	0.423	0.424	0.424
Industry FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Panel C: Industry expertise of mutual funds and firm diversification

	(1)	(2)	(3)	(4)	(5)
	MFO_ICI1	*MFO_ICI2*	*MFO_ICI3*	*MFO_ICI4*	*MFO_ICI5*
PurePlay	0.0101	0.0655	0.124	0.187**	0.214***
	(0.299)	(1.051)	(1.464)	(2.530)	(3.317)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	96,991	96,991	96,991	96,991	96,991
R-squared	0.233	0.213	0.233	0.263	0.208
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Table 2.4 (continued)

Panel D: Industry beta, industry R-squared and MFO

	(1)	(2)	(3)	(4)
	MFO	*MFO*	*MFO*	*MFO*
PurePlay		0.583**		0.639***
		(2.551)		(2.779)
IndBeta	1.071***	1.045***		
	(6.065)	(5.943)		
IndRsq			1.074	0.946
			(0.810)	(0.714)
Controls	Yes	Yes	Yes	Yes
Observations	95,789	95,789	95,789	95,789
R-squared	0.426	0.426	0.423	0.424
Industry FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Table 2.5: Fund level summary statistics

This table summarizes the characteristics of actively managed U.S. equity mutual funds. The sample includes fund-quarter level observations from 2003 to 2015, with 1,928 unique funds. The *Pureplayness* is defined as $Pureplayness = \sum(w_j \cdot Pureplay_j)$, where w_j is the portfolio weight of stock j in the fund, and *Pureplay* is a dummy that equals one for pure-play firms, and zero for conglomerates. Other variables are defined in Table B.1.

	Mean	Median	Min	Max	Std	N
Pureplayness	0.617	0.615	0.345	0.897	0.115	62,062
ICI	0.067	0.057	0.015	0.266	0.043	62,062
Fund turnover	0.763	0.610	0.040	3.170	0.601	62,062
Expense	0.012	0.012	0.005	0.022	0.003	62,062
Fund size	5.783	5.739	2.092	9.756	1.678	62,062
Fund size^2	36.256	32.934	4.376	95.175	19.998	62,062
#Stocks	64.818	49.000	11.000	365.000	57.405	62,062
Fund age	16.402	13.000	1.000	56.000	10.986	62,062
Team	0.694	1.000	0.000	1.000	0.461	62,062
Net inflow (3m)	0.001	-0.017	-0.240	0.514	0.109	62,062
Net inflow (1yr)	0.024	-0.047	-0.606	1.400	0.327	62,062
Fund return (3m)	1.152	1.183	0.334	1.870	0.334	62,062
Fund return (1yr)	1.091	1.106	0.537	1.675	0.238	62,062

Table 2.6: Persistence of the *Pureplayness*

This table reports the time-series means of the average *Pureplayness* for the current quarter and the subsequent three months, six months, and one to three years for each of the quintile portfolios sorted on the previous 8-quarter *Pureplayness*. It also reports the difference between the 1st and 5th quintile. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Time 0	3m	6m	1y	2y	3y
Low *Pureplayness*	0.491	0.499	0.506	0.515	0.528	0.541
2	0.563	0.565	0.568	0.572	0.579	0.588
3	0.613	0.613	0.613	0.612	0.616	0.621
4	0.667	0.665	0.665	0.664	0.660	0.662
High *Pureplayness*	0.754	0.748	0.744	0.738	0.730	0.728
Hi - Lo	0.262***	0.249***	0.239***	0.223***	0.202***	0.187***
(*t-stat*)	(250)	(220)	(200)	(170)	(130)	(110)

Table 2.7: Determinants of the *Pureplayness*

The dependent variable is *Pureplayness* for each fund-quarter observation. *Fund size* is the natural logarithm of end-of-quarter total net assets. *Fund return* is the raw return in the previous 3 months or 1 year. *Fund age* is fund age measured by years. *Team managed* is a dummy that equals one for funds that are managed by two or more managers. Other variables are defined in Table B.1. Time fixed effects and fund style fixed effects are included in all specifications. The t-statistics are based on standard errors clustered by fund and year-quarter. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) *Pureplayness*	(2) *Pureplayness*	(3) *Pureplayness*	(4) *Pureplayness*
Fund size	0.00334***	0.00332***	0.00201*	0.00170
	(2.937)	(2.931)	(1.674)	(1.460)
Fund turnover	0.0195***	0.0196***	0.0190***	0.0209***
	(6.677)	(6.639)	(6.393)	(7.252)
Expense	3.164***	3.175***	3.253***	2.773***
	(5.674)	(5.702)	(5.762)	(5.048)
Fund return (3m)		0.0109	0.00997	0.00495
		(0.417)	(0.387)	(0.197)
Fund return (1yr)		0.0135	0.0194	0.0194
		(0.330)	(0.469)	(0.504)
#Stocks			4.23e-05	0.000103***
			(1.374)	(3.266)
Fund age			0.000440**	0.000461**
			(2.480)	(2.655)
Team managed			-0.000163	0.00111
			(-0.0529)	(0.368)
Net inflow (3m)			-0.000269	-2.11e-05
			(-0.0205)	(-0.00167)
Net inflow (1yr)			-0.00386	-0.00451
			(-0.818)	(-0.984)
ICI				0.359***
				(8.717)
Observations	61,314	61,314	61,314	61,314
R-squared	0.266	0.267	0.269	0.284
Time FE	Yes	Yes	Yes	Yes
Fund Style FE	Yes	Yes	Yes	Yes

Table 2.8: Quintile portfolios - Factor-based performance measures

The sample includes monthly value-weighted portfolio returns from April 2003 to December 2015. The table reports the portfolio alphas and factor loadings using the Fama and French (2015) five-factor model plus the Carhart (1997) momentum factor. In Panel A, actively managed equity funds are sorted into quintiles based on their *Pureplayness* in the previous quarter. In Panel B, funds are sorted into quintiles first based on their Industry Concentration Index (*ICI*), and then within each quintile, funds are sorted based on their *Pureplayness* in the previous quarter. The portfolios are rebalanced quarterly, and the alphas are annualized. The standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Portfolios sorted on Pureplayness

	(1) Low *Pureplayness*	(2) 2	(3) 3	(4) 4	(5) High *Pureplayness*	(6) Hi - Lo
Market	0.989***	1.011***	1.006***	1.009***	0.993***	0.00369
	(76.38)	(66.43)	(84.99)	(76.19)	(61.75)	(0.202)
Size	-0.0224	0.0569***	0.185***	0.270***	0.378***	0.401***
	(-1.243)	(2.797)	(11.62)	(12.56)	(13.38)	(12.23)
Value	0.0337	-0.0112	-0.0308	-0.138***	-0.182***	-0.215***
	(1.281)	(-0.404)	(-1.430)	(-4.806)	(-7.234)	(-6.768)
Momentum	-0.0178*	-0.0227*	0.000536	0.0207	0.00776	0.0255
	(-1.896)	(-1.773)	(0.0527)	(1.420)	(0.441)	(1.643)
Profitability	0.00830	0.0341	0.0182	-0.0638**	-0.146***	-0.155***
	(0.343)	(1.293)	(0.747)	(-2.018)	(-3.809)	(-3.817)
Investment	-0.0339	-0.0749***	-0.140***	-0.147***	-0.245***	-0.211***
	(-1.030)	(-3.120)	(-5.168)	(-4.115)	(-5.657)	(-4.153)
Alpha	0.00481	-0.00386	-0.00291	0.0162***	0.0218***	0.0170**
	(1.043)	(-0.807)	(-0.723)	(2.950)	(3.318)	(2.261)
Observations	153	153	153	153	153	153
R-squared	0.990	0.991	0.992	0.987	0.981	0.710

Table 2.8 (continued)

Panel B: Double-sorted portfolios

	Low *Pureplayness*	2	3	4	High *Pureplayness*	Hi - Lo
Low *ICI*	0.00480	-0.00223	-0.00946**	0.00315	0.0218***	0.0170**
	(1.126)	(-0.489)	(-2.338)	(0.669)	(2.796)	(2.411)
2	-0.00152	-0.00680	0.000697	0.0216***	0.0166*	0.0181*
	(-0.251)	(-1.225)	(0.115)	(3.054)	(1.868)	(1.898)
3	0.00696	0.00490	-0.00111	0.0176**	0.0128	0.00581
	(1.408)	(0.640)	(-0.163)	(2.414)	(1.498)	(0.553)
4	0.00744	-0.00312	-0.00253	0.0134*	0.0312***	0.0238**
	(1.223)	(-0.490)	(-0.380)	(1.861)	(3.787)	(2.416)
High *ICI*	0.00582	-0.00307	-0.00358	0.0228***	0.0199**	0.0140
	(0.662)	(-0.390)	(-0.563)	(2.915)	(2.322)	(1.162)

Table 2.9: *Pureplayness* and alternative risk adjustments

This table reports the coefficients of the panel regression of fund performance measures on fund characteristics. The dependent variable is the Fama-French-Carhart six-factor alpha in column 1, and the industry selectivity measure in column 2, the industry timing measure in column 3, and the industry-adjusted stock selectivity measure in column 4. All independent variables are lagged by one quarter relative to the dependent variable. The definition of variables is available in Table B.1. The standard errors are clustered by fund and time and are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) Six-factor Alpha	(2) Industry Selectivity	(3) Industry Timing	(4) Stock Selectivity
Pureplayness	0.0607**	0.0585***	0.00861*	0.0250
	(2.324)	(3.119)	(1.844)	(0.737)
ICI	-0.00247	0.0666**	0.0115	-0.0276
	(-0.0888)	(2.220)	(1.174)	(-0.505)
Fund turnover	-0.000327	0.00241*	6.79e-05	0.00409
	(-0.0739)	(1.720)	(0.0657)	(1.144)
Expense	-0.256	0.00344	0.00947	0.593
	(-0.595)	(0.0135)	(0.102)	(1.416)
Fund size	-0.00103	-0.000466	0.000199	0.000723
	(-0.930)	(-0.846)	(1.177)	(0.991)
Fund age	0.000227**	4.49e-05	-4.84e-05*	5.60e-05
	(2.183)	(0.886)	(-1.783)	(0.621)
Net inflow (1yr)	0.00353	-0.00175	-6.03e-08	0.00643
	(0.630)	(-0.792)	(-5.96e-05)	(1.527)
Observations	46,155	46,155	46,155	46,155
R-squared	0.093	0.332	0.144	0.071
Time FE	Yes	Yes	Yes	Yes
Fund Style FE	Yes	Yes	Yes	Yes

2.8 Figures

Figure 2.1: Mutual fund ownership and firm diversification around M&As

This figure presents the dynamics in mutual fund ownership in the quarters around the conglomerating M&As. The estimation is from the following specification:

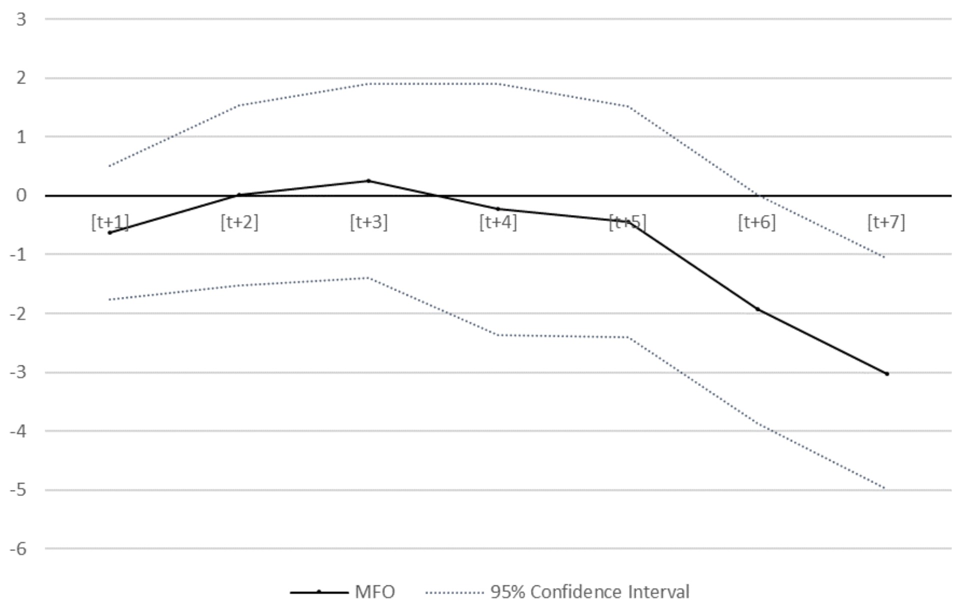
$$MFO_{i,t} = \beta_0 + \beta_1 After_{i,t} + \beta_2 After_{i,t} \times Conglo\ M\&A_{i,t} + \beta_3 Controls_{i,t} + \psi_i + \delta_t + \epsilon_{i,t}$$

The *After* dummy equals one if it is after the M&A effective quarter, and zero otherwise. *Conglo M&A* is a dummy that indicates it is a conglomerating M&A, which requires the target and acquire have different primary 3-digit SIC code, and at least one of them is a pure-play firm before the M&A. In each model, we include one quarter before and one quarter after the M&A. E.g. [t+4] plots the coefficient when we include the fourth quarter before and the fourth quarter after the M&A. We plot β_2 and the 95% confidence intervals in Panel A. We estimate the following specification and plot the coefficients in Panel B and Panel C:

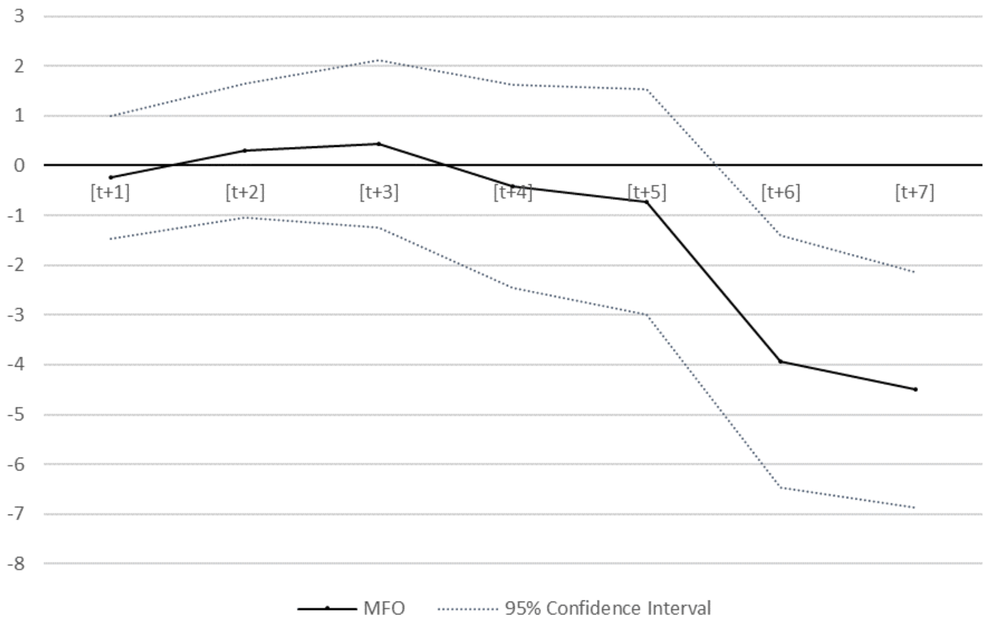
$$MFO_{i,t} = \beta_0 + \beta_1 After_{i,t} + \beta_2 After_{i,t} \times Conglo\ M\&A_{i,t} \times HHI\ drop_i + \beta_3 After_{i,t} \times Conglo\ M\&A_{i,t} \times (1 - HHI\ drop_i) + \beta_4 Controls_{i,t} + \psi_i + \delta_t + \epsilon_{i,t}$$

The *HHI drop* dummy equals one if the change of the Herfindahl index of segment sales around the M&A is below or equal to zero, i.e., the firm becomes more diversified. We plot the β_2 and β_3 coefficients and 95% confidence intervals in Panel B and Panel C, respectively. All the regressions control for firm fixed effects and time fixed effects. T-values are clustered by firm.

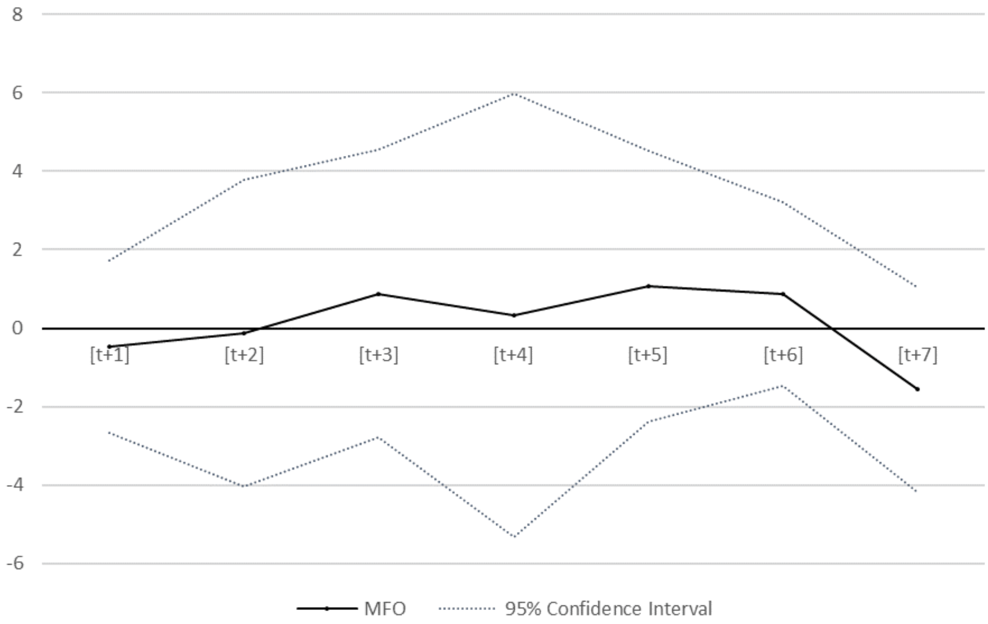
Panel A: Conglomerating M&As



Panel B: Conglomerating M&As – segment sales diversified



Panel C: Conglomerating M&As – segment sales not diversified



Chapter 3 Can Mergers and Acquisitions Internalize Positive Externalities in Funding Innovation?

3.1 Introduction

Technological innovation is a key driver of economic growth. However, several important innovations that, if successful, may lead to significant changes in how an industry or a sector of the economy operates, involve huge upfront costs. At the same time, their benefits may be spread across various sectors of society, such that the value of these benefits is not likely to be fully appropriable by the innovating firm alone. We will refer to such innovations as "fundamental innovations" in this paper. The positive externalities of fundamental innovations may cause individual firms to underinvest in developing such innovations. For example, Netflix, Inc. may work on a more efficient way of streaming high definition videos. If the technology is successfully developed, it can be used not only by Netflix for streaming its own content to its subscribers but may also help other companies in the industry or even outside the industry. For example, users may be likely to use this new technology to share videos on their Facebook pages and generate more website traffic for Facebook. While Netflix can benefit from charging royalties on direct usage of its patents, the spillover effects to Facebook or other companies may be hard for Netflix to appropriate fully. Therefore, the innovation producer (Netflix, in this example) may underinvest in the research and development (R&D) required for developing this more efficient video-streaming technology.[1]

Bayar, Chemmanur, and Liu (2019) theoretically analyze the implication of the above idea and show that venture capitalists can help mitigate the underinvestment in such fundamental innovations by internalizing the positive externalities from innovation by

[1] The positive externalities driven by knowledge spillovers are known to be large in magnitude and are discussed by Griffith, Harrison, and Van Reenen (2006) and Bloom, Schankerman, and Van Reenen (2013). They show that the social rate of return to R&D are more than two times higher than the private return, indicating an under-investment of R&D compared to the socially optimal level.

investing in firms in upstream and downstream industries.[2] In this paper, we empirically analyze a different mechanism to address the above problem, namely, mergers and acquisitions (M&As). We test whether M&As can internalize the positive externalities by merging firms from both the user industries and the producer industries of an innovation. After the merger or acquisition, the combined firm can capture the benefits associated with the innovation in both user industries and producer industries. Therefore, the combined firm should have greater incentives to fund innovation and file more patents.

The notion that innovative activities are difficult to finance in a freely competitive marketplace has been around for a long time, and this is a typical positive externality problem analyzed in economics.[3] Economists have realized that the market can sometimes (at least partly) internalize the positive externality. For example, beekeepers can collect honey from their hives, but the bees will also pollinate surrounding fields and thus aid farmers. If the beehives and fields are owned separately, the number of beehives may be lower than the socially optimal level. However, if the farmer also owns beehives, he or she will increase the number of beehives. Similarly, by funding firms from both the user industries and the producer industries of innovation, the combined firm has incentives to increase the innovation because it will capture a higher proportion of the benefits of the innovation than it would otherwise. However, empirical tests of solutions to this problem are much rarer in the literature. Our paper aims to fill this gap.

---

[2] Another example of a fundamental innovation provided by Bayar, Chemmanur, and Liu (2018) is the development of a new battery that can store electrical energy in a much more efficient way. While an electric car manufacturer such as Tesla may invest significant amounts in developing this innovation (and benefit from it if the battery development is successful), the benefits from the development of this innovation may spill over to not only other electric auto-manufacturers, but to many other sectors of the world economy (such as solar or wind powered electricity generating companies and any other industry that can benefit from the more efficient storage of electric energy). This means that individual firms like Tesla may currently be underinvesting in developing this fundamental innovation.

[3] The discussion perhaps starts with the classic articles of Nelson (1959) and Arrow (1962). Even in the presence of various mechanisms to increase appropriability such as patents and other forms of intellectual property protection, the underinvestment may not completely go away. For example, Mazzucato (2015) points out that fundamental innovations underlying such popular products as the iPod and iPhone manufactured by Apple, and Google's search algorithm, were funded by U.S. government in various ways (e.g., through agencies such as DARPA or the National Science Foundation).

Empirically, we test this hypothesis by using two datasets: M&As from the Thomson Reuters Securities Data Company (SDC) Platinum, which contains data on how firms merge, and a US patent and citation dataset, which contains data on firms that file the patents and citations received by the patents. We start by defining the upstream (innovation-producing) and downstream (innovation-using) firms for each industry using the patent citation dataset. If patents filed by firms from one industry are most cited by another industry in the previous 10 years, we define the former industry as the upstream industry and latter industry as the downstream industry. Firms in the upstream industry are the innovation producers, while firms in the downstream industry are the innovation users. We then show that after a merger between firms in upstream and downstream industries (i.e., both producers and users of an innovation), the combined firm becomes more innovative compared to the case where the two firms remain separate. We use various measures of innovation that are used in the literature.

Bena and Li (2014) ask a related question of how M&As affect innovation and find that if two firms share the same innovation knowledge base, the combined firm after the merger is more productive compared to the case where the two firms remain separate. One drawback of their result comes from the firm-level data. The acquiring firm and the target firm may both file patents in multiply technological classes in the year before the merger, but not all the technological classes of the acquirer are affected by this merger. The synergy of innovation-related mergers and the increase of innovation output should mainly come from the targeted technological classes that are in a user-producer relationship with the target firm's R&D. Our paper, therefore, takes a different approach to this problem, as we discuss below.

To better pin down the effect of mergers on patents, we further construct a firm tech-class level dataset. Instead of providing associated SIC codes in patent documents, the United States Patent and Trademark Office (USPTO) assigns patents to three-digit technological classes that are based on technology categorization instead of final product categorization. Because targets in about 90% of the M&A deals are private firms, their patent and technological class data are unavailable in our dataset. Therefore, in this setting, we know the technological classes of the acquirers and the industries of the targets. Similar to the industry-to-industry relationship, we define an industry-to-technological-class relationship using the patent and citation dataset. If patents filed by firms from one industry are most cited by (most likely to cite) patents from one technological class in the previous 10 years, we define that the industry and the class have a producer-user (user-producer) relationship. We can then compare the affected technological classes to other unaffected classes to better pin down the effect of mergers between innovation producers and users. Moreover, a technological class can be related to the target firm in one merger and unrelated in another merger at the same time, so we can then capture the effect of related M&A within the same technological class. We show that a tech-class been targeted in related mergers becomes more innovative compared to the same tech-class in unrelated mergers.

If M&As between innovation users and producers do internalize positive externalities and incentivize innovation, firms should fund more innovation in targeted technological classes and less in other classes after mergers. We find that this is indeed the case: we observe an increase in innovation output in the targeted classes and a decline in other classes. Moreover, financially unconstrained firms tend to reallocate resources to targeted classes from other unaffected classes.

Finally, we test the impact of innovation-related M&As on tech-class level innovation. If a technological class is more likely to be involved in related mergers, the innovation should be enhanced in that technological class. Our results support this hypothesis.

The first identification challenge comes from the concern that the increase may be mechanical. It may be the case that innovation increases after mergers not because of the synergy or internalizing positive externality, but simply due to the fact that target firm files patents anyway. For example, Sevilier and Tian (2012) show that firms undertake M&As for the purpose of acquiring innovation. To address this concern, first, we use unrelated mergers as the control group in the firm-level regressions so that the mechanical increase cancels out with each other. Using a control group may not entirely eliminate the concern of mechanical increase because one can argue that target firms in related mergers tend to be more innovative. We address this concern in the firm-class level regressions. Targets and acquirers in related mergers do not necessarily file patents in the same technological class. In fact, among the M&A deals with public target firms, the technological classes of acquirers and targets do not overlap in most of the deals. Moreover, even when acquirers and targets have overlapping technological classes, they are not necessarily defined as targeted classes but are included in the control group. Therefore, the increase in patents in targeted technological classes is not caused by mechanical reasons.

The second concern is that the increase in innovation and M&A activity can be endogenous. For example, a firm with a large amount of free cash may invest more in innovation and M&A activities at the same time. We address this concern by including M&A deal fixed effects and technological class fixed effects in the regressions. To further establish a causal effect between M&A activity and innovation, we use the method developed by Savor and Lu (2009) to compare the change of innovation around successful mergers to that of mergers withdrawn for reasons that are exogenous to innovation. For example, we exclude mergers that are withdrawn due to the disagreement on the future development strategy between the acquirer and the target because this reason may be related to future innovation strategy. By ruling out the systematic relation between a firm's innovation and the probability of a failed merger, this strategy can help identify the causal effect of a firm's M&A on its innovation output. A firm that decides to invest more in innovation can choose to acquire another firm with relevant knowledge to achieve this goal. However, the innovation output does not increase when the merger is withdrawn. Our result is consistent with existing literature and shows that the innovation output increases more after successful user-producer mergers relative to failed user-producer mergers.

Another identification challenge comes from shocks at the industry level. For instance, suppose an industry is growing fast, and its product market is becoming more and more competitive, firms will merge to exploit synergies to differentiate their products from their competitors. By the same token, the innovation of the industry may also reach the peak given the inverted-U shape relationship between competition and innovation (Aghion, et al. 2005). This industry, with more patents, is more likely to become the top producer or user of innovation of other industries. Therefore, industry-level shocks may be driving both innovation-related M&As and innovation outputs. Fortunately, the concern of such shocks can be mitigated by using firm tech-class level data because the tech-class level relationship is less correlated with the acquirers' industry condition. In addition, we control for any shocks to technological classes by including a full set of class-year fixed effects. The fixed effects are identified because a technological class can be involved in a related merger and an unrelated merger at the same time.

Our paper contributes to the M&A literature and the innovation literature in at least two ways. We are the first to develop a measure of innovation-user and innovation-producer industries and a user-producer relationship between industries and technological classes, which can be used in future research. And we show that M&As between an innovation-user and an innovation-producer can internalize the positive externalities associated with funding innovation, and the increase of innovation is driven by the targeted technological classes.

The remainder of the chapter is organized as follows. We discuss the related literature in section 3.2. Section 3.3 describes the sample construction, empirical methodology, and sample overview. Section 3.4 reports the empirical results on changes in innovations around M&As. Section 3.5 concludes the paper.

3.2 Literature Review

Our study is related both to studies in firm innovation and studies in M&As. Innovation is an important driver of firm performance (Bloom and Van Reenen 2002; Levine 2005), and the quality of patents correlates with a firm's market value (Hall, Jaffe, and Trajtenberg 2005). Kogan et al. (2017) also find that the stock market positively reacts to the approval of patents that are eventually highly cited, and that such patents predict firm productivity. Given the importance of firm innovation, it is essential to understand the factors that incentivize it. Some papers empirically show that CEOs' incentives have a significant impact on motivating innovation. For example, both corporate venture capital (Chemmanur, Loutskina, and Tian 2014) and stock options in CEOs' compensation (Chang et al. 2015) motivate managers to undertake innovative projects.

Despite the importance and incentive of innovation, funding it can be difficult. Brown, Fazzari, and Petersen (2009) is the first paper to show that cash flow and external equity are essential to financing innovations in young firms. Large and publicly traded firms also depend on banks, and they receive cheaper bank loans if they produce higher-quality patents (Hall and Lerner 2010). It indicates that firms and lenders consider the cost of innovation and the value associated with the innovation. If the benefit of innovation spread across industries and is not easily appropriable, this positive externality will cause firms to underinvest in innovation.

Several papers study the effect of M&As on innovation. Seru (2014) shows that innovation decreases after diversifying M&As because inventors become less productive. Sevilir and Tian (2012) find a positive relationship between M&As and innovation and show that acquiring innovation is an important motive for undertaking M&As. Our paper takes a different approach. We consider the citation relationship between acquirers and target firms before the M&As and study the synergy from this relationship. Phillips and Zhdanov (2012) find that acquiring firms that have successfully innovated can be a more efficient path to obtaining innovation than innovating directly by oneself. We take care of the mechanical increase of innovation after M&As by using unrelated M&As as a control group and using firm-class level data. The paper that is most relevant to ours is Bena and Li (2014). They show that M&As are more likely to be conducted between firms with technological overlap, and the innovation output increases after such mergers. Their identification strategy is to compare successful mergers to withdrawn mergers. Our paper is different in the sense that our main results come from all available mergers that include about 2,804 deals, compared to the 60 withdrawn mergers in their analysis. Additionally, we use firm-technological-class level data to better pin down the effect. We will elaborate on this idea in the empirical part of the paper.

This paper is the first to study the citation links between industries and technological classes using a novel patent citation dataset. We show that merging two firms from the user industry and the producer industry internalize the positive externalities of innovations and enhances the innovation output afterward. We are also the first to study the effect of M&As on firm tech-class level innovation, which allows us to mitigate the endogeneity concern.

3.3 Data

3.3.1 Sample selection

We use a novel dataset of patents and citations constructed by Kogan et al. (2017). The dataset includes the entire history of US patent documents from Google Patents. The United States Patent and Trademark Office (USPTO) allows only individuals to be the inventor, but an individual can assign granted patent to another person or to a corporation. Therefore, patents always have an inventor, and sometimes they have been assigned to one or more corporations. Kogan et al. then matched the corporation names to firms in the Center for Research in Security Prices (CRSP) stock return database. The dataset covers patents granted from 1926 to 2010 that are assigned to firms in the CRSP database. The USPTO also keeps track of all citations for patents granted from 1976 to 2010. Compared to the NBER patent project, this dataset provides 1.9 million patents that can be matched to companies, 27 percent of which are not included in the NBER data.[4] Another commonly used measure of innovation is R&D expenditures, but 65 percent of firm-year observations from Compustat have missing values. Missing R&D expenditures in financial statements do not necessarily mean that the firm is not innovative (Koh and Reeb 2015). Therefore, compared to R&D expenditures, patent-based metrics better reflect the productivity of R&D and more realistically reflect a firm's innovation performance. Using this patent and citation dataset allows us to measure the innovation output for every public firm in every year.

In addition, the USPTO has developed an elaborate classification system for the patented inventions. This system categories technologies into about 400 3-digit patent

[4] The dataset is provided by Noah Stoffman on his website (https://iu.app.box.com/v/patents). This dataset covers more patents and corresponding firms than NBER patents data mainly because the patent text files provided by Google have better quality than the files provided by USPTO, so more patent assignees can be identified. More details of the patent data construction can be found in the paper and the online appendix of Kogan et al. (2017).

classes, and each patent belongs to a technological class. Using the classification data, we can further measure the innovation output of each firm in each class and analyze the effect of M&A on the firm tech-class level.

To identify a sample of M&As, we begin with all completed US M&As with effective dates from 1984 to 2007, covered by the mergers and acquisitions database of Thomson Reuters' SDC Database[5]. We exclude the deal if: (1) the acquirer or the target firm is from the financial industry (Standard Industrial Classification [SIC] codes 6000 to 6999); (2) the transaction value of the deal is less than $10 million, to drop the small and economically insignificant deals; (3) the acquiring firm cannot be matched to Compustat/CRSP; (4) the acquiring firm did not file any patent the year before the merger; (5) the acquiring firm exists less than three years before or after the M&A because there is a few years' lag between starting an innovative project and filing patents. To make sure the acquirers are innovative before mergers, we require them to have at least one patent the year before the mergers. We do not require the target firms to be matched to Compustat/CRSP because most of the target firms are private firms; excluding them would result in a large drop in the sample size. Large firms tend to buy innovation by acquiring small private firms that are engaged in R&D, so excluding the sample of private target firms would lead to biased estimation. Our final M&A sample contains 2,804 deals for the period 1984 to 2007.

3.3.2 Related M&As and targeted tech-classes

To examine how mergers between innovation-user and innovation-producer can enhance innovation output, we need to define the user-producer relationship between industries. In year *t*, industry *i* and *j* are innovation related industries if in the previous 10 years[6], the patents filed by firms in industry *i* is among the top-3 industries that cite (are cited by) patents filed by firms in industry *j*. Similarly, we define the relationship

[5] Our sample begins in 1984 because information on M&As in SDC is less reliable before 1984. Our sample period ends on December 31, 2007, three years before our patent data end in 2010. Allowing a three-year period after the last merger can mitigate the potential truncation bias in our innovation output measures.

[6] Using previous 3 years does not change our results.

between industries and technological classes. In year *t*, industry *i* and technological class *j* are related if, in the previous 10 years, the patents filed by firms in technological class *j* is among the top-3 classes that cite (are cited by) patents filed in industry *i*. In this definition, the user-producer relationship updates every year and can capture the changing relationship over time. In our final sample, 1,567 deals are between firms from related industries (related M&As), and 12% of the technological classes of acquirers (targeted classes) are related to the target firms' industry.

3.3.3 Variables

We employ various measures of innovation to capture different aspects of a firm's innovation performance. We begin with the number of patents filed each year and the number of citations received by those patents. We count patents at the time when they are filed with the USPTO because inventors have the incentives to file the patent as soon as it is finished, the filing date is the closest to the actual time of innovation. The citations are counted after the grant date when a patent is revealed and starts to be cited by the public. Because the distribution is positively skewed, we use the natural logarithm of these innovation measures.

Some technological classes have more patents and receive more citations than other classes by nature. To adjust for this heterogeneity, we also create a citation-weighted number of patents and a patent index. The citation-weighted number of patents scales the patent citation by the average number of citations a patent received in the same year and the same technology class (Hall, Jaffe, and Trajtenberg 2001). We also follow Bena and Li (2014) to calculate the patent index of each firm. The patent index is the patent number adjusted by the median value of each technology class.[7] However, the measures of patents and citations are meaningful only when used comparatively, the fact that a firm files 10 or 100 patents does not tell you whether the firm is highly innovative. That is, the evaluation of the patent intensity needs to be made with references to some "benchmark" intensity. Therefore, we control for deal fixed effects

[7] Details of how to construct these measures are reported in the Appendix C.

or tech-class fixed effects in the regressions.

In the analyses, we use the measures of innovation output as the dependent variables. Our key test variables are the *After* dummy, which equals one if the observation is after M&As and zero otherwise, and the *Related* dummy, which equals one if the deal is a related M&A and zero otherwise. The *After* dummy captures the change of innovation around M&As, and the *Related* dummy captures the difference between related M&As and unrelated M&As. In the multivariate tests, we include control variables that may affect firm innovation. *R&D/Assets* is the R&D expense adjusted by total assets. *Sales* are the total sales as a measure of firm size. *ROA* is the return on assets, defined as the operating income before depreciation, divided by total assets. *Leverage* is the total debt divided by total assets. Capital expenditure, *Capex*, is the capital expenditure divided by total assets. *Tangibility* is the total gross property, plant, and equipment divided by total assets. *Tobin's Q* is the ratio of market value to book value of assets. *HHI* is the Herfindahl index that captures the competition of a 3-digit SIC industry.

Table 3.1 reports the summary statistics of firm innovations and control variables. Panel A shows a summary of all acquiring firms during the sample period. Panels B, C, and D report the M&A characteristics. From 1984 to 2007, there are 2,804 deals of M&As, of which 1,583 deals are innovation-related M&As. Panel E shows the summary statistics of firm-class level innovation. An average firm file patents in 3.8 technological classes in the year before mergers, the median number is 3. An average firm file patents in 12.3 technological classes during the entire sample period, the median number is 10.

3.4 Empirical Results

3.4.1 Univariate tests

Because firm innovation and the decision to merge are endogenously determined, it is difficult to compare innovation between firms involved in related M&As and unrelated M&As. Therefore, we use a panel structure approach to control for many factors that affect both innovation and M&A decisions, such as R&D, ROA, and leverage. Before doing so, we provide some simple summary evidence on differences in firm innovation between related and unrelated M&As.

We first conduct a before-after comparison among firms involved in related M&As and unrelated M&As, respectively. In Table 3.2, Panel A reports the change of firm innovation and characteristics around related M&As. The variables are the average value of three years' observations before or after M&As. The number of patents, citation-weighted number of patents, and patent index all increase significantly after related M&As. The originality of patents also increases. In Panel B, we conduct the before-after test among firms involved in unrelated M&As, and the change of generality and patent index is insignificant. Panel C reports comparisons of the change after related M&As and unrelated M&As. The results show that the number of patents, the citation-weighted patents, the patent index, originality, and generality all increase by a higher amount after related M&As compared to unrelated M&As.

3.4.2 Multivariate tests of firm-level innovation

We use a panel structure to analyze the innovation around related M&As. Because the increase of innovation output can be mechanical, we use unrelated M&As as the control group and perform the difference-in-difference tests in a multivariate setting. Specifically, we estimate the following regression using a sample of all related and unrelated M&As:

$$\begin{aligned} Innovation_{i,t} = \beta_0 + \beta_1 After_{i,t} + \beta_2 After_{i,t} \times Related\ M\&A_{i,t} \\ + \beta_3 Controls_{i,t} + \psi_i + \delta_t + \epsilon_{i,t} \end{aligned} \quad (3.1)$$

where $Innovation_{i,t}$ is one of the innovation measures of firm i in year t; $After_{i,t}$ is a dummy that equals one if it is after the M&As and zero otherwise; $Related_{i,t}$ equals one for mergers between innovation users and innovation producers (related M&As) and equals zero otherwise. We keep 6 years of data before and after each M&A because it usually takes a few years for the acquired knowledge to turn into patents. We include M&A deal fixed effects to difference away any time-invariant differences among deals[8], and we use year fixed effects to control for common trends in all M&A deals. Therefore,

[8] The coefficient on $Related_{i,t}$ is absorbed by deal fixed effects.

this model estimates the change over time in innovation output for the same cross-section units. The variable of interest is the interaction term of *After* dummy and *Related* dummy that captures the change of innovation after the acquiring firm merges with a target firm that is in the upstream or downstream industries of patent citations. The hypothesis that related M&As can internalize the positive externalities predicts that the coefficient β_2 is positive, meaning that the combined firm is more productive after related mergers relative to unrelated mergers.

We include control variables commonly used in the previous literature. Bena and Li (2014) find that firms that are larger, have higher ROA, and have higher market-to-book ratios are more likely to be the acquirers. We, therefore, include the variable Tobin's Q to control for investment opportunities; Sales to control for size; and ROA to control for profitability. Other factors include leverage, capital expenditures, tangibility, and R&D expenses. The results are reported in Table 3.3. In all regressions throughout the paper, we control for M&A deal fixed effects and year fixed effects, and the p-values reported are based on deal-level clustered standard errors.

In the first column of Table 3.3, the dependent variable is the total number of patents filed by a firm, and the coefficient on the interaction term is positive and significant at the 1 percent level. The result indicates that the increase of innovation after related M&As is greater than that after unrelated M&As. In columns 3 to 6, we use different measures of firm innovation, and the coefficients are all significantly positive. In column 2, the coefficients on the interaction term are insignificant but still show a positive sign. In summary, the difference-in-difference tests in Table 3.3 show that related M&As enhance the quantity and quality of innovation output more than unrelated M&As do. The results indicate that after the merger of two firms from related industries, the combined firm is more likely to benefit from the innovation, and thus innovation output is enhanced to a new, optimal level.

The coefficient on *After* (β_1) captures the effect of unrelated M&As on innovation output, and they are all significantly negative as shown in Table 3.3. This finding is consistent with Seru (2014), which shows that innovation ability decreases after conglomerating mergers. Because the combined firm may be reluctant to fund innovative ideas, inventors may leave the firm around diversifying mergers while stayers become less productive (Gompers, Lerner and Scharfstein, 2005). Other literature, including Jensen and Ruback (1983) and Cassiman et al. (2005), points out that M&A activities increase financial leverage and the opportunity cost of funding R&D, which leads to the elimination of R&D projects, especially for debt-financed M&As. Our results show that the decrease in innovation output is mitigated when the mergers are between an innovation user and an innovation producer.

3.4.3 Firm tech-class level test

The synergy between innovation user and innovation producer makes it easier for firms to appropriate the value of innovation, obtain relevant knowledge, internalize costs, more likely to fund innovation, and thus enhance innovation output after mergers. However, this synergy and enhancement should mainly occur in the targeted technological classes rather than every area of research within the firm. To examine the effect of M&As on the targeted technological class and other classes within the same firm, we estimate the following firm tech-class level difference-in-difference specification:

$$Innovation_{i,j,t} = \beta_0 + \beta_1 After_{i,t} + \beta_2 Targrted\ Class_{i,j,t} + \beta_3 After_{i,t} \times Targeted\ Class_{i,t} + \beta_3 Controls_{i,t} + \psi_i + \delta_{j,t} + \epsilon_{i,j,t} \quad (3.2)$$

where i indexes firms, j indexes technological classes, t indexes years. $After_{i,t}$ is a dummy that equals one if it is after the M&As; $Targeted\ Class_{i,j,t}$ equals one if the technological class j of the acquirer i is in a user-producer relationship with the target's industry. The coefficient of interest is β_3, which measures the effect of M&As on targeted classes relative to other classes. We include firm fixed effects and class by year fixed effects in the regressions. The class by year fixed effects control for any technological class level shocks.

The dependent variables are various measures of firm tech-class level innovation output, include the number of patents filed by firm i in year t that belong to technological class j; the number of citations received by these patents in the subsequent 3 years after granted; the citation-adjusted number of patents; and the patent index. The control variables are the same as those in equation (3.1). To mitigate the effect of outliers, we winsorize all variables at the 1st and 99th percentiles. To account for serial and cross-sectional dependence across classes within the same firm, we cluster standard errors at the deal level.

Table 3.4, Panel A shows the effect of M&As on innovation output of the targeted technological classes and other classes within the same firm. As shown in column 1, after mergers, the number of patents in the targeted class increases by 4.7% relative to the other tech-classes in the same firm. The coefficients on the interaction term are positive in columns 3 and 4, and they are both significant, meaning that both the citation-adjusted patents and the patent index increase significantly in the targeted class compared to the other tech-classes.

The interaction term in the above difference-in-difference model shows the difference between targeted classes and other classes, but does not tell the exact changes in each of the two categories. In four out of the six columns in Table 3.4, the coefficients on *After* (β_1) are significantly negative, meaning that the innovation output decreases in other technological classes. The sum of coefficients on *After* (β_1) and the interaction term (β_3) captures the effect of M&As on targeted technological classes, and they are all positive, meaning that the innovation output increases in targeted technological classes. These results support the hypothesis that firms reallocate resources from the other classes to the targeted class.

3.4.4 Reallocation of innovation within firms

Reallocation of resources within firms rests on the premise that firms are financially constrained so that it is necessary to winner-pick the best performing sectors (Stein 1997). Giroud and Mueller (2015) find that financially constrained firms reallocate capital and labor to plants experiencing positive shocks, but unconstrained firms use external funding and do not reallocate within the firm. Similarly, we expect to see the reallocation of innovation funding in financially constrained firms. To test this hypothesis, we examine the effect separately for financially constrained and unconstrained firms by estimating equation (3.2) in two subsamples. We use the KZ index (Kaplan and Zingales, 1997) to measure financial constraints. A firm is financially unconstrained (constrained) if the firm's KZ index is below (above) the median of all sample firms in the year before M&As[9]. Table 3.5, Panel A reports the results of constrained firms. Panel B reports the results using unconstrained firms.

In all six columns of Panel A, the coefficients on the *After* dummy are significantly negative, meaning that the quality and quantity of innovation output decrease in unaffected tech-classes. The sum of coefficients on *After* and the interaction term estimate the effect of M&As on targeted tech-classes, and the values are close to zero in column 1, 3, and 4, meaning that the quantity of innovation output does not change in targeted classes after M&As in financially constrained firms. However, in columns 5 and 6, the originality and generality increase significantly in targeted classes. The results indicate that although financially constrained firms have trouble in funding innovation after related M&As, the mergers between innovation users and innovation producers can enhance the quality of innovation output. In Panel B, the coefficients on *After* and the interaction term show that M&As enhance the innovation output in targeted classes in financially unconstrained firms, and the innovation output remains the same in the other unaffected classes. The difference of results in Panel A and B indicates that financially constrained firms reallocate resources from other classes to fund targeted classes, while unconstrained firms can fund more innovation in targeted classes and maintain the innovation level in other classes.

Another premise of reallocation is that firms conduct innovative research in multiple technological classes. A firm operating in many classes may be easier to move the resource to one important class. Therefore, we estimate the effect of mergers on innovation with respect to the number of technological classes. A dummy *Few Classes* equals one if, in the year before mergers, the number of technological classes of a firm

---

[9] The financial constrained dummy is generated conditioning on firms that conducted M&As. An acquirer can be defined as financially constrained compared to other acquirers but financially more flexible than an average firm of the whole population.

is below the median value of all firms and equals zero otherwise. Table 3.6 reports the results. The positive coefficient on the triple interaction term means that the innovation output of targeted classes increases more in firms with fewer technological classes. This result contradicts our reallocation hypothesis. Instead, it indicates that firms that concentrate innovation in fewer classes can conduct research in the targeted class where the patent value is more appropriable.

3.4.5 The aggregate effect of M&As

As shown in the previous tables, M&As between innovation-user and innovation-producer can enhance the innovation output of the targeted technological class as well as the combined firm. In this part, we consider the aggregate effect at the technological class level. To be more specific, a technological class involved in related M&As will experience an increase in innovation output. We then construct a class-year panel dataset. The dummy *Targeted Class* equals 1 if that class is involved in at least one innovation-related mergers and 0 otherwise. We control for class fixed effects in the regressions. The results are reported in Table 3.7. The positive coefficient means that after mergers, the innovation output quantity increases in technological classes that are involved in innovation-related mergers.

3.4.6 Withdrawn M&As

We also employ the identification strategy developed by Savor and Lu (2009). They compare the change of innovation around successful mergers to that of mergers withdrawn for reasons that are exogenous to innovation. We followed Bena and Li (2014) to construct the control group using withdrawn M&As. We begin with 191 unsuccessful friendly merger bids that are announced from 1984 to 2007. We then keep deals where the news articles from Factiva did not mention R&D activity as a reason for the failure. Table 3.8, Panel A presents the filters. The final control group includes 67 unsuccessful merger bids.

Next, we construct the treatment sample of completed deals that: (i) involve acquirers and target firms of which data are available in Compustat/CRSP; (ii) occur in related (unrelated) acquirer-target industry pairs that match related (unrelated) industry pairs of the failed mergers; and (iii) are announced within the three-year window centered at the announcement year of the failed mergers (514 deals). For the failed mergers that are matched to multiple deals, we select the completed deal with the closest relative size ratio, measured by the target firm's total assets divided by the acquirer's total assets.

To estimate the different effects from successful mergers and failed mergers, we create a dummy *Treated* that equals one for successful M&As; *Related* equals one for M&As of two firms from two related industries. Table 3.8, Panel B shows that after a successful M&As, the patents increase when the two firms are from related industries. Panel C reports the falsification test results in which we falsely assume that the onset of treatment occurs four years before it actually does. Results in Panel C show that the coefficients on the triple interaction term are insignificant as expected.

3.4.7 Robustness tests

This study focuses on the relationship between innovation user and innovation producer instead of shared knowledge or technology proximity. Although our measure of innovation-related M&As is different from M&As between firms doing similar innovative research, these two measures may be overlapped. An industry can be its own innovation user or producer, so our results captures the synergy between firms in the same industry instead of the innovation-related industries. To mitigate this concern, we exclude M&As between firms from the same industry and estimate equation (3.1) again. The results in Table 3.9 shows that our results still hold.

Next, we conduct the analysis using a higher-level classification of patents instead of technological classes. Because there are over 400 tech-classes, two different tech-classes can be very similar, and firms can switch between similar tech-classes effortlessly. Therefore, the narrow definition of the technological class may bias our estimations. To alleviate this concern, we re-estimate the effect of related M&As on innovations at the firm sub-category level. Hall, Jaffe, and Trajtenberg (2001) aggregate the 400 tech-classes into 36 two-digit technological sub-categories. We can then measure the innovation output of each firm in each sub-category. The results in Table 3.10 show a similar positive effect of M&A on targeted sub-categories.

Finally, we use different measures of related industries as robustness tests. In the main tests, we measure the citation relationship between industries using the previous 10 years of data and define the top 3 pairs as related. In Table 3.11, we use the previous 10 years or 3 years of data and define the top 1 pair, 3 pairs, and 5 pairs as related industries. For simplicity, we only report the results using the *Patent Index* as dependent variables. Using other measures of innovation output give qualitatively similar results to the main tables.

3.5 Conclusion

Fundamental innovations usually involve huge upfront costs, but the benefits are spread across various sectors and are difficult for the inventors to appropriate. That means innovations possess positive externalities and are underinvested in by individual firms. We empirically show that after M&As between firms from user industries and producer industries of innovations, the combined firm can internalize the positive externalities and enhance innovation output. We use failed M&As as an identification strategy, and the results are robust.

Using a firm tech-class level dataset, we find that the increase in firm-level patents is driven by the patents in the technological class that is in a user-producer relationship with the target firm. While financially unconstrained acquirers fund more innovation in the targeted tech-classes, constrained firms move resources from other unaffected tech-classes to the targeted classes. The firm-class panel data also help to mitigate the concerns of mechanical increase and endogeneity.

3.6 Tables

Table 3.1: Summary statistics

Panel A reports the summary statistics of key variables from 1984 to 2007, including 1,318 acquiring firms and 2,804 M&A deals. If an acquiring firm conducts multiple M&As within three years, we exclude all those deals. Based on our definition of industry-level innovation relationship, 1,567 M&As are between an innovation-user firm and an innovation-producer firm. The innovation measures are defined in Table B.1. Panel B, C, and D report the M&A summary statistics. The sample includes acquiring firms that filed at least one patent one year before the mergers. Panel E reports the firm technological-class level summary statistics.

Panel A: Firm summary

	Mean	Median	St. dev.
Innovation Measures			
#Patents	30.400	3.688	107.171
#Citations	54.932	4.000	221.505
#Citations per patent	1.283	0.800	1.763
#Patent - Citationweighted	61.403	6.157	220.182
NumPat	1.664	1.251	1.428
NumCited	1.495	1.011	1.498
CiteWeightPat	1.975	1.576	1.628
Patent Index	17.746	2.328	59.269
Originality	0.479	0.500	0.287
Generality	0.459	0.481	0.287
Control Variables			
Assets Total	4775.677	671.175	17351.520
ROA	0.084	0.123	0.169
Leverage	0.210	0.195	0.156
R&D/Assets	0.092	0.059	0.106
Capex/Assets	0.057	0.050	0.034
TobinQ	2.376	1.834	1.604
Tangibility	0.253	0.205	0.178
HHI	0.045	0.028	0.058

Table 3.1 (continued)

Panel B: Distribution by year of execution

Period	1984–1987	1988–1991	1992–1995	1996–1999	2000–2003	2004–2007	Total
#M&As	296	307	371	690	725	415	2,804
#related M&As	110	135	192	413	455	262	1,567

Panel C: Distribution of top five industries (4-digit SIC) for acquirers

Industry (SIC)	Prepackaged Software (7372)	Semiconductors and Related Devices (3674)	Pharmaceutical Preparations (2834)	Biological Products (2836)	Surgical and Medical Instruments (3841)
#M&As	215	182	153	80	75
#related M&As	184	155	139	69	56

Panel D: Distribution of top five industries (4-digit SIC) for targets

Industry (SIC)	Prepackaged Software (7372)	Semiconductors and Related Devices (3674)	Pharmaceutical Preparations (2834)	Surgical and Medical Instruments (3841)	Computer Peripheral Equipment (3577)
#M&As	228	138	110	83	56
#related M&As	179	128	106	68	52

Panel E: Firm tech-class summary

	All Classes		Related Classes		Other Classes	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
#Patents	6.643	20.445	10.279	29.128	5.530	16.768
#Citations	15.525	72.853	25.094	107.390	12.597	58.035
#Patent - Citationweighted	13.855	46.153	21.437	65.986	11.535	37.780
NumPat	0.449	0.820	0.812	1.090	0.383	0.742
NumCited	0.404	0.994	0.766	1.359	0.338	0.897
CiteWeightPat	0.577	1.039	1.026	1.348	0.496	0.951
Patent Index	1.262	5.847	2.685	9.063	1.004	5.008
Originality	0.143	0.272	0.251	0.321	0.123	0.258
Generality	0.137	0.267	0.237	0.317	0.119	0.253

Table 3.2: Univariate tests of innovation before and after M&As

This table provides univariate test results of the innovation and firm characteristics before and after M&As. Panel A includes all related M&As, and Panel B includes all unrelated M&As. An M&A is defined as "related" if the target and the acquirer are from two related industries; two industries are defined as "related" if patents filed by firms in one industry is among the top three industries citing or been cited by patents filed by firms in another industry in the previous 10 years. Panel C reports the results from a difference-in-difference univariate test by comparing the change of variables around related M&As to unrelated M&As. The variables are the mean value of three years before or after M&As.

Panel A: Related M&As

	After	Before	Diff	*p*-value	*N*
#Patents	62.499	52.403	10.097	0.002	1,567
#Citations	116.555	155.361	-38.807	0.003	1,567
#Patent - Citationweighted	130.816	116.135	14.681	0.041	1,567
NumPat	2.285	2.287	-0.001	0.955	1,567
NumCited	1.822	2.522	-0.701	0.000	1,567
CiteWeightPat	2.628	2.768	-0.141	0.000	1,567
Patent Index	35.170	32.202	2.968	0.099	1,567
Originality	0.578	0.568	0.010	0.075	1,567
Generality	0.544	0.553	-0.009	0.103	1,578
R&D/Assets	0.099	0.106	-0.008	0.000	1,408
Assets Total	7972.694	4376.186	3596.508	0.000	1,567
ROA	0.097	0.114	-0.017	0.000	1,567
Leverage	0.209	0.169	0.040	0.000	1,567
Capex/Assets	0.046	0.064	-0.017	0.000	1,567
TobinQ	2.236	3.110	-0.874	0.000	1,567
Tangibility	0.211	0.240	-0.030	0.000	1,567
HHI	0.036	0.030	0.005	0.000	1,567

Table 3.2 (continued)

Panel B: Unrelated M&As

	After	Before	Diff	*p*-value	*N*
#Patents	47.479	42.159	5.321	0.073	1,237
#Citations	76.525	90.845	-14.320	0.106	1,237
#Patent - Citationweighted	95.285	82.948	12.337	0.040	1,237
NumPat	2.042	2.131	-0.089	0.000	1,237
NumCited	1.652	2.090	-0.438	0.000	1,237
CiteWeightPat	2.328	2.487	-0.160	0.000	1,237
Patent Index	28.131	28.044	0.087	0.958	1,237
Originality	0.591	0.596	-0.004	0.514	1,237
Generality	0.561	0.586	-0.025	0.000	1,237
R&D/Assets	0.046	0.051	-0.005	0.000	951
Assets Total	11480.556	6995.843	4484.713	0.000	1,237
ROA	0.123	0.138	-0.015	0.000	1,237
Leverage	0.263	0.228	0.035	0.000	1,237
Capex/Assets	0.051	0.065	-0.014	0.000	1,237
TobinQ	1.740	1.985	-0.245	0.000	1,237
Tangibility	0.278	0.310	-0.032	0.000	1,237
HHI	0.062	0.058	0.004	0.000	1,237

Panel C: Difference-in-difference tests

	Related Diff	Unrelated Diff	DiD	*p*-value
#Patents	10.097	5.321	4.776	0.143
#Citations	-38.807	-14.320	-24.487	0.071
#Patent - Citationweighted	14.681	12.337	2.344	0.404
NumPat	-0.001	-0.089	0.088	0.008
NumCited	-0.701	-0.438	-0.263	0.000
CiteWeightPat	-0.141	-0.160	0.019	0.329
Patent Index	2.968	0.087	2.881	0.123
Originality	0.010	-0.004	0.014	0.048
Generality	-0.009	-0.025	0.016	0.032
R&D/Assets	-0.008	-0.005	-0.003	0.146
Assets Total	3596.508	4484.713	-888.205	0.035
ROA	-0.017	-0.015	-0.002	0.339
Leverage	0.040	0.035	0.005	0.175
Capex/Assets	-0.017	-0.014	-0.003	0.000
TobinQ	-0.874	-0.245	-0.629	0.000
Tangibility	-0.030	-0.032	0.002	0.167
HHI	0.005	0.004	0.001	0.247

Table 3.3: Firm-level innovation around related M&As

This table tests the effects of related M&As on firm innovation, using unrelated M&As as a control group. The sample contains observations for six years before and six years after M&A. The dependent variables are the six measures of firm innovations. The key independent variable, *After*, equals one if the observation is after the M&As and zero otherwise. *Related* equals one if it is the merger is between firms from an innovation-user industry and an innovation-producer industry and equals zero otherwise. We include M&A deal fixed effects and year fixed effects. The standard errors are clustered at the deal level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	-0.196***	-0.105***	-0.198***	-0.212***	-0.0274***	-0.0285***
	(-7.641)	(-2.934)	(-6.588)	(-8.522)	(-3.892)	(-4.195)
*After * Related*	0.185***	0.0583	0.138***	0.206***	0.0189**	0.0252***
	(5.121)	(1.160)	(3.306)	(5.870)	(2.131)	(2.942)
R&D/Assets	0.309**	0.517***	0.390**	0.387***	0.0435	0.0104
	(2.392)	(2.911)	(2.510)	(3.030)	(1.198)	(0.323)
Sales	-0.162***	-0.135**	-0.186***	-0.153***	-0.0412***	-0.0336***
	(-3.891)	(-2.320)	(-3.798)	(-4.013)	(-3.765)	(-3.173)
ROA	0.136*	0.188*	0.170*	0.172**	0.0267	0.0214
	(1.669)	(1.662)	(1.793)	(2.185)	(1.312)	(1.102)
Leverage	-0.205***	-0.178*	-0.298***	-0.200***	0.00852	-0.0278*
	(-2.945)	(-1.828)	(-3.562)	(-2.986)	(0.489)	(-1.708)
Capex/Assets	-0.215**	-0.573***	-0.339***	-0.223**	0.0620*	0.0578*
	(-1.967)	(-3.669)	(-2.580)	(-2.168)	(1.943)	(1.857)

Table 3.3 (continued)

TobinQ	-0.00843***	0.0232***	0.00552	-0.00505	-0.00367***	-0.00342***
	(-2.658)	(5.352)	(1.488)	(-1.559)	(-4.111)	(-4.026)
Tangibility	0.914***	1.496***	1.155***	0.776***	0.0549	0.00558
	(4.608)	(5.184)	(4.854)	(4.057)	(0.916)	(0.0950)
HHI	0.241**	-0.419**	0.186	0.257***	0.117***	0.108***
	(2.334)	(-2.540)	(1.367)	(2.582)	(3.471)	(3.125)
Observations	26,965	26,965	26,965	26,965	26,965	26,965
R-squared	0.872	0.823	0.855	0.860	0.648	0.683
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.4: Firm-class level innovation around M&As

This table tests the effects of M&As on firm-class level innovation output. The key independent variable, *After*, equals one if the observation is after the M&As and zero otherwise. *Targeted Class* is a dummy variable that equals one if it is the technological class of the acquirer and the industry of the target has the innovation user-producer relationship and equals zero otherwise. We include M&A deal fixed effects and class by year fixed effects. The standard errors are clustered at the deal level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	-0.0163***	0.00601	-0.0156**	-0.0131***	-0.00398**	-0.00184
	(-3.055)	(0.848)	(-2.338)	(-2.925)	(-2.218)	(-1.067)
Targeted Class	0.265***	0.296***	0.337***	0.213***	0.0679***	0.0685***
	(14.28)	(12.56)	(14.92)	(13.70)	(14.39)	(14.05)
*After * Targeted Class*	0.0473***	-0.0519**	0.0371*	0.0351**	0.0212***	0.0190***
	(2.650)	(-2.110)	(1.752)	(2.372)	(4.597)	(4.137)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	319,884	319,884	319,884	319,884	319,884	319,884
R-squared	0.533	0.502	0.509	0.509	0.403	0.416
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Class*Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.5: Firm heterogeneity – Financial constraints

This table reports the effect of M&As among financially constrained and non-constrained firms. Financially constrained (unconstrained) firms are firm whose measure of financial constraints lies above (below) the median across all firms one year before the mergers. Panel A include financially constrained firms. Panel B includes financially unconstrained firms. We include M&A deal fixed effects and class by year fixed effects. The standard errors are clustered at the deal level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Subsample of financially constrained firms

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	-0.0243***	-0.0161*	-0.0299***	-0.0200***	-0.00516*	-0.00485*
	(-3.214)	(-1.735)	(-3.193)	(-3.097)	(-1.877)	(-1.852)
Targeted Class	0.275***	0.286***	0.335***	0.223***	0.0722***	0.0742***
	(10.24)	(8.517)	(10.32)	(9.822)	(10.09)	(9.753)
*After * Targeted Class*	0.0300	-0.0793**	0.0315	0.0188	0.0189***	0.0224***
	(1.101)	(-2.144)	(0.987)	(0.820)	(2.592)	(3.095)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	145,667	145,667	145,667	145,667	145,667	145,667
R-squared	0.547	0.516	0.522	0.516	0.409	0.421
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Class*Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.5 (continued)

Panel B: Subsample of financially unconstrained firms

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	-0.00173	0.0233**	0.00374	-0.000959	-0.00233	0.00145
	(-0.227)	(2.178)	(0.388)	(-0.153)	(-0.933)	(0.603)
Targeted Class	0.253***	0.303***	0.335***	0.203***	0.0654***	0.0652***
	(10.09)	(9.390)	(10.90)	(9.674)	(10.40)	(10.28)
*After * Targeted Class*	0.0510**	-0.0433	0.0290	0.0405**	0.0176***	0.0111*
	(2.127)	(-1.301)	(1.009)	(2.063)	(2.838)	(1.812)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	166,805	166,805	166,805	166,805	166,805	166,805
R-squared	0.565	0.532	0.544	0.550	0.443	0.458
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Class*Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.6: Firm heterogeneity – Number of technological classes

This table reports the effect of M&As among firms that conduct innovation in a few technological classes and many technological classes. *Few Classes* is a dummy variable indicating whether the firm's number of technological classes is below the median across all firms one year before the mergers. We include M&A deal fixed effects and tech-class by year fixed effects. The standard errors are clustered at the deal level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	-0.0295***	-0.0450***	-0.0411***	-0.0250***	-0.00698***	-0.00839***
	(-4.558)	(-5.234)	(-5.147)	(-4.635)	(-3.145)	(-3.946)
Targeted Class	0.383***	0.459***	0.474***	0.306***	0.0879***	0.0898***
	(16.42)	(15.21)	(16.72)	(15.54)	(15.25)	(15.19)
*After * Targeted Class*	0.0261	-0.141***	0.00747	0.0188	0.0123**	0.0140**
	(1.138)	(-4.303)	(0.276)	(0.976)	(2.240)	(2.553)
*After * Few Classes*	0.0438***	0.166***	0.0841***	0.0402***	0.00940***	0.0216***
	(4.332)	(10.94)	(6.529)	(4.869)	(2.699)	(6.321)
*Targeted Class * Few Classes*	-0.399***	-0.545***	-0.459***	-0.310***	-0.0672***	-0.0715***
	(-14.75)	(-15.23)	(-13.71)	(-13.86)	(-8.858)	(-9.116)
*After * Targeted Class * Few Classes*	0.0731***	0.290***	0.100***	0.0553**	0.0275***	0.0139
	(2.583)	(7.365)	(2.957)	(2.398)	(3.032)	(1.600)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	316,849	316,849	316,849	316,849	316,849	316,849
R-squared	0.536	0.506	0.512	0.512	0.403	0.417
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Class*Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.7: Tech-class level regression

This table tests the effect of related M&As on technological level innovation output. The sample includes class-year observations. *Targeted Class* equals one if that class is involved in at least one innovation user-producer mergers and zero otherwise. The standard errors are clustered at the tech-class level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
	NumPat	*NumCited*	*CiteWeightPat*	*Originality*	*Generality*
Targeted Class	0.204***	0.0955***	0.217***	-0.00660**	0.00295
	(4.726)	(2.594)	(4.829)	(-2.126)	(1.190)
Observations	9,797	9,797	9,797	9,797	9,797
R-squared	0.886	0.847	0.877	0.717	0.727
Class FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Table 3.8: Failed Mergers

Panel A reports the steps to construct the control M&A deals. We exclude deals that are withdrawn due to reasons endogenous to innovation. In Panel B, *Treated* is a dummy equals one if it is a successful M&A and equals zero for a failed M&A. *Related* equals one if it is the merger is between firms from an innovation-user industry and an innovation-producer industry and equals zero otherwise. In Panel C, we falsely assume that the onset of treatment occurs six years before it actually happens. The standard errors are clustered at the M&A deal level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Control sample construction

191	All unsuccessful merger bids (excluding financial firms)
-12	Difference in corporate philosophy over growth strategy (not involving R&D)
-58	Other competing bids emerged and the acquisition with the competitor went through
-30	Valuation issues/Problem (not involving R&D) revealed over the course of negotiations
-4	Market/analysts expected the deal to fail
-20	Not enough information/negotiations not completed/exogenous events (e.g., 1987 crash)
67	Final control group

Table 3.8 (continued)

Panel B: Post-acquisition innovation

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	1.286***	0.281*	1.489***	0.432**	0.0743	0.0608
	(2.773)	(1.992)	(2.890)	(2.053)	(1.162)	(0.982)
Treated	0.829	0.139	0.837	0.289	0.112	0.0926
	(0.973)	(0.626)	(0.900)	(0.777)	(1.133)	(0.995)
*After * Treated*	-1.305***	-0.264*	-1.499***	-0.381***	-0.168***	-0.139**
	(-4.038)	(-1.989)	(-4.229)	(-3.410)	(-2.712)	(-2.661)
*After * Related*	-1.050**	-0.132	-1.199**	-0.510**	0.0374	0.0719
	(-2.523)	(-0.713)	(-2.513)	(-2.483)	(0.473)	(0.969)
*Treated * Related*	-0.253	0.130	-0.0556	-0.171	-0.0471	0.0479
	(-0.239)	(0.341)	(-0.0462)	(-0.385)	(-0.317)	(0.327)
*After * Treated * Related*	1.583**	0.112	1.728**	0.853***	0.0210	-0.0481
	(2.603)	(0.351)	(2.464)	(3.086)	(0.210)	(-0.514)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	472	472	472	472	472	472
R-squared	0.839	0.704	0.837	0.861	0.789	0.812
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.8 (continued)

Panel C: Falsification test

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	0.502	-0.0919	0.589	0.0295	0.112	-0.147
	(0.425)	(-0.103)	(0.358)	(0.130)	(0.488)	(-0.468)
Treated	0.175	-0.186	0.237	-0.463	0.0269	-0.220
	(0.115)	(-0.204)	(0.125)	(-0.739)	(0.107)	(-0.696)
*After * Treated*	-0.296	0.540	-0.186	0.245	0.0263	0.258
	(-0.230)	(0.598)	(-0.107)	(0.575)	(0.104)	(0.796)
*After * Related*	-0.599	0.282	-0.522	-0.146	-0.125	0.146
	(-0.478)	(0.297)	(-0.303)	(-0.509)	(-0.521)	(0.457)
*Treated * Related*	-0.725	0.977	-0.323	-0.248	-0.0873	0.251
	(-0.460)	(0.891)	(-0.156)	(-0.351)	(-0.324)	(0.776)
*After * Treated * Related*	1.251	-0.710	1.085	0.385	-0.0583	-0.259
	(0.871)	(-0.710)	(0.573)	(0.687)	(-0.222)	(-0.779)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	187	187	187	187	187	187
R-squared	0.927	0.840	0.921	0.932	0.918	0.899
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.9: Robustness check – Exclude M&As from the same industry

This table tests the effects of related M&As on firm innovation after excluding M&As where acquirers and targets are from the same industry. The sample contains observations for six years before and six years after M&As. The dependent variables are the six measures of firm innovations. The key independent variable, *After*, equals one if the observation is after the M&As and zero otherwise. *Related* equals one if it is the merger is between firms from an innovation-user industry and an innovation-producer industry and equals zero otherwise. We include M&A deal fixed effects and year fixed effects. The standard errors are clustered at the deal level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	-0.103***	-0.0530	-0.118***	-0.131***	-0.0146**	-0.0210***
	(-4.174)	(-1.552)	(-4.109)	(-5.380)	(-2.063)	(-3.110)
*After * Related*	0.157***	0.0113	0.114*	0.193***	0.00378	0.0215*
	(3.109)	(0.154)	(1.951)	(3.802)	(0.334)	(1.953)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,579	15,579	15,579	15,579	15,579	15,579
R-squared	0.880	0.834	0.865	0.864	0.646	0.686
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.10: Robustness check – Using Sub-category classification of patents

This table tests the effects of M&As on firm-subcategory level innovation output. The key independent variable, *After*, equals one if the observation is after the M&As and zero otherwise. *Targeted SubCat* is a dummy variable that equals one if it is the technological sub-category of the acquirer and the industry of the target has the innovation user-producer relationship and equals zero otherwise. We include M&A deal fixed effects and Sub-category by year fixed effects. The standard errors are clustered at the deal and subcategory level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	NumPat	*NumCited*	*CiteWeightPat*	*Ln(Patent Index)*	*Originality*	*Generality*
After	-0.181***	-0.128***	-0.230***	-0.154***	-0.0495***	-0.0427***
	(-10.30)	(-6.901)	(-11.42)	(-9.823)	(-10.57)	(-9.552)
Targeted SubCat	0.662***	0.713***	0.780***	0.573***	0.128***	0.131***
	(10.10)	(9.960)	(10.71)	(10.26)	(9.666)	(10.13)
*After * Targeted SubCat*	0.0952**	-0.000229	0.0983**	0.0835**	0.0149	0.0157*
	(2.324)	(-0.00552)	(2.099)	(2.216)	(1.612)	(1.732)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	148,179	148,179	148,179	148,179	148,179	148,179
R-squared	0.393	0.430	0.385	0.377	0.312	0.331
Deal FE	Yes	Yes	Yes	Yes	Yes	Yes
Subcat-Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.11: Robustness check – Different measure of related industries

This table re-estimate the effects of M&As on firm-level innovation output using different measures of related industries. We use the data in the previous 10 years or 3 years to calculate the citation relationship between industries. We then define the top 1, top 3, or top 5 pairs as the related industries. *Related* is a dummy variable that equals one if the acquirers and targets are from two related industries and equals zero otherwise. The key independent variable, *After*, equals one if the observation is after the M&As and zero otherwise. We include M&A deal fixed effects and year fixed effects. The standard errors are clustered at the deal level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Ln(Patent Index)	10yr top5	10yr top1	3yr top5	3yr top3	3yr top1
After	-0.192***	-0.111***	-0.213***	-0.200***	-0.0884***
	(-6.966)	(-5.594)	(-7.751)	(-8.016)	(-5.090)
*After * Related*	0.168***	0.0828**	0.200***	0.201***	0.0379
	(4.592)	(2.231)	(5.462)	(5.644)	(0.900)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	25,471	25,471	25,471	25,471	25,471
R-squared	0.862	0.861	0.862	0.862	0.861
Deal FE	Yes	Yes	Yes	Yes	Yes
Year FE	YES	YES	YES	YES	YES

Appendices

Appendix A: Chapter 1

Table A.1: Variable definitions

Innovation output	
#Patents	Natural logarithm of one plus firm *i*'s total number of patents filed (and eventually granted) in year *t*.
#Citations	Natural logarithm of one plus a firm's total number of citations received on the firm's patents filed in year *t*.
Patent Index	I follow Bena and Li (2014) to construct this measure in three steps. First, for each technology class *k* and patent application year *t*, I compute the median value of the number of awarded patents in technology class *k* with application year *t* across all firms that were awarded at least one patent in technology class *k* with application year *t*. Second, I scale the number of awarded patents to the acquirer/target firm in technology class *k* with application year *t* by the corresponding technology class-specific and application year-specific median value from the first step. Third, for each firm, I sum the scaled number of awarded patents from the second step across all technology classes and application years.
Generality	One minus the Herfindahl index of the citations received by the patents filed in year *t* based on technology classes.
Originality	One minus the Herfindahl index of the citations made by the patents filed in year *t* based on technology classes.

Board of directors	
Female Director Ratio	The fraction of directors that are female.
Female Independent Director Ratio	The fraction of directors that are female independent directors. ISS defines independent directors are as directors that have no material connection to the company other than a board seat. The connection should not potentially influence one's objectivity in the boardroom.

Table A.1 (continued)

Female Non-Independent Director Ratio	The fraction of directors that are female inside or affiliated directors. Inside directors are directors that are employees, highly paid, or beneficial owners of the company. The majority of affiliated directors are former CEO, non-CEO executives, family members, and directors that are linked to the firm through other relationships.
Local Female Director Ratio	The average female director ratio of firms within a 60 miles radius.
Connected Male Director Ratio	The fraction of male directors on the board who sit on other boards of nonlocal firms on which there are female directors. Nonlocal firms are firms that are more than 60 miles away.
R&D/Tech experts	Independent directors with corporate experience at firms with positive R&D or in the high-tech industry (Baginski, Hassell, and Kimbrough, 2004; Knyazeva, Knyazeva, and Masulis, 2013).

Table A.2: Primary and nonprimary firms of female and male directors

The table presents the difference between directors' primary employers and other firms they work for as independent directors. The sample includes directors that sit on more than one board, and that ISS clearly identifies one of the firms as the primary employer. The sample includes 6,761 director-primary-nonprimary-year observations. Panel A reports the firm and headquarter county characteristics of the primary and nonprimary firms of directors of a different gender. Panel B reports the distance between the primary firm and nonprimary firms. Panel C presents the percentage of nonprimary firms that are in the same industry as the primary firm. Industries are classified by different levels of SIC codes. Related industries are the industries that are major suppliers or customers industries. Panel D shows the percentage of primary and nonprimary firms of female (male) directors that are connected by another male director. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Firm and county characteristics by primary employer and gender

	Female Directors			Male Directors		
	Primary	Nonprimary	Diff	Primary	Nonprimary	Diff
Total Assets	9010.982	22791.758	13780.776***	10581.293	13440.547	2859.254***
Sales	6374.033	12163.415	5789.382***	6998.194	7801.847	803.653***
Market Cap	8187.201	16523.400	8336.199***	9001.343	9846.216	844.872***
M/B	2.113	2.692	0.579**	1.941	1.974	0.033
ROA	0.149	0.164	0.015***	0.146	0.151	0.005***
R&D	324.814	475.887	151.073***	240.368	257.755	17.387**
Firm Age	28.946	35.543	6.598**	31.703	35.603	3.9***
Leverage	0.228	0.207	-0.021	0.233	0.233	0.000
R&D/Assets	0.067	0.048	-0.02***	0.043	0.043	0.000
Tangibility	0.236	0.231	-0.005	0.282	0.285	0.003
Dividend Yield	0.011	0.020	0.009***	0.016	0.018	0.002***
Cash/Assets	0.159	0.146	-0.012	0.119	0.116	-0.003
Population	13.967	13.879	-0.088	13.725	13.772	0.048***
Personal Income Per Capita	10.516	10.512	-0.004	10.476	10.473	-0.003**
Personal Income Per Capita Growth	3.692	3.711	0.019	3.663	3.682	0.019

Table A.2 (continued)

Panel B: Distance between primary and nonprimary firms

	Mean	25%	50%	75%	St. Dev.	N
Female	620.274	19.981	254.750	881.992	789.537	217
Male	610.797	28.669	372.691	907.976	703.156	6,544
Diff (t-value)	9.477 (0.195)					

Panel C: Same industry

	SIC-4	SIC-3	SIC-2 and related industries	SIC-1	N
Female	3%	7%	12%	32%	217
Male	4%	6%	14%	38%	6,544

Panel D: Connected by male directors

	Connected	Not connected	N
Female	86%	14%	217
Male	9%	91%	6,544

Appendix B: Chapter 2

Table B.1: Variable definitions

Firm and stock level variables

Mutual fund ownership (MFO)	The percentage of shares outstanding that are held by mutual funds.
PurePlay	The organizational form dummy which equals one if the firm is a pure-play firm (operates in one 3-digit SIC industry), and zero otherwise.
Market Cap	Market capitalization, defined as the dollar value of equity at the end of each quarter.
Stock Return	The monthly return on the firm's stock at the end of the quarter.
Dividend Yield	The ratio of total dividend payout to stock price.
Return Volatility	The volatility of stock returns is calculated using monthly stock returns over the prior twelve months.
Stock Turnover	The quarterly share turnover ratio of the outstanding shares.
M/B	The market value of assets divided by the book value of assets.
Leverage	The ratio of total debt to the market value of assets.
Price	The end-of-quarter stock price.
Firm Age	The number of years since the firm first appears in CRSP.
S&P500	An S&P 500 index membership dummy, which equals one if the firm is in the S&P 500 Index, and zero otherwise.
ROA	The net income divided by the total assets.
Spread	The average of the monthly trading spread of the stock in each quarter.

Mutual fund level variables

Pureplayness	$Pureplayness = \sum_{j=1}^{N}(w_j * Pureplay_j)$, where w_j is the portfolio weight of stock j in the fund, and *Pureplay* is a dummy that equals one for pure-play firms, and zero for conglomerates.
Industry Concentration Index (ICI)	We follow Kacperczyk, Sialm, and Zheng (2005) to define it as the sum of the squared deviations of the value weights for each of the 10 different industries held by the mutual fund, $W_{j,t}$, relative to the industry weights of the total stock market, $\overline{W}_{j,t}$: $ICI_j = \sum_{j=1}^{10}(W_{j,t} - \overline{W}_{j,t})^2$.
Fund Size	The total net assets (TNA) in millions.
Fund Turnover	The turnover ratio in each quarter.
Expense	The annual expense ratio.
#Stocks	The total number of stocks reported in the quarterly holding file.
Fund Age	The number of years since the fund first appears in CRSP.
Team	A dummy that equals one if a fund is managed by two or more managers, and zero otherwise.
Net Inflow	$TNA_t = TNA_{t-1} * (1 + r_t)$
Fund Returns	Fund's percentage returns

Table B.2: Institutional ownership, hedge fund ownership, and single-segment firms

Tables in the appendix test whether institutional investors or hedge funds prefer pureplay firms. We use the same list of hedge funds used in Aiken, Clifford, Ellis, and Huang (2017). Panel A and B repeat Tables 2.2 and 2.3 but use hedge fund ownership and other institutional investors' ownership, respectively. Other institutions are institutional investors that are not actively managed mutual funds or hedge funds. Panel C and D repeat Panel C of Table 2.4 to test industry expertise and preference for pure-play firms. Standard errors are clustered by firm and quarter. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Organizational form and institutional ownership

	(1)	(2)	(3)	(4)
	IO other	*IO other*	*HFO*	*HFO*
PurePlay	0.209	0.662	0.323	0.297
	(0.370)	(1.243)	(1.014)	(0.991)
Ln(Mkt cap)	6.909***	4.547***	1.801***	1.257***
	(30.78)	(14.55)	(19.40)	(8.107)
Stock return	-3.656***	-0.907	-0.403	-0.256
	(-4.031)	(-1.127)	(-1.064)	(-0.722)
Dividend yield	22.49***	7.754*	-15.82***	-18.47***
	(4.353)	(1.784)	(-8.400)	(-8.203)
Return volatility	-12.58***	-18.44***	-5.625***	-6.740***
	(-5.136)	(-7.880)	(-4.457)	(-5.017)
M/B		-1.417***		-0.491***
		(-10.55)		(-6.901)
Leverage		7.273***		-0.893
		(6.363)		(-1.505)
Ln(Price)		3.527***		1.192***
		(8.500)		(5.593)
Ln(Firm age)		0.561*		0.573***
		(1.713)		(2.801)
S&P 500		-0.443		-6.521***
		(-0.442)		(-12.43)
ROA		-2.036***		-0.600*
		(-3.084)		(-1.734)
Stock turnover		0.679***		0.227***
		(13.73)		(10.04)
Spread		-104.9***		-51.65***
		(-7.319)		(-5.401)
Observations	99,464	96,991	99,464	96,991
R-squared	0.433	0.487	0.226	0.277
Industry FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Table B.2 (continued)

Panel B: Falsification tests

	(1)	(2)
	IO other	*HFO*
Actual	0.955	-1.004**
	(1.337)	(-2.384)
Controls	Yes	Yes
Observations	26,331	26,331
R-squared	0.711	0.538
Firm FE	Yes	Yes
Time FE	Yes	Yes

Panel C: Industry concentration of other institutional investors

	(1)	(2)	(3)	(4)	(5)
	IO Other ICI1	*IO Other ICI2*	*IO Other ICI3*	*IO Other ICI4*	*IO Other ICI5*
PurePlay	0.938**	-0.0728	0.0919	0.0641	0.104
	(2.213)	(-0.780)	(0.929)	(0.510)	(1.037)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	96,991	96,991	96,991	96,991	96,991
R-squared	0.542	0.090	0.090	0.053	0.084
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Panel D: Industry concentration of hedge funds

	(1)	(2)	(3)	(4)	(5)
	HFO ICI1	*HFO ICI2*	*HFO ICI3*	*HFO ICI4*	*HFO ICI5*
PurePlay	0.176	-0.0437	0.0743	0.0189	0.0273
	(0.989)	(-0.433)	(1.468)	(0.395)	(0.728)
	(-3.003)	(-6.069)	(-5.758)	(-6.962)	(-6.591)
Observations	96,991	96,991	96,991	96,991	96,991
R-squared	0.333	0.113	0.059	0.039	0.057
Industry FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Appendix C: Chapter 3

Table C.1: Variable definitions

NumPat: Natural logarithm of one plus firm i's total number of patents filed (and eventually granted) in year t.

NumCited: Natural logarithm of one plus a firm's total number of citations received on the firm's patents filed in year t.

CiteWeightPat: Natural logarithm of one plus the number of citation-weighted patents based on the total citations received by firm i in year t. Citation-weighted patents are calculated as $\sum_j (1 + C_j/\overline{C_j})$, where C_j is the number of citations to patent j and $\overline{C_j}$ is the mean number of citations to patents granted in the same year as patent j.

Generality: One minus the Herfindahl index of the citations received by the patent portfolio (patents filed by the firm in the previous five years) in year t based on technology classes.

Originality: One minus the Herfindahl index of the citations made by the patent portfolio (patents filed by the firm in the previous five years) in year t based on technology classes.

Patent Index: This measure is constructed in three steps. First, for each technology class k and patent application year t, we compute the median value of the number of awarded patents in technology class k with application year t across all firms that were awarded at least one patent in technology class k with application year t. Second, we scale the number of awarded patents to the acquirer/target firm in technology class k with application year t by the corresponding technology class-specific and application year-specific median value from the first step. Third, for the acquirer/target firm, we sum the scaled number of awarded patents from the second step across all technology classes and across application years.

References

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, *94*(2), 291-309.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., & Howitt, P. (2005). Competition and innovation: An inverted-U relationship. *The Quarterly Journal of Economics*, *120*(2), 701-728.

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica* 60(2), 323-351.

Arrow, K. J. (1971). The economic implications of learning by doing. In *Readings in the Theory of Growth* (pp. 131-149). Palgrave Macmillan, London.

Baginski, S. P., Hassell, J. M., & Kimbrough, M. D. (2004). Why do managers explain their earnings forecasts? *Journal of Accounting Research*, *42*(1), 1-29.

Baker, M., Stein, J. C., & Wurgler, J. (2003). When does the market matter? Stock prices and the investment of equity-dependent firms. *The Quarterly Journal of Economics*, *118*(3), 969-1005.

Balsmeier, B., Fleming, L., & Manso, G. (2017). Independent boards and innovation. *Journal of Financial Economics*, *123*(3), 536-557.

Bargeron, L., Lehn, K., & Smith, J. (2015). Employee–management trust and M&A activity. *Journal of Corporate Finance*, *35*, 389-406.

Bayar, O., Chemmanur, T. J., & Liu, M. H. (2019). How to Motivate Fundamental Innovation: Optimal Interactions between Entrepreneurs, Venture Capitalists, and the Government. *Available at SSRN 2799490*.

Beckman, C. M., & Haunschild, P. R. (2002). Network learning: The effects of partners' heterogeneity of experience on corporate acquisitions. *Administrative science quarterly*, *47*(1), 92-124.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2018). Who becomes an inventor in America? The importance of exposure to innovation. *The Quarterly Journal of Economics*, *134*(2), 647-713.

Bernile, G., Bhagwat, V., & Yonker, S. (2018). Board diversity, firm risk, and corporate policies. *Journal of Financial Economics*, *127*(3), 588-612.

Bernstein, S. (2015). Does going public affect innovation? *The Journal of*

Finance, *70*(4), 1365-1403.

Bena, J., & Li, K. (2014). Corporate innovations and mergers and acquisitions. *The Journal of Finance*, *69*(5), 1923-1960.

Berger, P. G., & Ofek, E. (1996). Bustup takeovers of value-destroying diversified firms. *The Journal of Finance*, *51*(4), 1175-1200.

Bloom, N., Schankerman, M., & Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, *81*(4), 1347-1393.

Bloom, N., & Van Reenen, J. (2002). Patents, real options and firm performance. *The Economic Journal*, *112*(478), C97-C116.

Brav, A., Jiang, W., Ma, S., & Tian, X. (2018). How does hedge fund activism reshape corporate innovation? *Journal of Financial Economics*, *130*(2), 237-264.

Brav, A., Jiang, W., Partnoy, F., & Thomas, R. (2008). Hedge fund activism, corporate governance, and firm performance. *The Journal of Finance*, *63*(4), 1729-1775.

Brown, J. R., Fazzari, S. M., & Petersen, B. C. (2009). Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. *The Journal of Finance*, *64*(1), 151-185.

Bushee, B. J., & Noe, C. F. (2000). Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research*, 171-202.

Campa, J. M., & Kedia, S. (2002). Explaining the diversification discount. *The Journal of Finance*, *57*(4), 1731-1762.

Campello, M. (2002). Internal capital markets in financial conglomerates: Evidence from small bank responses to monetary policy. *The Journal of Finance*, *57*(6), 2773-2805.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.

Cassiman, B., Colombo, M. G., Garrone, P., & Veugelers, R. (2005). The impact of M&A on the R&D process: An empirical analysis of the role of technological-and market-relatedness. *Research Policy*, *34*(2), 195-220.

Chang, X., Fu, K., Low, A., & Zhang, W. (2015). Non-executive employee stock options and corporate innovation. *Journal of Financial Economics*, *115*(1), 168-188.

Chakrabarty, B., Moulton, P. C., & Trzcinka, C. (2017). The performance of short-term institutional trades. *Journal of Financial and Quantitative Analysis*, *52*(4), 1403-1428.

Chemmanur, T. J., Loutskina, E., & Tian, X. (2014). Corporate venture capital, value creation, and innovation. *The Review of Financial Studies*, *27*(8), 2434-2473.

Chen, C., Chen, Y., Hsu, P. H., & Podolski, E. J. (2016). Be nice to your innovators: Employee treatment and corporate innovation performance. *Journal of Corporate Finance*, *39*, 78-98.

Chen, J., Leung, W. S., & Evans, K. P. (2018). Female board representation, corporate innovation and firm performance. *Journal of Empirical Finance*, *48*, 236-254.

Chhaochharia, V., & Grinstein, Y. (2009). CEO compensation and board structure. *The Journal of Finance*, *64*(1), 231-261.

Coase, R. H. (1995). The nature of the firm. In *Essential readings in economics* (pp. 37-54). Palgrave, London.

Coles, J. L., Daniel, N. D., & Naveen, L. (2014). Co-opted boards. *The Review of Financial Studies*, *27*(6), 1751-1796.

Cremers, K. M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *The Review of financial studies*, *22*(9), 3329-3365.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance*, *52*(3), 1035-1058.

Davis, G. F., Yoo, M., & Baker, W. E. (2003). The small world of the American corporate elite, 1982-2001. *Strategic Organization*, *1*(3), 301-326.

Ding, W. W., Murray, F., & Stuart, T. E. (2006). Gender differences in patenting in the academic life sciences. *Science*, *313*(5787), 665-667.

Del Guercio, D. (1996). The distorting effect of the prudent-man laws on institutional equity investments. *Journal of Financial Economics*, *40*(1), 31-62.

Denis, D. J., Denis, D. K., & Sarin, A. (1997). Agency problems, equity ownership, and corporate diversification. *The Journal of Finance*, *52*(1), 135-160.

Duchin, R. (2010). Cash holdings and corporate diversification. *The Journal of*

Finance, *65*(3), 955-992.

Eckbo, B. E., Nygaard, K., & Thorburn, K. S. (2016). Board gender-balancing and firm value. *European Corporate Governance Institute (ECGI)-Finance Working Paper*, (463).

Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics*, *101*(3), 621-640.

Fahlenbrach, R., Low, A., & Stulz, R. M. (2017). Do independent director departures predict future bad events? *The Review of Financial Studies*, *30*(7), 2313-2358.

Falkenstein, E. G. (1996). Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *The Journal of Finance*, *51*(1), 111-135.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1-22.

Farrell, M. (2018). Read Between the Filings: Daily Mutual Fund Holdings and Liquidity Provision. *Available at SSRN 3241664*.

Farrell, K. A., & Hersch, P. L. (2005). Additions to corporate boards: the effect of gender. *Journal of Corporate Finance*, *11*(1-2), 85-106.

Gaspar, J. M., & Massa, M. (2011). The role of commonality between CEO and divisional managers in internal capital markets. *Journal of Financial and Quantitative Analysis*, *46*(3), 841-869.

Giannetti, M., & Wang, T. Y. (2019). Public Attention to Gender Equality and the Demand for Female Directors. *Available at SSRN 3341518*.

Giroud, X., & Mueller, H. M. (2015). Capital and labor reallocation within firms. *The Journal of Finance*, *70*(4), 1767-1804.

Gompers, P., Lerner, J., & Scharfstein, D. (2005). Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *The Journal of Finance*, *60*(2), 577-614.

Griffith, R., Harrison, R., & Van Reenen, J. (2006). How special is the special relationship? Using the impact of US R&D spillovers on UK firms as a test of technology sourcing. *American Economic Review*, *96*(5), 1859-1875.

Griffin, D., Li, K., & Xu, T. (2019). Board Gender Diversity and Corporate Innovation: International Evidence. *Journal of Financial and Quantitative Analysis*, 1-68.

Güner, A. B., Malmendier, U., & Tate, G. (2008). Financial expertise of directors. *Journal of Financial Economics*, *88*(2), 323-354.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: Lessons, insights and methodological tools* (No. w8498). National Bureau of Economic Research.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 16-38.

Hall, B. H., & Lerner, J. (2010). The financing of R&D and innovation. In *Handbook of the Economics of Innovation* (Vol. 1, pp. 609-639). North-Holland.

Hann, R. N., Ogneva, M., & Ozbas, O. (2013). Corporate diversification and the cost of capital. *The Journal of Finance*, *68*(5), 1961-1999.

Harris, M., Kriebel, C. H., & Raviv, A. (1982). Asymmetric information, incentives and intrafirm resource allocation. *Management Science*, *28*(6), 604-620.

He, J., & Huang, Z. (2011). Board informal hierarchy and firm financial performance: Exploring a tacit structure guiding boardroom interactions. *Academy of Management Journal*, *54*(6), 1119-1139.

Hermalin, B. E., & Weisbach, M. S. (1988). The determinants of board composition. *The RAND Journal of Economics*, 589-606.

Jame, R. (2017). Liquidity provision and the cross section of hedge fund returns. *Management Science*, *64*(7), 3288-3312.

Jensen, K., Kovács, B., & Sorenson, O. (2018). Gender differences in obtaining and maintaining patent rights. *Nature Biotechnology*, *36*(4), 307.

Ruback, R. S., & Jensen, M. C. (1983). The market for corporate control: The scientific evidence. *Journal of Financial Economics*, *11*, 5-50.

Jiang, W. (2017). Have instrumental variables brought us closer to the truth. *The Review of Corporate Finance Studies*, *6*(2), 127-140.

Joenväärä, J., Kosowski, R., & Tolonen, P. (2012). Revisiting stylized facts about hedge funds. *Imperial College Business School*.

Jordan, B. D., Liu, M. H., & Wu, Q. (2018). Organizational form and corporate payout policy. *Journal of Financial and Quantitative Analysis*, *53*(2), 789-813.

Jordan, B. D., & Riley, T. B. (2015). Volatility and mutual fund manager skill. *Journal of Financial Economics*, *118*(2), 289-298.

Kacperczyk, M., Sialm, C., & Zheng, L. (2005). On the industry concentration of actively managed equity mutual funds. *The Journal of Finance*, *60*(4), 1983-2011.

Kaplan, S. N., & Zingales, L. (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints?. *The Quarterly Journal of Economics*, *112*(1), 169-215.

Knyazeva, A., Knyazeva, D., & Masulis, R. W. (2013). The supply of corporate directors and board independence. *The Review of Financial Studies*, *26*(6), 1561-1605.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, *132*(2), 665-712.

Koh, P. S., & Reeb, D. M. (2015). Missing r&d. *Journal of Accounting and Economics*, *60*(1), 73-94.

Kurtulus, F. A., & Tomaskovic-Devey, D. (2012). Do female top managers help women to advance? A panel study using EEO-1 records. *The ANNALS of the American Academy of Political and Social Science*, *639*(1), 173-197.

Lang, L. H., & Stulz, R. M. (1994). Tobin's q, corporate diversification, and firm performance. *Journal of Political Economy*, *102*(6), 1248-1280.

Levi, M., Li, K., & Zhang, F. (2014). Director gender and mergers and acquisitions. *Journal of Corporate Finance*, *28*, 185-200.

Levine, R. (2005). Finance and growth: theory and evidence. *Handbook of economic growth*, *1*, 865-934.

Li, G. C., Lai, R., D'Amour, A., Doolin, D. M., Sun, Y., Torvik, V. I., Amy Z. Y., & Fleming, L. (2014). Disambiguation and co-authorship networks of the US patent inventor database (1975–2010). *Research Policy*, *43*(6), 941-955.

Li, K., Ortiz-Molina, H., & Zhao, X. (2008). Do voting rights affect institutional investment decisions? Evidence from dual-class firms. *Financial Management*, *37*(4), 713-745.

Liu, M. H. (2011). Analysts' incentives to produce industry-level versus firm-specific information. *Journal of Financial and Quantitative Analysis*, *46*(3), 757-784.

Maksimovic, V., & Phillips, G. M. (2013). Conglomerate firms, internal capital markets, and the theory of the firm. *Annu. Rev. Financ. Econ.*, *5*(1), 225-244.

Manso, G. (2011). Motivating innovation. *The Journal of Finance*, *66*(5), 1823-1860.

Masulis, R. W., & Mobbs, S. (2011). Are all inside directors the same? Evidence from the external directorship market. *The Journal of Finance*, *66*(3), 823-872.

Matsa, D. A., & Miller, A. R. (2011). Chipping away at the glass ceiling: Gender spillovers in corporate leadership. *American Economic Review*, *101*(3), 635-39.

Mayer, R. C., Warr, R. S., & Zhao, J. (2018). Do Pro-Diversity Policies Improve Corporate Innovation? *Financial Management*, *47*(3), 617-650.

Mazzucato, M. (2015). The Entrepreneurial State: Debunking Public vs. Private Sector Myths (revised ed.). *Philadelphia, PA: PublicAffairs*.

McNeil, C., Niehaus, G., & Powers, E. (2004). Management turnover in subsidiaries of conglomerates versus stand-alone firms. *Journal of Financial Economics*, *72*(1), 63-96.

Mitton, T., & Vorkink, K. (2010). Why do firms with diversification discounts have higher expected returns?. *Journal of Financial and Quantitative Analysis*, *45*(6), 1367-1390.

Medland, D. (2004). Small steps for womankind. *Corporate Board Member Europe, Winter*.

Nelson, R. R. (1959). The simple economics of basic scientific research. *Journal of Political Economy*, *67*(3), 297-306.

Nguyen, B. D., & Nielsen, K. M. (2010). The value of independent directors: Evidence from sudden deaths. *Journal of Financial Economics*, *98*(3), 550-567.

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2015). Scale and skill in active management. *Journal of Financial Economics*, *116*(1), 23-45.

Pinnuck, M. (2003). An examination of the performance of the trades and stock holdings of fund managers: Further evidence. *Journal of Financial and Quantitative Analysis*, *38*(4), 811-828.

Phillips, G. M., & Zhdanov, A. (2013). R&D and the incentives from merger and acquisition activity. *The Review of Financial Studies*, *26*(1), 34-78.

Rajan, R., Servaes, H., & Zingales, L. (2000). The cost of diversity: The diversification discount and inefficient investment. *The Journal of Finance*, *55*(1), 35-80.

Savor, P. G., & Lu, Q. (2009). Do stock mergers create value for acquirers?. *The Journal of Finance*, *64*(3), 1061-1097.

Schwartz-Ziv, M. (2017). Gender and board activeness: The role of a critical mass. *Journal of Financial and Quantitative Analysis*, *52*(2), 751-780.

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

Sevilir, M., & Tian, X. (2012). Acquiring Innovation. *Available at SSRN 1731722*.

Staines, G., Tavris, C., & Jayaratne, T. E. (1974). The queen bee syndrome. *Psychol Today*, 7, 55-60.

Stein, J. C. (1997). Internal capital markets and the competition for corporate resources. *The Journal of Finance*, *52*(1), 111-133.

Tate, G., & Yang, L. (2015). Female leadership and gender equity: Evidence from plant closure. *Journal of Financial Economics*, *117*(1), 77-97.

Torchia, M., Calabrò, A., & Huse, M. (2011). Women directors on corporate boards: From tokenism to critical mass. *Journal of Business Ethics*, *102*(2), 299-317.

Villalonga, B. (2004). Diversification discount or premium? New evidence from the business information tracking series. *The Journal of Finance*, *59*(2), 479-506.

Wermers, R. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *The Journal of Finance*, *55*(4), 1655-1695.

Zhou, X. (2001). Understanding the determinants of managerial ownership and the link between ownership and performance: comment. *Journal of Financial Economics*, *62*(3), 559-571.

Vita

Ang Li

Education

City University of Hong Kong
M.Sc. in Financial Engineering 2013
Zhongnan University of Economics and Law
B.S. in Financial Engineering 2012

Teaching Experience

Independent Instructor
- FIN 410: Investment Analysis Summer 2018; 2019-2020
- FIN 300: Corporate Finance Summer 2019; 2019-2020

Teaching Assistant
- B&E 105: Technology for Business Solutions 2017-2018
- FIN 600: Corporate Financial Policy (MBA) 2017-2018
- FIN 464: Real Estate Finance 2016-2017

Honors and Awards

Certifications
- CFA Level II Candidate Jun. 2020
- Financial Risk Manager (FRM) Jun. 2017
- Bloomberg Essentials Certification Sep. 2015

Awards
- Graduate Assistantship 2015-2020
- Fellowship 2015-2019

Professional Activities

Conferences
- Eastern Finance Association (Program Committee) 2020
- FMA (Presenter); MFA (Presenter) 2019
- FMA Asia/Pacific (Presenter, Discussant, and Session Chair);
  MFA (Presenter and Discussant); EasternFA (Presenter and Discussant) 2018
- FMA (Discussant) 2016

Reviewer
- Pacific-Basin Finance Journal (*4) 2018, 2019