Social Capital and Innovation:

Evidence from Connected Holdings

Click Here for the Latest Version

Hongxun Ruan*

November 12, 2017

Abstract

This paper investigates how social capital affects innovation. I measure a firm's social capital with connected holdings, which is the fraction of equity of a particular firm held by mutual funds whose managers are connected to the firm's board members through educational networks. I use plausibly exogenous variation in the size of board members' networks as an instrument for connected holdings. I find higher connected holdings lead to larger number of patents granted, more patent citations, and higher firm value created by patents. Connected holdings foster innovation by helping to reduce short-term capital market pressures and to increase management job security.

Key words: social capital, innovation, patents, mutual funds, corporate governance
JEL codes: G23, G34, O32

*The Wharton School, University of Pennsylvania, hongxunr@wharton.upenn.edu. I am grateful to Jules van Binsbergen, David Musto and Nikolai Roussanov for their invaluable guidance and support. I also thank Aymeric Bellon, Xingyi Chen, Andrea Contigiani, Anna Cororaton, Adrian Aycan Corum, Paul Decaire, Winston Dou, Amora Elsaify, Xiang Fang, Erik Gilje, Joao Gomes, Qing Gong, Deeksha Gupta, Mete Kilic, Doron Levit, Karen Lewis, Jianan Liu, Tong Liu, Yang Liu, Don Keim, Christian Opp, Sebastien Plante, Krishna Ramaswamy, Lin Shen, Luke Taylor, Rudi Weber, Yanhao Wei, Ram Yamarthy, Weilong Zhang, Yicheng Zhu, audience at Wharton and the Mack Innovation Doctoral Association. Special thanks to Prof Luke Taylor, Wei Jiang, Bob Parrino for sharing the data. I am grateful to Mack Institute for Innovation Management, Jacobs Levy Equity Management Center for Quantitative Financial Research and Rodney L. White Center for Financial Research for financial support. All errors are mine.

"Virtually every commercial transaction has within itself an element of trust, certainly any transaction conducted over a period of time."
Kenneth Arrow (1972)

1 Introduction

Economists have long been interested in the impact of social capital on economic outcomes. In the literature, there are two conflicting views: on one hand, Putnam (1993) argues that "associations instill in their members habits of cooperation, solidarity, and public-spiritedness". The cooperative attitude and trust among group members could mitigate informational frictions and enable welfare-improving transactions.[1] On the other hand, Olson (1982) argues that associational activities foster emergence of self-serving interest groups such as cartels, colluding elites and lobbies which distort markets and hurt economic growth.[2] The goal of this paper is to shed new light on the relative importance of these two conflicting views. Using a measure of social connections between board members of publicly traded firms and mutual fund managers, I find that more social capital leads to more innovation at the firm level. The evidence suggests that the positive aspects of social capital dominate the potential negative ones, at least with respect to corporate innovation.

There are two major challenges associated with empirically estimating the effect of social capital on economic outcomes. The first one is measurement of social capital. The second is identification, as purely exogenous variation in social capital is rare. Previous empirical research on the question usually has focused on country-level measures of trust and has used cross-country variation to identify the impact on economic outcomes (e.g., Knack & Keefer 1997; Guiso *et al.* 2009).

To address the first challenge, I construct a firm-level measure of social capital: connected holdings, which is defined as the fraction of equity of a particular firm held by mutual funds whose managers are connected to the firm's board members through educational networks. To address the second challenge, I exploit the plausibly exogenous variation in connected mutual fund managers' entry and exit of asset management industry and use it as an instrument for connected holdings. Since innovation is important in generating long-term economic growth, I examine the impact of connected holdings on innovation.

In a *frictionless* economy, social capital should not matter because financial markets can allocate the efficient amount of capital to innovative firms. In the presence of frictions such as asymmetric information, communication costs, incomplete contracting, etc., social capital could either encourage or impede firm innovation. On one hand, social capital could promote innova-

[1] See Knack & Keefer (1997), Guiso *et al.* (2004), Guiso *et al.* (2008), Bloom *et al.* (2012).
[2] See Fisman (2001), Khwaja & Mian (2005), Haselmann *et al.* (2017).

tion by alleviating short-term capital market pressures. This might be the case if capital market investors value short-term growth in firm profits or find it hard to evaluate long-term innovative projects.[3] In such an environment, management chooses to underinvest in research and development (R&D). Social capital in the form of connected holdings could alleviate capital market pressures by acting like long-term investment and supporting firms in adverse situations ("the insurance effect"). On the other hand, social capital might hurt innovation (and, more generally, firm value) because connected holdings could increase management entrenchment and weaken capital market's disciplinary effects. Classical moral hazard models inform us that under weaker monitoring, management could enjoy private benefits (e.g., through shirking or a quiet life) at the cost of firm value, which can hurt innovation (Holmström 1979, Shleifer & Vishny 1989, Bertrand & Mullainathan 2003).

Identifying the effect of social capital on firm actions is complicated by potentially endogenous selection. Connected mutual fund managers might select firms to invest in based on the information that is observable to them, but not to the econometrician. An omitted factor could be the quality of inventors at the firm, the long-term innovation plans, or the prototypes of unpatented inventions.

I classify a firm and a mutual fund as "connected" if a fund manager and a board member attended the same university at the same time and received the same type of degree. I address the potential selection issues by instrumenting the connected holdings with the total number of connected mutual fund managers (including those who do not hold any shares of the firm in their portfolio). More specifically, the instrument is defined as follows: in each year, for each firm in my sample, I count the number of mutual fund managers who are currently working in the asset management industry and connected to the firm. The main idea behind the instrument is that the time-series changes of the number of connected mutual fund managers affect connected holdings, but are not affected by firm innovation.

For an instrument to be valid, it needs to satisfy both the relevance condition and the exclusion condition. I formally test for the relevance condition. First-stage F-statistic estimates indicate that the instrument is highly correlated with the endogenous variable connected holdings. For the exclusion condition to hold, the instrument should not affect firm innovation through any channel other than connected holdings. After teasing out the *average* number of connected mutual fund managers for each board member using university fixed effects, the time-series variation in the instrument comes from the within variation of each board member's *individual* connections to the mutual fund managers. And the time-series variation of each board member's *individual* connections to the mutual fund managers is generated by the entry and exit of mutual fund managers, which is plausibly exogenous to firm innovation due to the fact that each connected firm on average constitutes less than 50 bps of the connected mutual fund manager's portfolio.

[3] See, for example, Shleifer & Vishny (1990), Stein (1989), Bebchuk & Stole (1993), Narayanan (1985).

Using the instrumental variable approach, I establish a causal link between connected holdings and innovation. First, I investigate the impact of connected holdings on innovation outcomes. A one standard deviation increase in connected holdings causes a 10.61% increase in the total number of patents granted, a 14.86% increase in the total number of citations, and a 2.21% increase in firm value created by patents that equates to about 23 million dollars for the average firm. Two factors affect innovation outcomes: input into innovation (Hausman *et al.* 1984) and innovation efficiency (Hirshleifer *et al.* 2013). A one standard deviation increase in connected holdings translates into a 12.61% increase in firm's input into innovation as measured by R&D expenditure. This result suggests that more connected holdings enable firms to pursue riskier innovative projects, a finding consistent with the insurance effect of connected holdings. Relationship between connected holdings and commonly used measures of innovation efficiency, such as the ratios of either patents or citations to R&D capital, is not statistically significant. This result indicates that the negative effects induced by connected holdings are offset by positive effects, otherwise we should see a significant drop in innovation efficiency. More interestingly, innovation efficiency as measured by the ratio between the value created by patents and R&D capital is positively associated with connected holdings. This implies that connected holdings might incentivize firms to produce more valuable patents, but not more cited patents.

The impact of connected holdings on innovation is heterogeneous across firms and funds. I find that firms in industries that most rely on patents (such as pharmaceutical and IT industry), their innovations more sensitively react to the change in connected holdings. This suggests that firms in more innovative industries might be more constrained by capital market pressures. Consistent with Aghion *et al.* (2013), I find that the connected holdings coming from relatively passive funds have no impact on innovation, whereas connected holdings from truly active funds positively affect innovation. This result is intuitive: in bad situations, active funds have more degree of freedom to support the firm while passive funds need to follow some exogenously set rules which might restrain them from supporting the firm. In addition to innovation, I check how connected holdings affect other dimensions of the firm performance. I find that higher connected holdings predict higher firm growth in profits, output, and the number of employees. This finding is consistent with Lins *et al.* (2017).

How do connected holdings encourage innovation? Based on previous theories, I focus on two explanations: (1) connected holdings alleviate short-term capital market pressures (Stein 1989) and (2) connected holdings increase management's job security (Manso 2011; Stein 1988). Capital market pressures force firms to focus excessively on short-term earnings at the cost of long-term investment, such as innovation. I hypothesize that connected holdings could potentially reduce firm short-termism through loyal investments. Meanwhile, connected holdings could increase job security of the management by reducing takeover risk through their deterrence to corporate raiders.

Consistent with both theories, I find that, first, after a firm misses their quarterly earnings target, connected mutual funds stick with the firm, whereas non-connected funds divest significantly. Second, I find that, after firms miss their quarterly earnings targets, the stock returns (as measured by one-month-ahead abnormal returns) drop less for the firms with connected holdings. Third, I find connected holdings reduce firm's takeover exposure which increases job security of the board members. In addition, I also find that, on average, connected funds are more likely to vote against shareholder-initiated proposals on various governance issues than are non-connected funds.

My findings have two broad implications. The literature on corporate governance typically views management entrenchment as harmful to shareholder values. In this paper, I show that, when the firm wants to incentivize management to be more innovative, the firm should allow for some degree of management entrenchment. The increase in job security could increase management's incentives to take on more risks associated with innovation. My paper also sheds new light on the role of active funds in the economy. Previous literature holds the view that active funds cannot deliver outperformance to investors, despite charging high fees. Aggregate welfare might be higher in a counterfactual world without active funds. In this paper, I show that some active funds (the connected ones) can encourage firms to be more innovative. Since innovation can promote long-run economic growth. Without those connected *active* funds, the economy as a whole might be worse off.

1.1 Related Literature

This paper links to several strands of the literature. First, it is related to papers examining the impact of social capital on economic outcomes. Putnam (1993) finds that local government performance is positively associated with people's participation in public activities in Italy. Knack & Keefer (1997) use trust as a measure of social capital and find that social capital fosters economic growth in a cross-country regression. Similarly, La Porta *et al.* (1997) examine the effect of trust on the performance of large organizations in a cross-country setting and find a positive relationship. In the finance setting, Guiso, Sapienza, and Zingales (2004, 2008) document that trust affects stock market participation and international trade. In explaining the positive relationship between social capital and economic growth, Bloom *et al.* (2012) show that high social capital in an area increases decentralized decision-making within firms. To the best of my knowledge, this paper is the first to examine the role of innovation as a potential channel by which to explain how social capital affects aggregate economic growth. And I provide identification for this mechanism.

Second, my paper links to the literature on managerial short-termism. Theoretical papers argue that managers are biased toward short-term gains because of reputation concerns (Narayanan 1985, Holmström 1999), takeover threats (Stein 1988), and concerns about stock price (Stein 1989).

Empirical paper finds that the majority of financial executives would give up positive net present value (NPV) projects to avoid missing their quarterly earnings target (Graham *et al.* 2005). Also, some papers call into question the existence of "short-termism" (Kaplan 2017). I contribute to this literature by showing that social capital, as measured by connected holdings, could alleviate short-term capital market pressures, increase management job security, and lengthen firms' planning horizon.

Third, this paper also contributes to a voluminous literature that explores the impact of financial market on corporate innovation. Bernstein (2015) investigates the impact of going public on innovation. Seru (2014) examines the effects of mergers and acquisitions on innovation. He & Tian (2013) show that financial analyst coverage causes firms to innovate less. Brav *et al.* (2017) demonstrate how hedge fund activism reshapes innovation. My contribution to this literature is to show how the social networks of the board members could influence innovation. Or, more broadly, how important is external investor's trust to innovation?

Lastly, this paper links to a growing literature investigating the real impact of institutional investors on corporate policies, such as leverage (Michaely *et al.* 2014), dividends (Grinstein & Michaely 2005), R&D (Bushee 1998, Aghion *et al.* 2013), and governance (Appel *et al.* 2016). The closest paper to mine is Aghion *et al.* (2013). The authors study the total impact of institutional investors on innovation outcomes. I deepen the understanding of institutional investors' role on innovation by showing that among all institutional ownership, connected mutual fund ownership matters *a lot* for innovation. And surprisingly, for some innovation measures, after accounting for connected holdings, institutional ownership is not statistically significant anymore, which indicates that connected holdings might be the part of the institutional ownership that drive innovation.

The rest of the paper is organized as follows: Section 2 describes the data and the measures. Section 3 presents the identification strategy. Section 4 shows the results. Section 5 discusses the potential mechanisms. Section 6 concludes.

2 Data and Variable Constructions

I use several sources to collect data on patents, mutual fund holdings, the educational background of corporate officers and mutual fund managers, and firm-specific and fund-specific characteristics.

2.1 Patent Data

2.1.1 Innovation Measures

To measure firms' innovative activities, I construct three measures. The first measure is the firm's total number of patent applications filed in a given year that are eventually granted. There are two important dates for a patent, the filing date and the issue date. The filing date issued by the USPTO on a patent application is of critical importance to an inventor or patent owner because it determines who has priority over the right to file a patent application for an invention. Once the original or amended claims of a patent application have been approved by the examiner, the patent is granted and the specific date of grant is called the issue date. Here, I use the patent's filing year because it allows me to capture the actual time of innovation.[4]

However, not all patents are of equal importance. Simple patent counts cannot distinguish between groundbreaking innovation and incremental discoveries (e.g., Griliches 1990). Hence, I construct the second measure, which is a firm's total number of citations for patents filed in a given year. For example, suppose IBM filed for 10 patents in 1990. Then I track all the citations that those 10 patents received in subsequent years (until the end of the sample) and aggregate those citations together to obtain IBM's total number of citations for 1990.

The above two measures mainly utilize the patent data. The last measure I used is based on Kogan *et al.* (2017) (KPSS, hereafter). Kogan *et al.* (2017) propose a new measure of the private, economic value of new innovations that is based on the stock market's reaction to patent grants. The basic idea is to first compute the abnormal stock returns that can be attributed to patent issuance and then multiply the return with the market cap to get the value of the patent. This measure can assess the importance of each patent but also label each patent with its economic value. This measure helps me better investigating more quantitative questions related to innovation, such as how much value is created by innovative activities for a given firm. Another advantage of this measure is that, because of the forward-looking feature of the stock market's reaction, I don't need a long period of time to accumulate citations to access the importance of the patent. (For more details on how Kogan *et al.* (2017) construct their measure, please refer to their paper.)

2.1.2 Patent Data Source

I downloaded the patent data from Professor Noah Stoffman's Web site.[5] A detailed description of this dataset can be found in Kogan *et al.* (2017). Kogan *et al.* (2017) begin with all patents downloaded from Google Patents. They matched patents to Center for Research in Security Prices (CRSP) firms by the assignee's name. In their final dataset, they have 1,928,123 matched patents, of which 523,301 (27%) are new compared with the commonly used NBER patent dataset (see Hall *et al.* 2001 for more details).

The patent data contain truncation problems that arises because patents appear in the database only after they are granted. There is a lag between the patent filing date and the patent issuing date. We can see from Table A2 that the mean lag days in the 1980s and in the 1990s is approximately

[4] For the reasons to use the filing date, see Griliches *et al.* 1986.
[5] https://iu.app.box.com/v/patents

2 years. But it significantly increases to more than 3 years in the 2000s with the 99th percentile reaching almost 8 years. In Figure A1, I plot the total number of patent applications that are eventually granted by year. We can see a significant drop starting from year 2003. This drop is due to the lag days: a lot of the patents applied for after 2003 are still under review and had not been granted by 2010, which is the last year of the patent data. To deal with this problem, I choose 2003 as the last year of my sample. This will guarantee that 99% of patents applied for before 2003 (including 2003) have been issued before 2010. I also check the growth rate of the total number of patent applications, and find that it is relatively stable from 1980 to 2003. Starting from 2003, it significantly decreases. For robustness, in the result section, I also rerun all the exercises with 2004 or 2005 as the last year of the sample. All the results are quantitatively the same. The truncation problem for the patent citations arises as patents tend to receive citations over an extended period of time (e.g., 20 years), but I only observe the citations up to 2010. Then patents issued near the end of the dataset have no time to accumulate citations. To deal with this problem, I follow Hall *et al.* (2001) and Hall *et al.* (2005). I estimate the citation-lag distribution for different technology categories. For patents that haven't received 20 years of citations, I use a citation-lag distribution to infer the total number of citations that those patents should receive.[6] Other methods can be used to address the citations truncation problem. For example, one can only count the total number of citations a patent received in the first N (N = 0, 1, 2, 3) years after it is granted. For robustness, I also check this measure, and all the results are qualitatively the same.

Kogan *et al.* (2017) measure of the value of patent is based on the stock market's reaction on the patent *issue* date. As discussed before, the patent *filing* date is more informative about the actual time the innovation occurred. So I move the KPSS measure to the patent application year.

Following the innovation literature, I set the patent and citation counts to zero for firms without available patent or citation information from the patent database. The distribution for both patent applications and patent citations is right skewed, with its median at zero. To deal with skewness, first I winsorize those variables at the 99^{th} percentile, and, second, I take the natural logarithm of both patent applications ($LnPatApp$) and patent citations ($LnPatCite$). This creates another problem: a lot of the observations with zero patents become missing values after taking the natural logarithm. I take two approaches: (1) I drop the observations with missing values (Aghion *et al.* 2013 also takes this approach), and (2) to avoid losing observations, I add one to the actual values before calculating the natural logarithm. I denote those cases as ($Ln\widetilde{PatApp}$) and ($Ln\widetilde{PatCite}$). I also apply the above procedures to the KPSS measure of the patent value. I performed a similar transformation, which I denote as $LnKPSS$ and $Ln\widetilde{KPSS}$.

[6]For more details, see Appendix G, where I provide the citation-lag distribution and adjustment methods.

2.2 Mutual Fund Holdings and Fund Manager Data

My data on mutual fund holdings come from the Thomson Reuters CDA/Spectrum S12 database, which includes all registered mutual funds filing with the SEC. I only include common stock holdings of mutual funds (i.e., a stock with the share code 10 or 11 in CRSP).

I obtain portfolio managers' biographical information from Morningstar, Inc. For each mutual fund manager, Morningstar provides the manager's name, all college and graduate degrees he or she received, the year in which the degrees were granted, and the granting institution. Morningstar also provides the employment history for each mutual fund manager, including the fund name, starting date, and end date. I first merge Morningstar dataset with CRSP mutual fund dataset, then I use MFLINKS data link provided by Wharton Research Data Services to merge it with Thomson Reuters fund holdings dataset (see Wermers 2000 for details on how to merge these two databases). My final mutual fund sample includes survivorship-bias-free data on holdings and biographical information for 3,094 U.S. mutual funds and 5,369 mutual fund managers between January 1980 and December 2003.[7]

[INSERT FIGURE A3 HERE]

2.3 Company Officers Data

The senior biographical information of company officers (defined as Chief Executive Officer [CEO], Chief Financial Officer [CFO], Chief Technological Officer [CTO], Chief Operating Officer [COO], and Chairman) and board of directors was obtained from BoardEx of Management Diagnostic Limited, a private research company specializing in collecting and disseminating social network data on company officials in U.S. and European public and private companies and other types of organizations (e.g., charity). For each senior company officer and board of directors, the Boardex database provides all the college and graduate degrees received, the year in which the degrees were granted, and the granting institution. Boardex also provides the employment history for each company official. This information includes his or her current and past roles with a start date and end date, as well as a dummy indicating whether the individual serves (served) on the board of directors in the current (past) employment position. Given this paper's focus, I restrict the sample to U.S. public firms.

[7]Notice that I keep all the matched funds instead of only active domestic equity mutual funds because I assume all kinds of funds' holdings could potentially affect firm's innovation. Out of those 3,094 mutual funds, 2,037 are domestic, well-diversified equity mutual funds. The domestic, well-diversified equity mutual funds compose about 90% of all the Asset Under Management (AUM) in the final dataset. The time series of total AUM of domestic active funds as a fraction of all funds can be found in Figure A3.

2.4 Connections and Connected Holdings

Following Cohen *et al.* (2008), I define the social networks over educational institutions. I group the degrees for mutual fund managers, senior company officers, and board of directors into six categories: (1) undergraduate, (2) Master of Business Administration (MBA), (3) general graduate (MA or MS), (4) doctorate (PhD), (5) law school (JD), and (6) medical school (MD).[8] The *broad connection* dummy equals 1 if the fund manager and a senior officer of the firm and/or a board member of the firm attended the same university. The *narrow connection* dummy equals 1 if the fund manager and a senior officer of the firm and/or a board member of the firm attended the same university at the same time and received the same type of degree (from the above-defined six groups of degrees). In terms of connectedness, narrow connections should dominate broad connections.

After I define the connection, connected holdings for a firm in a given year is defined by the following formula:

$$\text{ConHold}_{i,t}^{k} = \frac{\sum_{j=1}^{N_t} \text{CONNECTED}_{i,j,t}^{k} \cdot S_{i,j,t}}{\bar{S}_{i,t}} \qquad k = narrow, broad, \tag{1}$$

where $\text{CONNECTED}_{i,j,t}^{k}$ is a dummy variable that indicates whether fund manager j is connected to firm i through a type k connection in year t. $S_{i,j,t}$ is the shares held by the mutual fund manager j in firm i at time t. $\bar{S}_{i,t}$ is the total shares outstanding for firm i at time t. N_t is the total number of mutual fund managers at time t. $\text{ConHold}_{i,t}^{k}$ measures the fraction of a firm's equity that is held by connected mutual fund managers.

2.5 Other Data and Control Variables

Following the literature, I control for a set of variables that may affect a firm's innovation. To calculate the control variables, I collect financial statement items from Compustat, institutional holdings data from Thomsons CDA/Spectrum database (Form 13F), stock price information from CRSP, and institutional investor classification data from Brian Bushee's Web site.[9] The control variables include firm age, firm sale, ROA, asset tangibility, R&D stock, R&D investment intensity, leverage, Tobin's Q, investment intensity, institutional investor's ownership, dedicated institutional holdings, transitory institutional holdings, and quasi-indexed institutional holdings. I provide detailed variable definitions in Table A1.

[INSERT TABLE A1 HERE]

[8] I correct the inconsistencies in the educational institution names. The appendix provides the details.

[9] http://acct3.wharton.upenn.edu/faculty/bushee

2.6 Summary Statistics

Table A3 reports the number of connections for the top 15 educational institutes according to the definition of CONNECTED(broad) and CONNECTED(narrow), respectively. Harvard University and the University of Pennsylvania have the most connections under both definitions.

[INSERT TABLE A3 HERE]

[INSERT TABLE A4 HERE]

To mitigate the impact of outliers, I winsorize all variables at the 1st and 99th percentiles. Table A5, Panel A, reports the summary statistics for the full sample. Panel B separately shows statistics for those with positive CONNECTED(narrow) holdings and those with no CONNECTED(narrow) holdings. A salient feature in the data is that the firm-year observations with positive CONNECTED(narrow) holdings have significantly higher innovation outcomes as measured by the total number of patent applications (which were eventually approved), the total patent citations, and the KPSS value. Firms with positive CONNECTED(narrow) holdings also have a higher ROA, higher R&D stocks, and a higher Tobin's Q. But across the measures regarding routine investment, such as PPE or CAPX, the firms with positive CONNECTED(narrow) holdings have lower values.

The previous literature (Bushee 1998; Aghion *et al.* 2013) documents that among all different types of institutional investors, *dedicated* investors are important in terms of encouraging R&D investment and fostering innovation. The ratio between ConHold(broad) and dedicated institutional holdings in the sub-sample with positive CONNECTED(narrow) holdings is about 62.8%. And the ratio between ConHold(narrow) and dedicated institutional holdings is about 6.7%. So connected holdings constitute a sizable fraction of the dedicated holdings. In the following sections, without special notice, all the connected holdings are ConHold(narrow).

[INSERT TABLE A5 HERE]

3 Empirical Strategy

3.1 Specification

To explore the relationship between connected holdings and innovation, my baseline specification is as follows:

$$y_{i,t} = \alpha + \beta ConHold_{i,t} + X'_{i,t}\gamma + \eta_i + \mu_t + U_{i,j,t} + \epsilon_{i,t}, \tag{2}$$

where the indices i, j, and t correspond to the firm, the university, and the year, respectively. $y_{i,t}$ is various measures of innovation for firm i in year t. $ConHold_{i,t}$ is the fraction of the firm's equity held by connected mutual funds as defined in equation (1). $X'_{i,t}$ is a vector of firm and industry control variables that may affect a firm's innovation output. η_i is firm fixed effects. μ_t is the time fixed effects, and $U_{i,j,t}$ is a vector of dummies for each university in my sample. It equals 1 if one of the board members from firm i graduated from university j at time t. I cluster standard errors at the firm level. The appendix examines the other specifications, such as the Poisson model.

3.2 Identification

To establish the causal link between connected holdings (the proxy for social capital of the firm) and innovation is challenging. There are two selection issues. The first one is the endogenously chosen level of connected holdings. Connected mutual funds might select firms in which to invest based on the information observable to them, but not to me, the econometrician. An omitted factor could be the quality of inventors at the firm, the long-term innovation plan, or the prototype of some unpatented inventions. Such factors clearly affect innovation and are correlated with connected holdings. Apparently, constructing direct measures of those factors is difficult. But without the appropriate control, my results will be biased.

The second selection issue is that the educational network could be endogenous. People often self-select into groups (Manski 1993; Angrist & Pischke 2008). In my setting, mutual fund managers and corporate officers might choose to attend a university based on characteristics that are unobservable to the econometrician. For example, the average risk appetite of Harvard graduates might be different from that of Stanford graduates. And those unobserved characteristics might affect the connected holdings and the innovation and, in doing so, would cause omitted variable bias.

To tackle the first selection issue, I use an instrumental variable (IV) method. To address the second issue, I include university fixed effects as a control. I will first discuss the IV method, and in the exclusion section of my IV method, I discuss the importance of using university fixed effects and how they address the second selection issue.

My IV is the total number of connected mutual fund managers to the firm (including mutual funds that do not hold shares in the firm). More specifically, it is defined as follows: in each year, for each firm in my sample, I count the number of mutual fund managers who are "active" (i.e., working) and connected to the firm through an educational link. In the following two subsections, I discuss the necessary assumptions that need to hold for the instrument to be valid.

3.2.1 Relevance Condition

For the instrument to be valid, it must strongly affect connected holdings. This point has been partially established in Cohen *et al.* (2008). There, the authors show that mutual fund managers place larger bets on connected firms. Everything else equal, a firm with more connected mutual fund managers should have higher connected holdings. To formally test the relevance condition, I estimate the following first-stage regression:

$$ConHold_{i,t} = \delta z_{i,t} + X'_{i,t}\gamma + \eta_i + \mu_t + U_{i,j,t} + \epsilon_{i,t}, \tag{3}$$

where $z_{i,t}$ stands for the number of mutual fund managers connected to firm i at time t. For the IV to be valid, we need $\delta \neq 0$. In Table 1, I show the first-stage results. I find that the coefficient of the number of connected mutual fund managers is positive and significant at the 1% level. The F-statistic equals 48.44 and exceeds the threshold of $F = 10$, suggesting that the instrument is strong (Stock & Yogo 2005).[10] Using results from column 2, I show that an increase of one standard deviation in the number of connected mutual fund managers translates into an about 0.28-standard-deviation increase in CONNECTED(narrow) holdings. For robustness, I also check the first-stage regression for CONNECTED(broad). The results are similar.

[INSERT TABLE 1 HERE]

3.2.2 Exclusion Condition

The instrument not only needs to affect connected holdings, but importantly, it must satisfy the exclusion restriction. That is, it should be uncorrelated with the residual in equation (2). In other words, the instrument should influence the outcomes of firm-level innovation only through its effect on connected holdings. To validate the exclusion condition, I check the following dimensions:

3.2.2.1 Endogenous Choice of University Does Not Affect Time-Series Change in IV

The network used in this paper is based on the institutions that corporate officers and mutual fund managers both attended. There is a concern that agents (CEO or mutual fund managers) endogenously choose to attend a university based on unobserved characteristics. And those unobserved characteristics will be correlated with my instrumental variables, subsequently biasing my results (omitted variable bias). In this section, I analyze this possibility and show how to use university fixed effects to control for unobserved characteristics.

In my setting, my IV is the total number of mutual fund managers who are connected to the firm's board members. A connection is defined as people who attend the same university at the same time to obtain the same type of degree. Along three dimensions that determine connections,

[10] Here, since I assume that the error term ϵ is allowed to be correlated within cluster, the standard Cragg-Donald-based weak instrument test is no longer valid. Instead, I report the corresponding robust Kleibergen-Paap-Wald rank F-statistic. For more details, refer to Kleibergen & Paap (2006).

university choice is the one most likely to be endogenous.[11] From now on, I will focus on the endogeneity problem induced by the agent's choice of the university.

Assume the true data-generating process for firm-level innovation is as follows:

$$y_{i,t} = \alpha + \beta ConHold_{i,t} + X'_{i,t}\gamma + \eta_i + \mu_t + \underbrace{\overrightarrow{\lambda}_{i,t} + \epsilon_{i,t}}_{u_{i,t}},$$

where all the other variables share the same definition used in Section 3.1. $\overrightarrow{\lambda}_{i,t}$ is a vector that includes all the individual fixed effects for the board members for firm i at time t. In the simplest case, suppose there is only one CEO on the board. Then $\overrightarrow{\lambda}_{i,t}$ is equal to the CEO's individual fixed effects. Notice that if, throughout the sample, the firm has only that CEO, then the firm fixed effects and $\overrightarrow{\lambda}_{i,t}$ cannot be separately identified, because there is no time variation in either variable at the firm level. But if the firm replaces their CEO at some time, then $\overrightarrow{\lambda}_{i,t}$ captures the CEO's fixed effects at the corresponding periods. We can generate it to multiple person's case. $\epsilon_{i,t}$ is an error term that is uncorrelated with any regressors. I denote $\overrightarrow{\lambda}_{i,t} + \epsilon_{i,t}$ as the composite error $u_{i,t}$ that could be correlated with my IV.

The instrumental variable I used is the total number of mutual fund managers who are connected to firm i at time t. To make the IV invalid, we should have

$$\begin{aligned} E[z_{i,t} \cdot u_{i,t}] &= E[z_{i,t} \cdot (\overrightarrow{\lambda}_{i,t} + \epsilon_{i,t})] \\ &= E[z_{i,t} \cdot \overrightarrow{\lambda}_{i,t}] \\ &\neq 0, \end{aligned}$$

where the second equality comes from the fact that $\epsilon_{i,t}$ is an error term that is uncorrelated with any regressors. Because of the way $z_{i,t}$ is constructed, the only possible way $\overrightarrow{\lambda}_{i,t}$ is correlated with $z_{i,t}$ must be the endogenous choice of universities of the board members based on $\overrightarrow{\lambda}_{i,t}$. But $\overrightarrow{\lambda}_{i,t}$, at most, should be correlated with $\overline{z}_{i,t}$, which is the sum of the average network size for each network of the board member. In other words, the manager's individual fixed effects might affect his or her choice to attend Harvard University versus Stanford University. This could affect his or her *average* network size . But these individual fixed effects should be uncorrelated with the time-series variation of the network size. In other words, the CEO's individual fixed effects could not affect how many mutual fund managers are currently working in the mutual fund industry at any period in time. I should find that $(z_{i,t} - \overline{z}_{i,t})$ is uncorrelated with $\overrightarrow{\lambda}_{i,t}$. The proof is as follows: given $E[(z_{i,t} - \overline{z}_{i,t})|\overrightarrow{\lambda}_{i,t}] = 0$ because $(\overline{z}_{i,t})$ is the mean of $z_{i,t}$.

[11] I admit that the choice to attend university in a specific year also might be endogenous. But the time fixed effects should be able to control for this concern.

$$
\begin{aligned}
E[(z_{i,t}-\overline{z}_{i,t})|\overrightarrow{\lambda}_{i,t}] &= E[E[(z_{i,t}-\overline{z}_{i,t})\cdot\overrightarrow{\lambda}_{i,t}|\overrightarrow{\lambda}_{i,t}]] \\
&= E[\overrightarrow{\lambda}_{i,t}E[(z_{i,t}-\overline{z}_{i,t})\cdot\overrightarrow{\lambda}_{i,t}]] \\
&= E[\overrightarrow{\lambda}_{i,t}\cdot 0] \\
&= 0
\end{aligned}
$$

This shows that the demeaned instrument is uncorrelated with the individual unobserved characteristics. The university fixed effects could capture this demeaning effect. I create a full set of dummies for each university and set it to 1 if any company officer receives a degree from this university.

3.2.2.2 Board Members Cannot Affect Mutual Fund Manager's Career

In the last section, I argue that the board members' individual fixed effects will not affect the time-series change in the number of connected mutual fund managers after controlling for the university fixed effects. But if board members could affect the career path of the mutual fund manager, then this point could invalidate my IV. For example, board members could lengthen the tenure of connected mutual fund managers by feeding them insider information. I check this story by investigating the relationship between fund manager's termination and their connectedness. In the appendix, I provide the results. There is no statistically significant relationship between fund manager's connectedness (defined as the number of board member connections) and their termination (defined as not working in the current fund in the next year). The potential reason is that, usually, fund managers don't hold a lot of equity from connected firms because of the diversification restriction imposed by the fund family. So even if they could profit from connected holdings, those outperformances only marginally contribute to their total performances.

3.2.2.3 Dynamic Formation of Network Might Not Exist

Agents might form social ties in anticipation of future economic benefits (Manski 1993). If the links are dynamically formed (Jackson 2010), then there could be an endogenity problem. For example, assume the mutual fund manager A wants to know more information about firm F. And fund manager A endogenously formed a link with firm F through CEO B. If this sort of thing happens in the data, then the IV (the number of connected mutual fund managers) could be possibly driven by the profitability of the firm, a finding that could be correlated with both the IV and the innovation. But according to my network definition, this scenario is not possible because whether A and B attended the same school for the same degree at the same year is determined many years in the *past*. To investigate this point, I check the age distribution for the company officers and the mutual fund managers. In my data, the median age for the fund manager and company officers is

45 and 54, respectively. Most likely, they have passed the age for education. This evidence rules out the story that maybe a mutual fund manager or a company officer would endogenously choose their education to form the link.

[INSERT TABLE A6 HERE]

4 Result

In this section, I examine the impacts of connected holdings on firm innovation using the instrumental variable approach described in Section 3.2. Throughout all specifications, I control for firm, year, and university fixed effects. I also control for the following variables: log sales, firm age, asset tangibility, leverage, Tobin's Q, ROA, and investment rate. Table A1 provides the variable definitions.

4.1 Innovation Output

In this section, I examine the relationship between connected holdings and various measures of innovation outcomes. In brief, I use three sets of dependent variables to measure the outcome of innovation: (1) the total granted patents for a firm in a given year; (2) the total number of citations received by those patents; and (3) the value created by those patents for the firm as measured by $KPSS$.[12]

Table 2 reports the results. I find a statistically significant positive relationship between connected holdings and various measures of innovation output. With more connected holdings, firms produce more patents and these patents receive more citations. Meanwhile, more firm value is created by these patents. In terms of economic significance, a one-standard-deviation increase in connected holdings causes a 10.61% (= $(42.46 \times 0.0025 \times 100)\%$) increase in the number of patents generated by the firm, a 14.86% (= $(59.47 \times 0.0025 \times 100)\%$) increase in the number of citations, and a 2.21% increase in the firm value created by the patents. To put those results into perspective, a 10.61% increase in the number of patents for the average firm is equal to 6 more patents. And a 2.21% increase in firm value for the average firm is equal to 23 millions dollars (= $2.21\% \times 1,089$).

[INSERT TABLE 2 HERE]

[12] For detailed variable definitions, see Table A1.

4.1.1 Innovative Industries versus the Rest

Innovations do not equally spread across industries.[13] Connected holdings could potentially have heterogeneous impacts on firms in different industries. I split my sample into two groups by innovativeness. To begin with, I use Fama-French 12 industry categorization. For the high innovation group, I pick the Business Equipment industry, which covers Computers, Software, and Electronic Equipment. I also pick the health care industry because this industry traditionally puts more emphasis into inventions (think about all the new drugs). I put the remaining industries into the low innovation group. I find that for the high innovation group, the impact of connected holdings is much more pronounced. For patent applications, the coefficient doubled, and for patent citations, the coefficient tripled. This means that the connected holdings are much more important for the innovation process in the industry in which innovation is important. One way to interpret this result is that the firms in the high innovation group might be more constrained. When they have more support, they are more free to innovate. But for the firms in the low innovation group, the incentive to innovate is low. So even if the firms have more connected holdings, the production of innovation still could be low. In an untabulated table, I repeat this above analysis using the Fama-French 30-industry portfolio, and the results are quantitatively similar.

[INSERT TABLE 3 HERE]

4.1.2 Active Funds versus Closet Indexers

In the previous sections, I examine the impact of *total* connected holdings on innovation. However, there exists significant heterogeneity across the fund types, that is, active funds versus passive funds. Active funds can influence firms' behaviors through their actions in the public equity market, for instance, investment and divestment (Edmans (2009)). Passive funds don't have the divest option. They mainly influence firms through the "voice" channel (Levit & Malenko (2011), Appel *et al.* (2016)).

In terms of passive funds, there are mainly two types: (1) "true" passive funds (their goals are to replicate some existing indexes, e.g., S&P 500 Index Funds.) and (2) closet indexers (they don't claim they are index funds, but their investment styles are similar to those of index funds). I first check the presence of index funds in my sample by mainly focusing on the number and total asset management. Table A18 in the appendix provides the results. Because MFLINKS from WRDS

[13] According to Kogan *et al.* (2017), "developments in computing and telecommunication have brought about the latest wave of technological progress in the 1990s and 2000s, which coincides with the high values of our measure. In particular, it is argued that this is a period when innovations in telecommunications and computer networking spawned a vast computer hardware and software industry and revolutionized the way many industries operate. We find that firms that are main contributors to our measure belong to these sectors with firms such as Sun Microsystems, Oracle, EMC, Dell, Intel, IBM, AT&T, Cisco, Microsoft and Apple being the leaders of the pack."

mainly focus on matching active funds between CRSP and the Thomson Reuters dataset, there are very few index funds in my sample. In terms of fund numbers, the time-series average of passive funds is below 3%. In terms of total AUM, the time-series average of passive funds is below 2%. So within passive funds, I focus on closet indexers.

Currently, there is no universal way to detect closet indexers. Cremers & Petajisto (2009) argue that a fund with active shares between 20% and 60% are more likely to be a closet indexer.[14] I choose the cut-off as 50%; that is, I label the funds with active shares lower than 50% as closet indexers. The time-series average of the AUM of closet indexers versus the total AUM in the data is 7.27%.

I redo the analysis regarding innovation by including both active funds' connected holdings and closet indexers' connected holdings:

$$y_{i,t} = \alpha + \beta_1 ConHold^a_{i,t} + \beta_2 ConHold^c_{i,t} + X'_{i,t}\gamma + \eta_i + \mu_t + U_{i,j,t} + \epsilon_{i,t},$$

where $ConHold^a_{i,t}$ is the connected holdings for firm i in year t from active funds, and $ConHold^c_{i,t}$ is the connected holdings from closet indexers. All the other variable definitions are similar to those in the main specification in equation (2). The instrumental variable for $ConHold^a_{i,t}$ is the connected mutual fund managers for firm i in year t from active funds, and the instrumental variable for $ConHold^c_{i,t}$ is the connected mutual fund managers for firm i in year t from closet indexers.

Table 4 reports the results. I find that the impact of connected holdings mainly comes from active funds instead of closet indexers. This result is similar to that of Aghion *et al.* (2013), who find quasi-indexed institutional investors have no impact on firms innovation. This result indicates that, *quantitatively*, active funds exert more impacts on firm's innovation policy than passive funds.

[INSERT TABLE 4 HERE]

4.2 Innovation Input

Potentially, there could be two factors contributing to the increase in innovation outcomes following an increase in connected holdings. Connected holdings cause (1) an increase in innovation input (i.e., R&D expenditure increases) and (2) an increase in innovation efficiency (i.e., patents per R&D dollar). There are three sets of combined reasons to explain the increases in innovation outcomes: (1) an increase in input, but a decrease or no effect in efficiency; (2) an increase in efficiency, but a decrease or no effect in input; and (3) both an increase in efficiency and in input.

It has long been recognized that input into R&D matters for the production of patents (Hausman *et al.* 1984). Because of the skewness of R&D, I take the natural logarithm of R&D expenditure.

[14] Active shares are the shares of portfolio holdings of a fund that differ from benchmark index holdings.

There are two problems with the R&D data: (1) because of firms' incentives to disguise their behaviors, a lot of firms choose not to report R&D. In Compustat, they are missing numbers. I drop these observations.[15] And (2) for some observations, the value of R&D is zero. After calculating the natural logarithm, these values become missing values. I take two approaches: (1) I drop the observations with missing values (Aghion *et al.* 2013 also takes this approach), and (2) to avoid losing observations, I add one to the actual values before calculating the natural logarithm. I denote those cases as ($Ln\widetilde{R\&D}$).

In Table 5, I report the results. In columns (1) and (3), I report the baseline results. I find a statistically significant positive relation between connected holdings and firm's investment in R&D. In columns (2) and (4), after controlling for the university fixed effects, the coefficients are still significant and only marginally decrease in magnitude. The results are robust to using either $LnR\&D$ or $Ln\widetilde{R\&D}$. Since I use a log-level specification, the interpretation of the coefficient is semi-elasticity. A one-standard-deviation increase in connected holdings leads to a 12.61% (= $(50.46 \times 0.0025 \times 100)$%) increase in firms' R&D expenditure. This increase is economically sizable. To put the results in context, the average annual growth rate of R&D in my sample is 21.37%. So the increase in R&D for the firm that experiences a one-standard-deviation increase in connected holdings is about 59% of it.

To explore an interesting question about how persistent these impacts are, I relate R&D expenditure in the next 2 years (R&D in year $t+2$, not the cumulative R&D from year t to year $t+2$) to connected holdings. And I find a positive significant result. This means that the connected holdings have a long-term impact on the firm's R&D policy. And those long-term impacts align with the long-term project feature of the innovation process, which could lead to more innovation.

[INSERT TABLE 5 HERE]

4.3 Innovation Efficiency

In this section, I check for innovation efficiency. The null hypothesis is that connected holdings have no impact on innovation efficiency. There are two alternative hypotheses.

Hypothesis A1: Connected holdings encourage management to take on more risk. This does not mean that management takes on the risk in a discrete way. They could invest in a lot of projects, both promising and unpromising ones. Doing so leads to a decrease in innovation efficiency.

Hypothesis A2: Connected mutual fund managers provide useful information to management regarding innovation. Doing so could induce management to pick the right projects. In this case, innovation efficiency increases.

[15] In an unreported analysis, I replace all the missing values for R&D with zero, since my sample is post-FASB 1975 R&D reporting requirement. The results are quantitatively similar.

I construct three proxies for innovation efficiency: The first measure is the ratio between total patents and R&D expenditures. The second measure is the ratio between citations and R&D expenditures. The third measure is the ratio between KPSS and R&D expenditures.

Following Hirshleifer *et al.* (2013), I define innovation efficiency (IE) in the following way:

$$IE_{i,t} = \frac{Innovation_{i,t}}{(R\&D_{i,t} + 0.8 \times R\&D_{i,t-1} + 0.6 \times R\&D_{i,t-2} + 0.4 \times R\&D_{i,t-3} + 0.2 \times R\&D_{i,t-4})},$$

where the numerator is the total number of patents, the total number of citations, and the value created by patents (KPSS) for firm i in year t. The denominator is R&D capital, which is the 5-year cumulative R&D expenses, assuming an annual depreciation rate of 20% like in Chan *et al.* (2001) and Hirshleifer *et al.* (2013).

Table 6 reports the results. There is no statistically significant relationship between connected holdings and the first two measures of innovation efficiency: patents granted scaled by R&D capital and the total number of citations scaled by R&D capital. But there is a positive and statistically significant relationship between connected holdings and the value created by patents scaled by R&D capital. The results are mixed. To draw a conservative conclusion, connected holdings don't improve firms' innovation efficiency . So the observed increases in innovation outcomes are mainly due to more input into innovation.

[INSERT TABLE 6 HERE]

4.4 Patent Natures

Connected holdings might potentially affect the nature of the patent. I check the impact of connected holdings on the properties of the patents. Trajtenberg *et al.* (1997) develop the *originality* and *generality* measures based on the distribution of citations. A patent that cites diversified technological classes of patents is viewed as more original. On the other hand, a patent that is cited by diversified technologically classes is viewed as more general.

I find that the generality of patents increases when there are more connected holdings. This means that the patents produced by the firms are cited by patents coming from different technological classes. This evidence is consistent with the higher citation results.

[INSERT TABLE 7 HERE]

4.5 Real Outcomes

Numerous endogenous growth models imply that firm growth is related to innovation. Since results in the previous sections demonstrate that connected holdings cause an increase in innovation, here

I explore whether connected holdings can lead to firm growth. One important caveat is that innovation doesn't have to be the only channel through which connected holdings affect firm's growth. Connected holdings might affect the cost of capital of the firm or the firm's contracting environment. But it is still interesting to investigate whether, besides innovation, connected holdings affect firm's growth?

I focus on growth of (a) profits, (b) the nominal value of output, and (c) the number of employees. I estimate the following specification:

$$\log Y_{i,t+\tau} - \log Y_{i,t} = \alpha + \beta ConHold_{i,t} + X'_{i,t}\gamma + \eta_i + \mu_t + U_{i,j,t} + \epsilon_{i,t}, \tag{4}$$

where all the variables on the right-hand side are the same as those in the main specification equation (2). τ is the length of the horizon. Here, I explore $\tau = 1, 3, 5$.

Table 8 reports the results. I find that future firm growth is strongly related to connected holdings. This result is related to Lins *et al.* (2017). Lins *et al.* (2017) measure firm social capital by corporate social responsibility (CSR) intensity. They find that during a crisis, high-CSR firms experience higher profitability, growth, and sales per employee relative to low- CSR firms. Here, my finding shows that for a given firm that has more connected holdings, another measure of social capital, they also have higher profit growth, output growth, and employment growth. The results show that connected holdings are, in general, good for the firm.

[INSERT TABLE 8 HERE]

4.6 Robustness

4.6.1 Alternative Specification

The patent applications and patent citations for a firm in a certain year are count data, in other words, non-negative integers. To explain count data, linear models have shortcomings: the predicted value of a linear model might be negative, whereas the dependent variable is always positive. For count data, in the literature, people also use a Poisson regression model. My Poisson model specification is as follows:

$$E(y_{i,t}|\boldsymbol{x}_{i,t}) = exp(\beta ConHold_{i,t} + X'_{i,t}\gamma + \eta_i + \mu_t + \epsilon_{i,t}). \tag{5}$$

Different assumptions about the error term will generate alternative estimators even though equation (5) remains the same. A Poisson model assumes the mean equals the variance, but I also consider alternatives, such as a negative binomial regression.

Following Aghion *et al.* (2013), I implement the instrumental variable estimator by using the control function approach for the Poisson regression models. For details, please see the appendix.

Table A7 reports the results. The results are qualitatively similar. Connected holdings lead to more innovation as measured by patents granted and patent citations.

4.6.2 Only Focus on the Senior Management

For all the previous analysis, I include all board members. If I restrict my attention to only top management team (defined as CEO, CFO, COO, CTO, and Chairman), do the results hold? Table A9 reports the results. Comparing columns (1) to (3) with (4) to (6), it is clear that the results are quantitatively similar.

4.6.3 Dot-Com Bubble

I explore whether the results are mostly driven by the Internet bubble, since, in those years, the number of patents jumped. Meanwhile, it is the latter part of my sample, so there are more connected holdings (because more people in my data became CEO or a mutual fund manager). So all the results potentially could be driven by those years. Following the literature, I define the Dot-com bubble years as those from 1999 to 2003. I rerun some of the regressions in the two subperiods (1980 to 1998 and 1999 to 2003), and the results (in untabulated table) are both positive and statistically significant.

4.6.4 Management Ownership

Morck *et al.* (1988) demonstrate that management ownership affects the market valuation of the firm. Innovation, as an important constitute of firm value, might also be influenced by management ownership. Meanwhile, management ownership might also be correlated with connected holdings. To avoid omitted variable bias, I include management ownership as a control variable. I gather management ownership data from the Execucomp Annual Compensation dataset. I construct two measures for management ownership: (1) total ownership by all directors and (2) ownership by CEO. Ownership is defined as the ratio between the shares owned by management and the total shares outstanding.

I include the two additional control variables in the main regression with the total number of patents as the dependent variable. In untabulated tables, I find that the coefficients of connected holdings are still positively statistically significant. And management ownership is also positively statistically significant. These results show that management ownership also matters for innovation.

4.6.5 Are Connected Mutual Funds Local Investors?

Coval & Moskowitz (2001) document that local investors have an informational advantage in trading local firms. If all the connected holdings are also local holdings, then it is difficult to argue that it is the *connected* mutual fund managers, instead of the *local* mutual fund managers, who affect the firm's innovation .

To address this concern, I first compute the localness of the holdings. I obtain the headquarters' ZIP codes for mutual funds using the CRSP Mutual Fund dataset from WRDS. I obtain the corporate headquarters' ZIP codes from Compustat. I obtain the mapping between ZIP codes and the corresponding latitude and longitude information from the United States Postal Service (USPS). The location of headquarters is used as opposed to the state of incorporation for the simple reason that firms tend to incorporate in a state with favorable tax, bankruptcy, and takeover laws, rather than for operational reasons, and typically do not have the majority of their operations in their state of incorporation. I use 'Haversine' formula to calculate the great-circle distance between two points-that is, the shortest distance over the earth's surface.

$$d = 2R\arctan 2\left(\sqrt{a}, \sqrt{(1-a)}\right)$$
$$a = sin^2\left(\frac{lat_2 - lat_1}{360\pi}\right) + cos\left(\frac{lat_2}{180\pi}\right) cos\left(\frac{lat_1}{180\pi}\right) sin^2\left(\frac{lon_2 - lon_1}{360\pi}\right),$$

where d is the distance between two coordinates. R is the earth's radius, equals 6,371 km. lat stands for the latitude, whereas lon stands for the longitude. The subscripts $\{1,2\}$ stand for the places.

[INSERT FIGURE A4 HERE]

In the data, among all the connected holdings, the local holdings (i.e., a distance of less than 100 km between firms and funds) compose about 9.8%. The distant holdings compose the remaining 90.2%. Next, I repeat the main analysis with the distant connected holdings, and all the results are quantitatively similar.

5 Potential Mechanisms

The empirical findings thus far show that more connected holdings lead to more input in R&D and more innovation output, and those effects are stronger in the more innovative industries.

In this section, I discuss several potential explanations for my findings. Since social capital, as measured by connected holdings, might affect multiple dimensions of the firm, all potential stories may co-exist. I focus on three such explanations. First, the literature on corporate short-termism argues that short-term capital market pressure could be detrimental to innovation (Holmström (1989), Stein (1989), and Terry (2016)). I check the impacts of connected holdings on capital market pressures. Second, Manso (2011) predicts that higher job security enables management to take on more innovative projects. Firm exposure to a takeover could be a major job risk to management. I explore the relationship between connected holdings and takeover risks. Third, besides capital market and corporate control market, I investigate the behavior of connected mutual funds in the corporate governance market, especially voting behaviors. If connected funds are more likely to vote in favor of management, this could ease management's plan in pursuing innovative plans for the firm.

5.1 Capital Market Pressures

5.1.1 Hypotheses Development

Earning targets matter a lot to the firm. Each fiscal period, public firms must disclose their earnings. Before a disclosure, financial analysts forecast the earnings and the financial press widely reports a consensus forecast for a given firm. This consensus forecast (or earnings target) works as an externally set performance benchmark for the firm. Whether firms beat, maintain, or miss the earnings target matters a lot. Actually, around 90% of surveyed U.S. executives report pressure to meet short-term earnings targets (Graham *et al.* 2005). But why would management care so much about those earnings targets? Previous research (Kasznik & McNichols, 2002; Bartov *et al.* (2002)) demonstrates that missing quarterly earnings benchmarks leads to significantly lower abnormal returns (quarterly and annual). And because of management's stock-based compensation scheme, low stock returns have a significant impact on management's compensations.

How does meeting an earnings target matter for innovation? The previous literature shows that firms sacrifice their long-term investments, such as R&D, to meet the short-term earnings targets. It is because the benefits of R&D appear much later, whereas the costs show up in the current quarter through lower earnings. Almost half of surveyed U.S. executives report that they would prefer to reject a positive net present value project over missing their analyst target (Graham *et al.* 2005). The opportunistic cuts to meet short-term targets can be detrimental to innovation.

The role of institutional investors in creating short-termism is mixed. Some institutional investors might exuberate to the managerial myopia behavior because the institutional investors will dump a company's stock as soon as its earnings are not quite up to par (Stein 1988). Divestments encourage management to meet targets. However, I hypothesize that the connected mutual fund managers should behave differently. The social capital between the connected mutual fund managers and the firm's management should encourage the connected mutual fund managers to stay with the firm through adverse situations. So, in this section, I examine the investment behavior of connected mutual funds versus non-connected institutional investors following a missed earnings target.

5.1.2 Trading Behaviors

Utilizing the Institutional Brokers Estimates System (IBES) dataset, I document two new stylized facts: (1) when a firm misses its earnings expectations, connected mutual funds stay, whereas non-connected institutional investors divest. (2) Compared with firms with no connected holdings, firms with connected holding see their returns drop less.

My data on analysts' forecasts come from the IBES database. I focus on the quarterly earnings forecasts because it is widely studied in the finance and accounting literature.

Following the literature, I define an earnings surprise as:

$$ES_{i,t} = EPS_{i,t} - \text{median}\{EPS^f_{i,k,t}\},$$

where i indexes firm, t indexes time, and k indexes analyst. $EPS_{i,t}$ stands for the *realized* earnings per share. $EPS^f_{i,k,t}$ is the forecast made by analyst k for firm i quarter t's earning. median$\{EPS^f_{i,k,t}\}$ serves as the consensus forecast for all the analysts.

First, I examine the trading behaviors of connected mutual funds and non-connected funds. To investigate the trading behaviors of different types of investors, I follow a standard event-study methodology. I define trading as the change in the fraction of firm equity from quarter $t-1$ to quarter t. For connected holdings, it is

$$T^c_{i,t} = ConHold_{i,t} - ConHold_{i,t-1}.$$

For non-connected institutional holdings, it is

$$T^{nc}_{i,t} = nonConHold_{i,t} - nonConHold_{i,t-1}.$$

For the difference between connected holdings and non-connected institutional Holdings, it is

$$T^{diff}_{i,t} = (ConHold_{i,t} - nonConHold_{i,t}) - (ConHold_{i,t-1} - nonConHold_{i,t-1}).$$

My empirical specification is

$$y_{i,t} = MISS_{i,t-1} + \gamma X'_{i,t-1} + \eta_j + \mu_t + \epsilon_{i,t},$$

where y stands for trading by different types of investors and the difference between connected holdings and non-connected holdings. $MISS_{i,t-1}$ is a dummy variable that equals 1 if firms miss their quarterly earnings target (i.e., $ES_{i,t-1} < 0$). $X'_{i,t-1}$ is a vector of control variables that includes firm size, book-to-market ratio, and lagged 12 month returns. η_j stands for industry fixed effects. μ_t stands for time fixed effects. Standard errors are clustered at the firm level. Table 9 presents the results. I find that missing the earnings target triggers a sell off from non-connected institutional investors. The average magnitude is 35 basis points and is statistically significant. For the connected mutual funds, there is no evidence that their trading is affected by these events. And, as a natural result, for the difference between connected holdings and non-connected holdings, the gap becomes wider after firms miss their earning targets. The above evidence shows that after firms miss their earning targets, connected holdings stay and non-connected holdings divest. The results are robust when controlling for firm-level characteristics, as well as time and industry fixed effects.

[INSERT TABLE 9 HERE]

5.1.3 Return Behaviors

The previous section demonstrates that connected holdings are "loyal" to the firm when bad shocks are realized. In this section, I investigate whether connected holdings can actually alleviate the drop in stock price when firms miss their earnings target.

My empirical specification is as follows:

$$y_{i,t} = \alpha + MISS_{i,t} + AnyCon_{i,t} + MISS_{i,t} \times AnyCon_{i,t} + \epsilon_{i,t},$$

where y stands for the next month's abnormal returns estimated by the following models: CAPM, the Fama-French 3-factor model, the Fama-French-Carhart model, the Fama-French 5-factor model, and the characteristics- adjusted model (DGTW).

The variable of interest is the interaction term $MISS_{i,t} \times AnyCon_{i,t}$, which captures the cross-sectional difference between the firm with connected holdings versus the firm without connected holdings conditioned on firms experiencing a negative earnings surprise.

Table 10 presents the result. It is clear that across all kinds of measures of the abnormal return, the coefficient in front of the interaction term is greater than 0. This means that the firms' return drops less when they are invested with connected holdings. The evidence supports the argument that the presence of connected holdings could actually alleviate the price drop when firms miss the earnings target.

[INSERT TABLE 10 HERE]

5.2 Management Job Security and Takeover Exposure

In the last section, I show that missed earnings targets lead to lower subsequent stock returns, possibly leading to a reduction in management's compensation. Another risk associated with missed earnings targets is takeover risk. Stein (1988) argues that "Relatively patient stockholders may not be discouraged by a low earnings report; they may attribute it to a policy of long-term investment by the firm .... Impatient shareholders, on the other hand, may become very distressed by low earnings reports and may try to dump a stock as soon as such a report is issued.... managers will be more fearful of undervaluation and the accompanying possibility of rip-off by a raider".

Along the chain of reasoning, at its origin, the attitude of the institutional investor is very important. In the last section, I demonstrate that connected mutual funds are more patient with the firm. In this section, I examine whether relationships with connected holdings would lead to lower takeover risk.

I measure the probability of a firm being acquired (i.e., takeover exposure), following Cremers *et al.* (2008). I use the expected probability of a firm being acquired instead of actual takeovers because, according to Stein (1988), it is the threat (or likelihood) of a takeover that affects management's incentives to invest in innovation.

First, I estimate the firm's takeover exposure by running a logit model:

$$Target_{i,t+1} = \frac{1}{1+e^{-K_{i,t}}},$$

$$\begin{aligned}\text{where } K_{i,t} &= \alpha + \beta_1 Q_{i,t} + \beta_2 PPEAsset_{i,t} + \beta_3 LnCash_{i,t} + \beta_4 BLOCK_{i,t} + \beta_5 LnMV_{i,t} \\ &\quad + \beta_6 IndMA_{i,t} + \beta_7 Lev_{i,t} + \beta_8 ROA_{i,t} + Ind_j + Year_t + \epsilon_{i,t},\end{aligned}$$

where i indexes firm, j indexes industry, and t indexes time. The dependent variable $Target$ is a dummy variable equal to 1 if a firm is a target of an acquisition. All the independent variables are constructed following Cremers *et al.* (2008).[16] The acquisition data are obtained from the Securities Data Corporations (SDC) database.

After estimating the model using maximum likelihood, I calculate takeover exposure as the predicted probability of the model. Next, I use takeover exposure as the dependent variable in my main specification, like in equation (2). I find that connected holdings are statistically *negatively* associated with takeover exposure. A one-standard- deviation increase in connected holdings leads to a 0.09 ($= -0.8278 * 0.0025/0.0222$) standard deviation decrease in takeover exposure. This is economically sizable.

According to Stein (1988), if there is less takeover exposure, there is less incentive for management to sacrifice long-term investment for short-term gains. So my findings show that connected holdings could benefit innovation through reducing takeover risk channel.

[INSERT TABLE 11 HERE]

[16]For detailed variable definitions, see the footnote of table 1 in Cremers *et al.* (2008).

5.3 Connected Funds' Voting Behaviors

Voting is an important dimension of corporate governance. To continue exploring the idea that connected mutual funds are "supportive" of the board members, I examine connected mutual funds' voting behaviors. I find that, on average, connected funds are more likely to vote against shareholder-initiated proposals on various governance issues than are non-connected funds.

5.3.1 Empirical Test

In 2003, the Securities and Exchange Commission (SEC) introduced a new rule requiring mutual funds to report their votes on all shares held. I gather the data from the Institutional Shareholder Services' (ISS) Voting Analytics database. This database covers U.S. mutual fund voting records for all institutions filing the SEC N-PX form. It also contains company vote results. I focus on the period from 2005 to 2011.[17] Funds have the option of voting "for", "against", "abstain", "withhold", or "do not vote". To better assess voter preference, I only focus on the vote with either "for" or "against".[18] Among all the different types of proposals, I focus on four types of shareholder proposals: (1) requires independent board chairman; (2) requires a majority vote for the election of directors; (3) amends articles/bylaws/charter -- calls for special meetings; and (4) restores or provides for cumulative voting. I choose these proposals for two reasons: (1) the board's attitude is clear; for almost all above mentioned proposals, management recommend the shareholders to vote against it.[19] (2) Besides proposals on director election, auditor appointments and executive compensations, these four types of proposals are the most frequently voted proposals.[20] I merge the voting dataset with the CRSP mutual fund dataset to obtain the fund's size, expense ratio, turnover ratio, and fund family size. For each firm, I obtain the financial information from CRSP and Compustat. I also obtain institutional holding data from the Thompson Reuters 13F dataset. Lastly, in each year, I label the fund-firm pair connected or not according to the connection definition in Section 2.4. The final dataset contains 219,040 mutual fund votes that cover 164 public firms on 250 proposals. There are 8,289 different funds voted.

I explore the impact of connectedness between fund and firm on fund voting behavior. Following the literature, I employ a probit model. Table 12 reports the results. The dependent variable is a dummy that equals one if a mutual fund's vote is consistent with the board recommendation and

[17] I start with 2005 instead of 2003 because only after 2005 (inclusive) can I reliably merge ISS mutual fund voting data with the CRSP mutual fund database. The appendix provides the matching methods . I stop the sample at 2011 because 2010 is the last year in which innovation data are available, and I don't want to extend the voting analysis in a fashion that might not be relevant for my main (firm innovation) results.

[18] "For" and "against" votes compose 95.07% of all mutual fund votes.

[19] See the tables in the appendix for management recommendations for these four proposal types.

[20] For proposals on director election, please check Iliev & Lowry (2014), Matvos & Ostrovsky (2010), and Cai *et al.* (2009). For proposals on executive compensation, see Malenko & Shen (2016) and Butler & Gurun (2012).

zero otherwise. The variable of interest is $Connected(broad)$. It is a dummy that equals one if a mutual fund is connected to firm. For a detailed definition of the connection types, refer to Section 2.4. The control variables include ISS recommendation, firm size, book-to-market ratio, the previous year's total stock return, fund size, fund turnover ratio, fund expense ratio, and fund family size. In some specifications, I also include year and industry fixed effects. To avoid the incidental problem in estimating non-linear model, I use the Fama-French 12-industry classification instead of the 4-digit SIC industry. I cluster standard errors at the fund level.

[INSERT TABLE 12 HERE]

Column (1) reports the results for the case without a control. There is a significant positive relationship between connection and support to the management. In column (2), I control for both firm and fund characteristics, as well as industry and year fixed effects. The connection coefficient becomes larger and more statistically significant. In terms of magnitude, the marginal effect of $Connected(broad)$ is 6.8% (column (2)). This means that if a fund is connected with the firm, it is 6.8% more likely to vote with management on shareholder proposals. In the sample, only 48% of the votes support the management. Holding the marginal effects fixed, if all the funds are connected, the supporting rate increases by 20%.

In columns (3) to (6), I split the sample by specific proposals. The coefficient in front of $Connected(broad)$ is positive and statistically significant for the first three proposals: requires independent board chairman; requires a majority vote for the election of directors; and calls for special meetings. But it is not significant for the proposal restores or provides for cumulative voting.

In the appendix, I also report the results of the impact of connectedness on firm's "say-on-pay" proposals in 2011. I find that connected funds, on average, are more likely to vote in favor of management's proposals on compensation issues.

6 Conclusion

In this paper, I provide evidence that firm-level social capital, as measured by connected holdings affect innovation. In particular, I find that more connected holdings lead to more patents, higher patent impacts, and larger firm value as created by the patents. I also find that the above effects are more pronounced for firms in high-tech industries and for connected holdings coming from "truly" active funds. I also find that social capital improves innovation outcomes by encouraging investment in R&D. Connected holdings foster innovation by helping to reduce capital market pressures and to increase management job security.

This work has many exciting future directions. One would be to investigate how institutional investors affect other corporate policies, such as investments and capital structures utilizing the identification strategy developed here. Another would be to investigate the asset pricing implications of social capital. Third, the current setting could be used as a laboratory to examine the impacts of different kinds of networks (neighbor networks, club networks, religion networks, etc.) on corporate real outcomes.

References

Aghion, Philippe, Van Reenen, John, & Zingales, Luigi. 2013. Innovation and institutional ownership. *American Economic Review*, **103**(1), 277–304.

Angrist, Joshua D., & Pischke, Jörn-Steffen. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Appel, Ian R., Gormley, Todd A., & Keim, Donald B. 2016. Passive investors, not passive owners. *Journal of Financial Economics*, **121**(1), 111–141.

Bartov, Eli, Givoly, Dan, & Hayn, Carla. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics*, **33**(2), 173–204.

Bebchuk, Lucian Arye, & Stole, Lars A. 1993. Do short-term objectives lead to under-or overinvestment in long-term projects? *Journal of Finance*, **48**(2), 719–729.

Bernstein, Shai. 2015. Does going public affect innovation? *Journal of Finance*, **70**(4), 1365–1403.

Bertrand, Marianne, & Mullainathan, Sendhil. 2003. Enjoying the quiet life? Corporate governance and managerial preferences. *Journal of Political Economy*, **111**(5), 1043–1075.

Bloom, Nicholas, Sadun, Raffaella, & Van Reenen, John. 2012. The organization of firms across countries. *Quarterly Journal of Economics*, **127**(4), 1663–1705.

Brav, Alon, Jiang, Wei, Partnoy, Frank, & Thomas, Randall. 2008. Hedge fund activism, corporate governance, and firm performance. *Journal of Finance*, **63**(4), 1729–1775.

Brav, Alon, Jiang, Wei, Ma, Song, & Tian, Xuan. 2017. How does hedge fund activism reshape corporate innovation? *Journal of Financial Economics, Forthcoming*.

Bushee, Brian J. 1998. The influence of institutional investors on myopic R&D investment behavior. *Accounting Review*, 305–333.

Butler, Alexander W, & Gurun, Umit G. 2012. Educational networks, mutual fund voting patterns, and CEO compensation. *Review of Financial Studies*, hhs067.

Cai, Jie, Garner, Jacqueline L, & Walkling, Ralph A. 2009. Electing directors. *Journal of Finance*, **64**(5), 2389–2421.

Chan, Louis KC, Lakonishok, Josef, & Sougiannis, Theodore. 2001. The stock market valuation of research and development expenditures. *Journal of Finance*, **56**(6), 2431–2456.

Chevalier, Judith, & Ellison, Glenn. 1999. Career concerns of mutual fund managers. *Quarterly Journal of Economics*, **114**(2), 389–432.

Cohen, Lauren, Frazzini, Andrea, & Malloy, Christopher. 2008. The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy*, **116**(5), 951–979.

Coval, Joshua D, & Moskowitz, Tobias J. 2001. The geography of investment: Informed trading and asset prices. *Journal of Political Economy*, **109**(4), 811–841.

Cremers, KJ Martijn, & Petajisto, Antti. 2009. How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, **22**(9), 3329–3365.

Cremers, KJ Martijn, Nair, Vinay B, & John, Kose. 2008. Takeovers and the cross-section of returns. *Review of Financial Studies*, **22**(4), 1409–1445.

Edmans, Alex. 2009. Blockholder trading, market efficiency, and managerial myopia. *Journal of Finance*, **64**(6), 2481–2513.

Fisman, Raymond. 2001. Estimating the value of political connections. *American Economic Review*, **91**(4), 1095–1102.

Gennaioli, Nicola, Shleifer, Andrei, & Vishny, Robert. 2015. Money doctors. *Journal of Finance*, **70**(1), 91–114.

Graham, John R, Harvey, Campbell R, & Rajgopal, Shiva. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, **40**(1), 3–73.

Griliches, Zvi. 1990. Patent statistics as economic indicators: a survey.

Griliches, Zvi, Pakes, Ariel, & Hall, Bronwyn H. 1986. The value of patents as indicators of inventive activity.

Grinstein, Yaniv, & Michaely, Roni. 2005. Institutional holdings and payout policy. *Journal of Finance*, **60**(3), 1389–1426.

Guiso, Luigi, Sapienza, Paola, & Zingales, Luigi. 2004. The role of social capital in financial development. *American Economic Review*, **94**(3), 526–556.

Guiso, Luigi, Sapienza, Paola, & Zingales, Luigi. 2008. Trusting the stock market. *Journal of Finance*, **63**(6), 2557–2600.

Guiso, Luigi, Sapienza, Paola, & Zingales, Luigi. 2009. Cultural biases in economic exchange? *Quarterly Journal of Economics*, 1095–1131.

Hall, Bronwyn H, Jaffe, Adam B, & Trajtenberg, Manuel. 2001. The NBER patent citation data file: Lessons, insights and methodological tools.

Hall, Bronwyn H, Jaffe, Adam, & Trajtenberg, Manuel. 2005. Market value and patent citations. *RAND Journal of economics*, 16–38.

Haselmann, Rainer, Schoenherr, David, & Vig, Vikrant. 2017. Rent-Seeking in Elite Networks. *Journal of Political Economy*, forthcoming.

Hausman, Jerry A, Hall, Bronwyn H, & Griliches, Zvi. 1984. Econometric models for count data with an application to the patents-R&D relationship.

He, Jie Jack, & Tian, Xuan. 2013. The dark side of analyst coverage: The case of innovation. *Journal of Financial Economics*, **109**(3), 856–878.

Hirshleifer, David, Hsu, Po-Hsuan, & Li, Dongmei. 2013. Innovative efficiency and stock returns. *Journal of Financial Economics*, **107**(3), 632–654.

Holmström, Bengt. 1979. Moral hazard and observability. *Bell Journal of Economics*, 74–91.

Holmström, Bengt. 1989. Agency costs and innovation. *Journal of Economic Behavior & Organization*, **12**(3), 305–327.

Holmström, Bengt. 1999. Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, **66**(1), 169–182.

Iliev, Peter, & Lowry, Michelle. 2014. Are mutual funds active voters? *Review of Financial Studies*, **28**(2), 446–485.

Jackson, Matthew O. 2010. *Social and economic networks*. Princeton University Press.

Kaplan, Steven N. 2017. Are US companies too short-term oriented? Some thoughts.

Kasznik, Ron, & McNichols, Maureen F. 2002. Does meeting earnings expectations matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research*, **40**(3), 727–759.

Khwaja, Asim Ijaz, & Mian, Atif. 2005. Do lenders favor politically connected firms? Rent provision in an emerging financial market. *Quarterly Journal of Economics*, **120**(4), 1371–1411.

Kleibergen, Frank, & Paap, Richard. 2006. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, **133**(1), 97–126.

Knack, Stephen, & Keefer, Philip. 1997. Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics*, **112**(4), 1251–1288.

Kogan, Leonid, Papanikolaou, Dimitris, Seru, Amit, & Stoffman, Noah. 2017. Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics*, **132**(2), 665–712.

Levit, Doron, & Malenko, Nadya. 2011. Nonbinding voting for shareholder proposals. *Journal of Finance*, **66**(5), 1579–1614.

Lins, Karl V, Servaes, Henri, & Tamayo, Ane. 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *Journal of Finance*, **72**(4), 1785–1824.

Malenko, Nadya, & Shen, Yao. 2016. The role of proxy advisory firms: Evidence from a regression-discontinuity design. *Review of Financial Studies*, **29**(12), 3394–3427.

Manski, Charles F. 1993. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, **60**(3), 531–542.

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

Matvos, Gregor, & Ostrovsky, Michael. 2010. Heterogeneity and peer effects in mutual fund proxy voting. *Journal of Financial Economics*, **98**(1), 90–112.

Michaely, Roni, Popadak, Jillian A, & Vincent, Christopher J. 2014. The deleveraging of US firms and institutional investors' role.

Morck, Randall, Shleifer, Andrei, & Vishny, Robert W. 1988. Management ownership and market valuation: An empirical analysis. *Journal of Financial Economics*, **20**, 293–315.

Narayanan, MP. 1985. Managerial incentives for short-term results. *Journal of Finance*, **40**(5), 1469–1484.

Olson, Mancur. 1982. The rise and decline of nations: Economic growth, stagnation, and social rigidities.

Porta, Rafael La, Lopez-De-Silane, Florencio, Shleifer, Andrei, & Vishny, Robert W. 1997. Trust in large organizations. *American Economic Review Papers and Proceedings*, **87**, 333–338.

Putnam, Robert D, Leonardi, Robert, & Nanetti, Raffaella Y. 1993. *Making democracy work: Civic traditions in modern Italy*. Princeton University Press.

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

Shleifer, Andrei, & Vishny, Robert W. 1989. Management entrenchment: The case of manager-specific investments. *Journal of Financial Economics*, **25**(1), 123–139.

Shleifer, Andrei, & Vishny, Robert W. 1990. Equilibrium short horizons of investors and firms. *American Economic Review*, **80**(2), 148–153.

Stein, Jeremy C. 1988. Takeover threats and managerial myopia. *Journal of Political Economy*, **96**(1), 61–80.

Stein, Jeremy C. 1989. Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *Quarterly Journal of Economics*, **104**(4), 655–669.

Stock, James H, & Yogo, Motohiro. 2005. Testing for weak instruments in linear IV regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, edited by Donald W.K. Andrews and James H. Stock, Cambridge: Cambridge University Press*.

Terry, Stephen J. 2016. The macro impact of short-termism. *Working paper*.

Trajtenberg, Manuel, Henderson, Rebecca, & Jaffe, Adam. 1997. University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, **5**(1), 19–50.

Wermers, Russ. 2000. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *Journal of Finance*, **55**(4), 1655–1703.

Table 1: First Stage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Connection type	narrow	narrow	narrow	narrow	broad	broad	broad	broad
Sample	Full	Full	Full	Pre-1998	Full	Full	Full	Pre-1998
IV	18.95***	19.00***	17.79***	19.06***	6.34***	6.57***	7.49***	6.43***
	(2.72e-5)	(2.71e-5)	(2.12e-5)	(3.83e-5)	(3.77e-6)	(3.91e-6)	(3.20e-6)	(4.97e-6)
Firm FE	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes
University FE	yes	yes	no	yes	yes	yes	no	yes
Control Variables	yes	no	yes	yes	yes	no	yes	yes
Observation	110,663	112,241	110,663	85,489	110,663	112,241	110,663	85,489
R-squared	0.36	0.36	0.34	0.34	0.65	0.63	0.63	0.60
F-stat	48.44	49.18	70.84	24.76	282.14	282.04	548.81	223.60

This table reports the first-stage estimation of the instrumental variables analysis. The dependent variable is connected holdings. Two types of connected holdings are studied here. A narrow connection refers to board members and mutual fund managers who are connected by the same university, attended at the same time, received the same type of degree. A broad connection refers to board members and mutual fund managers who are connected by the same university. IV is the total number of connected mutual fund managers to firm i in year t according to the specific connection types. Section 2 defines all the variables. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. The F-stat is cluster adjusted. University FE are defined in Section 3.2.2. The estimated model is an ordinary least-squares (OLS) model with fixed effects. Robust standard errors were calculated and are provided in parentheses. Standard errors are clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. I multiply the coefficients for IV by 10^5 to make them easier to read.

Table 2: Connected Holdings and Innovation

	(1)	(2)	(3)
Dependent variable	$\widetilde{LnPatApp}$	$\widetilde{LnPatCite}$	$\widetilde{LnKPSS}$
$ConHold$	42.26***	59.47***	8.83***
	(12.83)	(22.60)	(2.10)
Year FE	yes	yes	yes
Firm FE	yes	yes	yes
University FE	yes	yes	yes
Control variables	yes	yes	yes
R-squared	0.84	0.77	0.59
Observations	110,663	110,663	110,663

This table reports the results of regressions of the innovation outcomes on connected holdings and other control variables. Table A1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 3.2.2. The estimated model is two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate that the coefficient is significant at the 1%, 5%, and 10% level, respectively.

Table 3: Disaggregating Industry: High-Tech Industry versus the Rest

	High-tech industry		Non-high-tech industry	
Dependent Variable	$\widetilde{LnPatApp}$	$\widetilde{LnPatCite}$	$\widetilde{LnPatApp}$	$\widetilde{LnPatCite}$
$ConHold$	101.55***	172.97**	28.29**	31.83
	(37.88)	(70.29)	(11.62)	(20.27)
Year FE	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes
University FE	yes	yes	yes	yes
Control variables	yes	yes	yes	yes
R-squared	0.79	0.71	0.86	0.80
Observations	35,534	35,534	75,126	75,126

This table reports regressions of the innovation outcome on connected holdings for high-tech industries and the remaining industries. In this table, I use the Fama-French 12-industry categorization. The high-tech industry contains: the Fama-French Industry 6 industry classification (Business Equipment -- Computers, Software, and Electronic Equipment) and the Fama-French 10-industry classification (Healthcare, Medical Equipment, and Drugs). Table A1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 3.2.2. The estimated model is a two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate that the coefficient is significant at the 1%, 5%, and 10% level, respectively.

Table 4: Disaggregating Fund Type: Active versus Closet Indexers

	(1)	(2)	(3)
Dependent variable	$\widetilde{LnPatApp}$	$\widetilde{LnPatCite}$	$\widetilde{LnKPSS}$
$ConHold$ of active funds	51.85***	77.10***	12.06***
	(16.53)	(28.56)	(3.09)
$ConHold$ of closet indexers	-9.97	-196.05	-94.43**
	(204.76)	(386.82)	(43.55)
Year FE	yes	yes	yes
Firm FE	yes	yes	yes
University FE	yes	yes	yes
Control variables	yes	yes	yes
R-squared	0.84	0.77	0.55
Observations	110,663	110,663	110,663

This table reports regressions of the innovation outcomes on different types of connected holdings and other control variables. Active funds are those funds with active share higher than 0.5. Closet indexers are those funds with active shares lower than 0.5. For the definition of active share, please refer to Cremers and Petajisto (2009). Table A1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in 3.2.2. The estimated model is a two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5: Connected Holdings and R&D

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	$LnR\&D$	$LnR\&D$	$\widetilde{LnR\&D}$	$\widetilde{LnR\&D}$	$LnR\&D_{t+2}$	$LnR\&D$	$LnR\&D$
Connection type	Narrow	Narrow	Narrow	Narrow	Narrow	Broad	Broad
$ConHold$	73.39***	50.46***	63.29***	42.23***	43.53***	7.56***	6.09***
	(22.93)	(16.07)	(16.63)	(13.24)	(16.23)	(1.94)	(1.47)
Year FE	yes	yes	yes	yes	yes	yes	yes
Firm FE	yes	yes	yes	yes	yes	yes	yes
University FE	yes	yes	yes	yes	yes	yes	yes
Control variables	no	yes	no	yes	yes	no	yes
R-squared	0.91	0.93	0.93	0.95	0.94	0.91	0.93
Observations	52,507	51,749	63,391	62,521	43,470	52,507	51,749

This table reports results of the regressions of the R&D expenditure on connected holdings and other control variables. Table A1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 3.2.2. The estimated model is a two-stage least squares model with firm and year fixed effect, and robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 6: Connected Holdings and Innovation Efficiency

	(1)	(2)	(3)
Dependent variable	Patent/RD	Citation/RD	KPSS/RD
$ConHold$	1.96	7.00	20.82***
	(2.43)	(9.41)	(7.31)
Year FE	yes	yes	yes
Firm FE	yes	yes	yes
University FE	yes	yes	yes
Control variables	yes	yes	yes
R-squared	0.41	0.55	0.62
Observations	110,663	110,663	110,663

This table reports the results of regressions of the innovation efficiency. Definitions of dependent variables can be found in section 4.3. Table A1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. The estimated model is a two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7: Connected Holdings and Patent Nature

Dependent variable	Originality	Generality
$ConHold$	1.21	4.37*
	(2.29)	(2.42)
Year FE	yes	yes
Firm FE	yes	yes
University FE	yes	yes
Control variables	yes	yes
R-squared	0.41	0.44
Observations	24,927	23,456

This table reports regressions of the patent Originality and Generality on connected holdings and other control variables. Table A1 provides variable definitions. Control variables are: annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, CAPX/Asset. University FE are defined in Section 3.2.2. The estimated model is a two-stage least squares with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 8: Connected Holdings and Firm Growth

Panel A: Profit		
$\tau = 1$	$\tau = 3$	$\tau = 5$
5.57	8.70	6.74
(3.55)	(3.08)	(2.45)
Panel B: Output		
$\tau = 1$	$\tau = 3$	$\tau = 5$
5.93	8.17	5.89
(4.14)	(3.31)	(2.41)
Panel C: Employment		
$\tau = 1$	$\tau = 3$	$\tau = 5$
3.23	6.63	5.75
(2.94)	(2.65)	(2.11)

This table reports regression results for equation 4. The dependent variables are firm profit growth, sales growth, and employment growth. τ is the length of the horizon, in unit of years. Table A1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. The estimated model is a two-stage least squares model with firm and year fixed effects. *t*-statistics were calculated and are provided in parentheses.

Table 9: Trading Behavior

Dependent variable	(1) T^c	(2) T^{nc}	(3) T^{diff}	(4) T^{diff}	(5) T^{diff}
$MISS$	-0.1359	-35.03***	34.89***	13.09***	15.94***
	(0.1490)	(3.531)	(3.537)	(3.665)	(3.504
Control variables	no	no	no	yes	yes
Time FE	no	no	no	no	yes
Industry FE	no	no	no	no	yes
R-squared	0.0000	0.0008	0.0008	0.0093	0.0998
Observations	127,879	127,879	127,879	107,858	105,964

The dependent variable in column (1) is the quarterly change of connected holdings; in column (2), it is the quarterly change of non-connected institutional holdings; and, in columns (3) to (5), it is the difference between the quarterly change of connected holdings and the quarterly change of non-connected institutional holdings. $MISS$ is a dummy variable that equals 1 if firms miss their quarterly earnings targets. Control variables include market cap, book-to-market ratio, and the lagged twelve-month total returns. Industry refers to the Fama-French 48-industry classification. Coefficients are reported in basis points. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 10: Return Behavior

Dependent variable	(1) α^{CAPM}	(2) α^{FF3}	(3) α^{FFC}	(4) α^{FF5}	(5) α^{DGTW}
$MISS$	-267.44***	-270.33***	-259.53***	-266.28***	-269.46***
	(8.92)	(8.44)	(8.21)	(8.33)	(9.28)
$AnyCon$	14.67	7.66	80.39	0.87	-31.43
	(11.69)	(11.07)	(17.86)	(10.98)	(11.90)
$MISS \times AnyCon$	90.59***	83.70***	80.39***	79.57***	93.63***
	(19.58)	(18.36)	(17.86)	(17.92)	(19.56)
R-squared	0.0074	0.0083	0.0079	0.0083	0.0081
Observations	147,226	147,226	147,226	147,226	147,226

The dependent variable in columns (1) to (5) is the monthly alpha estimated by the following models: CAPM, Fama-French 3-factor model, Fama-French-Carhart model, Fama-French 5-factor model, and the characteristics-adjusted model (DGTW). Dependent variables are in basis points. $MISS$ is a dummy variable that equals 1 if firms miss their quarterly earnings targets. $AnyCon$ is a dummy variable that equals 1 if there is positive connected holdings for firm i in quarter t. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 11: Connected Holdings and Takeover Exposure

Dependent variable	Takeover exposure
$ConHold$	-0.8278***
	(0.1986)
Year FE	yes
Firm FE	yes
University FE	yes
Control variables	yes
R-squared	0.8841
Observations	109,574

The dependent variable is a firm's takeover exposure. It is computed as the predicted value of the following logit regression:

$$Target_{i,t+1} = \alpha + \beta_1 Q_{i,t} + \beta_2 PPEAsset_{i,t} + \beta_3 LnCash_{i,t} + \beta_4 BLOCK_{i,t} + \beta_5 LnMV_{i,t} \\ + \beta_6 IndMA_{i,t} + \beta_7 Lev_{i,t} + \beta_8 ROA_{i,t} + Ind_j + Year_t + \epsilon_{i,t}.$$

The specification of the regression is the main specification in the paper and can be found in Section 3.1. $ConHold$ is the annual connected holdings of the firm. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 3.2.2. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 12: Connected Mutual Fund Votes on Governance Issues

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable = 1 if vote with board recommendation	All	All	Require independent board chairman	Require a majority vote for the election of directors	Call special meetings	Restore or provide for cumulative voting
$Connected(broad)$	0.0940**	0.1728***	0.1763***	0.2805***	0.1974***	0.0374
	(0.0374)	(0.0335)	(0.0457)	(0.0541)	(0.0539)	(0.0547)
Control variables	no	yes	yes	yes	yes	yes
Industry FE	no	yes	yes	yes	yes	yes
Year FE	no	yes	yes	yes	yes	yes
Pseudo R-squared	0.0004	0.1086	0.1531	0.0958	0.0659	0.0535
Observations	219,040	219,040	63,330	60,097	54,011	41,602
Marginal effect on $Connected(broad)$	0.0374	0.0689	0.0635	0.1052	0.0727	0.0147

Each observation represents the vote of a mutual fund on a proposal at a shareholder meeting at a company. The dependent variable equals 1 if mutual fund's vote follows board's recommendation. $Connected(broad)$ is a dummy variable that equals 1 if the mutual fund manager attended the same university as at least one of the firm's board members . Control variables include ISS recommendation, firm size, book-to-market ratio, last year total stock return, fund size, fund turnover ratio, fund expense ratio, and fund family size. Industry refers to the Fama-French 12-industry classification. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix

A A Simple Model

In this section, we present a simple model that builds on Gennaioli *et al.* (2015). The goal of the model is to capture, in the simplest way, the effect of social capital on R&D investment. Here, in my setting, I interpret social capital as trust. On trust, Gennaioli *et al.* (2015) write "Rather, trust describes confidence in the manager that is based on personal relationships, familiarity, persuasive advertising, connections to friends and colleagues, communication, and schmoozing. There are (at least) two distinct aspects of such trust. The first, stressed by Guiso, Sapienza, and Zingales (2004, 2008) and Georgarakos and Inderst (2011), sees trust as security from expropriation or theft. The other aspect, emphasized here, has to do with reducing investor anxiety about taking risk. With U.S. securities laws, most investors in mutual funds probably do not fear that their money will be stolen; rather, they want to be "in good hands."

A.1 The Basic Setup

There are two periods, $t = 0, 1$, and two types of projects in which the firm can invest, an innovative project (IP) and a non-innovative project (NIP). For the NIP, I assume it is riskless and yields $R_f > 1$ at $t = 1$. IP is risky; it yields an expected excess return, R, over the riskless asset and has a variance of σ. Here, I take the projects' payoff as exogenously given.

There are two types of investors: connected investors (C) and non-connected investors (NC). I assume the connected investors trust the firm more than the non-connected investors. This differential in trust could stem from the fact that connected investors attended the same school as the management so that they know the management's quality. Said another way, connected investors "speak the same language" as the firm's management. So based on the same information, they believe they can better understand the firm. Here, I take trust as exogenously given, and I believe it is reasonable to assume, everything else equal, the connected investors should trust the firm more. Out of a mass one of investors, I assume a λ fraction of them are connected investors, where $\lambda \in [0, 1]$. When $\lambda = 0$, all the investors are non-connected.

Investors have mean variance utility:

$$u_i = E(c_i) - \frac{a_i}{2} Var(c_i) \qquad i = \text{C,NC},$$

where c_i stands for consumption at time 1 for type i investors. Here, we assume $a_C < a_{NC}$. This means that the perceived risk for the IP causes lower disutility for the connected investors than for the non-connected investors. Of course, other conceptions of trust exist, and thus other modeling choices. Like in the Guiso, Sapienza, and Zingales (2004) framework, trust affects the expected return, but with the mean variance utility. These two types of specifications are similar.

There is one unit of resource at firm at $t = 0$. Firms maximize the utility of their investors by choosing the optimal amount of resources, x, invested in the innovative project (IP). We can formulate the firm's problem as follows:

$$\max_{x} \quad \lambda(R_f + xR - \frac{a_C}{2}x^2\sigma) + (1-\lambda)(R_f + xR - \frac{a_{NC}}{2}x^2\sigma).$$

The optimal investment in an IP is

$$\hat{x} = \frac{R}{(\lambda a_C + (1-\lambda)a_{NC})\sigma}.$$

Proposition 1. *Investment in innovative projects increases with the fraction of connected investors, λ, and the expected payoff of innovative projects, R, and decreases with the riskiness of the innovative projects, σ.*

Proof. We have the following partial derivatives:

$$\frac{\partial \hat{x}}{\partial \lambda} = \frac{R(a_{NC} - a_C)}{(\lambda a_C + (1-\lambda)a_{NC})^2\sigma^2} > 0,$$

$$\frac{\partial \hat{x}}{\partial R} = \frac{1}{(\lambda a_C + (1-\lambda)a_{NC})\sigma} > 0,$$

$$\frac{\partial \hat{x}}{\partial \sigma} = \frac{-R}{(\lambda a_C + (1-\lambda)a_{NC})\sigma^2} < 0.$$

□

We can use a_{NC} to proxy for capital market pressure. The higher the capital market pressure, the smaller the investment in innovative projects.

Proposition 2. *Investment in innovative projects decreases with capital market pressure a_{NC}.*

Proof. We have the following partial derivative:

$$\frac{\partial \hat{x}}{\partial \lambda} = \frac{-(1-\lambda)\sigma}{(\lambda a_C + (1-\lambda)a_{NC})^2\sigma^2} < 0.$$

□

Proposition 2 supports the reasoning in Guiso *et al.* (2008). The less trustworthy is the firm, the more anxious the non-connected investors would feel. Then the firm will not invest in R&D.

Claim 1. The above effect should be stronger for small, less analyst-covered firms.

The above analysis is based on trust. One way to endogenize this trust could be differential information. Small firms, or firms with more information asymmetries, should benefit more from the connected holdings.

B Mutual Fund Manager's Connections and Career Path

In this section, I investigate the relationship between the number of connections mutual fund managers have and their career outcome, with a special focus on termination, promotion, and demotion.

I follow Chevalier & Ellison (1999) by focusing on single-manager-managed funds. In my sample, there are 19,993 fund-year observations. Notice that this sample is not at the fund-manager-year level because one manager could manage more than one fund. To compute fund manager age, I require the manager's birth year. I obtain birth years in the following ways: (1) Morningstar provides the birth year for a subset of fund managers or (2) if Morningstar provides the manager's BA year, I subtract 21 from the year to get the birth year. For the performance of the fund, I use the Pastor, Stambaugh, and Taylor (2016) method. I subtract the corresponding monthly index return from the fund's monthly gross return. Then I aggregate across months to obtain the annual outperformance. This method allows me to avoid estimating the beta, and some studies show that this method is better at controlling for risks than is the standard linear factor model. For the investment category of the fund, I use the Morningstar category. To measure the mutual fund managers' connections to corporate I used two measures: (1) the number of connected board members and (2) the number of connected firms. The connection definition is CONNECTED(narrow). The dependent variable $Termination_{i,t}$ is set to one if the manager responsible for fund i in January of year t is no longer in charge of the fund at the beginning of year $t+1$. I also construct a promotion dummy that equals 1 if the manager managed more funds in year t than in year $t-1$ and a demotion dummy that equals 1 if the manager managed fewer funds in year t than in year $t-1$. The specification is as follows:

$$\begin{aligned} Y_{i,t} &= \beta_0 + \beta_1 Connection + \beta_2 Alpha_{i,t} + \beta_3 Alpha_{i,t} \times (MgrAge_{i,t} - \overline{Age}) \\ &\quad + \beta_4 Alpha_{i,t-1} + \beta_5 Alpha_{i.t-2} + \beta_6 MgrAge_{i,t} + InvCat_j + \mu_t + \epsilon_{i,t}, \end{aligned} \tag{6}$$

where Y is a dummy for termination, promotion, or demotion. $InvCat_j$ is the fund investment category fixed effects and μ_t is the time fixed effects. Since the dependent variable is a binary outcome, I use a probit model to estimate it. Table A15 presents the results. For the three types of career outcomes, connections to the firm are not statistically significant. This means that connections with board members are not important enough to affect the career path of mutual fund managers .

[INSERT TABLE A15 HERE]

C Connected Mutual Fund and Say on Pay Votes

The Dodd-Frank Act mandates that, starting from 2011, the management of public firms should submit proposals about the top five executives' compensation and shareholders can cast advisory vote of "yes" or "no" on the pay of the company's top named executives during the prior fiscal year. Table A16 reports the results. Connected mutual funds vote in favor of management.

[INSERT TABLE A16 HERE]

D Resistance to Hedge Fund Activism

Activists go after companies with vulnerabilities. Most companies have some vulnerabilities, which could be related to financial performance, such as missing quarterly numbers, a stagnant stock price, or comparatively weak revenue growth. Weaknesses, such as missing quarterly numbers, could be a result of a firm engaging in long-term plans. But hedge funds could step in and "increase" the efficiency of the firm. The activists' agenda usually includes increasing leverage, returning excess cash to shareholders, selling off non-core corporate assets, cutting operating costs, and demanding representation on the board. Some of those procedures could be harmful to the firm's long-term value such as innovation.

In this section, I check connected mutual fund's trading behavior when facing hedge fund activism. I find that during the period when hedge funds accumulate their shares, connected funds stick with their equity in the firm, whereas non-connected funds tend to sell.

I obtain hedge fund activism data from Professor Wei Jiang. Brav *et al.* (2008) provide details on how they construct the dataset. The current hedge fund activism data I am using cover the period from 1994 to 2011. Following the literature, I use the date that the hedge fund files a Schedule 13D as the event date. I aggregate data into quarterly frequency. For a given firm, in a given quarter, if there is at least one hedge fund targeting them, I label this quarter as an event quarter for the firm.[21] In the final sample, there are 1,822 hedge fund activism events affecting 1,452 different firms.

Table A17 shows the results. In column (1), the dependent variable is the change in $ConHold$ from quarter $t-3$ to quarter t, and, in column (2), it is the change in $Non-ConHold$ from quarter $t-3$ to quarter t. $Non-ConHold$ is the difference between institutional holdings and $ConHold$. The dependent variable in columns (3) to (5) is the difference between the above two variables. All

[21] In the data, there exists the case in which more than one hedge fund activism event occurred at a firm within a quarter. There are only 63 cases.

the dependent variables are in basis points. Hedge fund activism Is a dummy variable that equals 1 if the firm in quarter t has a hedge fund activism. Control variables include firm size, book-to-market ratio, and the previous year's total stock return. Industry refers to the Fama-French 12-industry classifications. Standard errors are clustered at the firm level.

In column (1), I find that connected holdings are not affected by hedge fund activism. This means that when the hedge fund tries to accumulate shares in the firm, connected mutual fund are less likely to sell. Meanwhile, for the non-connected funds, their shares significantly decrease in the period before hedge fund activism actually happens. The difference between the two types of funds trading can amount to 1% of a firm's equity. Usually, hedge funds accumulate around 5% of a firm's equity before they initiate their activism. This means that institutional investors who are not connected to the firm provide at least 1%. The results are robust to include more control variables and industry and time fixed effects.

[INSERT TABLE A17 HERE]

E Passive Funds in the Data

Table A18 provides information about active and passive funds in my dataset. Passive funds compose small amount of AUM in the data.

F Possion Model and Control Function Method

The basic assumption is that y's distribution, conditional on all $\boldsymbol{x}$, follows the Poisson distribution. This is similar to the MLE world. We assume that, conditional on all the $\boldsymbol{x}$, y follows a normal distribution. Then for the Poisson distribution, it is as follows:

$$f(y|x) = exp[-\mu(x)][\mu(x)]^y / y!,$$

where $\mu(x)$ decides the mean of the Poisson distribution. We can choose a parametric form for the $\mu(x) = exp(x\beta)$. This is because the $\mu(x)$ has to be positive all the time.

In my situation, because I assume $ConHold$ is endogenous. I will use a control function method to deal with the endogenity problem in the exponential regression setting.

First, assume the structural model follows:

$$E(y|z, ConHold, c_1) = exp(z_1\boldsymbol{\delta}_1 + ConHold\gamma_1 + c_1), \tag{7}$$

where c_1 is the unobserved latent variable that affects innovation and potentially causes the endogeneity problem. $z_1$is a $1 \times L_1$subset of z containing unity. Here, we assume z contains both an exogenous regressor, z_1 (i.e., all the control variables), and instrument variables, z_2 (i.e., $ConMFM$). On the one hand, z is assumed to be uncorrelated with c_1. On the other hand, I allow the correlation between $ConHold$ and c_1. Further, I assume $ConHold$ has a linear reduced form satisfying the following assumption:

$$ConHold = z\mathbf{\Pi}_2 + v_2. \tag{8}$$

This is exactly the first-stage regression in the IV analysis. In addition, I assume that (c_1, v_2) is independent of z and

$$c_1 = \rho_1 v_2 + e_1, \tag{9}$$

where e_1 is independent of v_2. Note that $ConHold$ is exogenous if and only if $\rho_1 = 0$.

Then I can write the original structural model as

$$E(y|z, ConHold, c_1) = exp(z_1\boldsymbol{\delta}_1 + ConHold\gamma_1 + v_2\rho_1), \tag{10}$$

and estimating this equation using the standard quasi-maximum likelihood estimate (QMLE) method could allow me to obtain consistent estimates for $\boldsymbol{\delta}_1, \gamma_1, \rho_1$. In implementation, estimate expression (8) and then compute the residual $\hat{v}_2 = ConHold - z\hat{\mathbf{\Pi}}_2$. Then plug the residual $\hat{v}$ along with z_1 and $ConHold$ into equation (10).

G Citation Adjustment Method

I use the patents filed between 1980 and 1990 as my sample for estimation. For each technology class, I first compute the total number of citations. Then I aggregate the number of citations a specific technology class patents receives in each subsequent year after being filed. I divided each year's citation number by the total number of citations to obtain the adjustment factors, which are presented in Table A12. Then I adjust the number of citations that a patent currently has based on the number of years the patent has been filed and the adjustment factors.

Table A1: Variable Definitions

Variable	Definition
Measures of innovation	
$PatCite$	Total number of citations received by the filed patents for firm i in year t
$PatApp$	Total number of patents filed (and eventually granted) by firm i in year t
$KPSS$	Total dollar value of innovation produced by firm i in year t, based on stock market's reaction to the patents granted news, scaled by firm's total asset
Measures of connected holdings and other control variables	
$ConHold$	Fraction of company i 's equity held by connected mutual funds in year t
Inst holding	Calculated as the mean of the four quarterly institutional holdings reported on form 13F
DED inst holding	The institutional holdings by dedicated institutional investors. Classification refers to Bushee (2001)
QIX inst holding	The institutional holdings by quasi-indexer institutional investors. Classification refers to Bushee (2001)
TRA inst holding	The institutional holdings by transitory institutional investors. Classification refers to Bushee (2001)
ROA	Return-on-assets ratio defined as operating income before depreciation (oibdp) divided by total assets (at)
Vol	Firm i's idiosyncratic volatility measured as the standard deviation for the idiosyncratic part of firm return defined as $\tilde{r} = r - r_m$, where r is firm's daily return, r_m is the daily market return
R&D stock	Following Hall, Jaffe, and Trajtenberg (2005), this variable is calculated using a perpetual inventory method: $G_t = R_t + (1-\delta)G_{t-1}$, where R_t is the R&D expenditure in year t and $\delta = 0.15$, the private depreciation rate of knowledge
PPE/Assets	Property, plant, & and equipment (ppent) divided by total assets (at)
Leverage	Book value of debt (dlc+dltt) divided by total assets (at)
Q	Tobin's Q, calculated as market value of equity (csho*prcc_c) plus total assets (at) minus book value of equity (ceq) minus balance sheet deferred taxes (pstkl, set to 0 if missing), divided by total assets (at)
R&D/Asset	Research and development expenditures (xrd+rdip) divided by total assets (at), set to 0 if missing
Capex/Assets	Capital expenditure (capx) scaled by total assets (at)
Age	The number of years since initial public offering (IPO) (proxied by the first appearance in Compustat)
Sale	Firm's sale

Table A2: Days between Patent Filing Date and Grant date

	Mean	SD	P1	P5	P25	P50	P75	P95	P99	N
1980–1989	728	387	246	337	518	671	854	1,275	2,002	235,577
1990–1999	714	385	235	316	495	652	854	1,289	1,925	352,370
2000–2010	1,141	561	292	433	732	1,037	1,453	2,171	2,828	730,670

This table reports the days between patent filing date and patent grant date in the dataset for the period from 1980 to 2010. Here, I only include patents granted to public firms. (Other type of entities, for example, universities, also have patents, but are not included). Means, standard deviations, and key percentiles are provided in the table.

Table A3: Number of Connections

Broad connection: Same university		Narrow connection: Same univ/same degree/same time	
Harvard University	3,977,753	Harvard University	49,273
University of Pennsylvania	2,524,600	University of Pennsylvania	40,600
Columbia University	1,269,328	University of Chicago	19,530
University of Chicago	1,263,519	Stanford University	15,968
New York University	1,177,280	University of California, Berkeley	13,150
Stanford University	954,470	Columbia University	11,166
University of California, Berkeley	839,847	University of Wisconsin, Madison	10,940
Northwestern University	579,842	Princeton University	10,284
University of Michigan	490,253	University of Michigan	9,980
University of Wisconsin, Madison	377,480	Yale University	9,360
Cornell University	374,640	University of Virginia	9,175
Massachusetts Institute of Technology	362,230	University of Illinois	9,169
University of Texas at Austin	326,298	Cornell University	9,099
Yale University	317,646	New York University	8,566
University of Virginia	311,120	Dartmouth College	7,823

This table reports the number of connections for the top 15 educational institutions according to two definitions: broad connection broad, that is, mutual fund managers and corporate board members attend the same school, and narrow connection, that is, mutual fund managers and corporate board members attended the same school for the same type of degree at the same time.

Table A4: Top 10 Most Connected University

Connected firms		Connected funds	
Harvard University	780	Harvard University	178
University of Pennsylvania	344	University of Pennsylvania	172
Stanford University	330	University of Chicago	115
Columbia University	257	Columbia University	110
Yale University	237	New York University	105
New York University	226	University of Wisconsin, Madison	78
Massachusetts Institute of Technology	196	Stanford University	71
Princeton University	194	Yale University	67
University of Michigan	182	Northwestern University	65
University of Chicago	180	University of California, Berkeley	62

This table reports the time-series average of connected firms and connected funds for the sample from 1980 to 2003. *Connected* means board members and fund managers obtained degrees from the same university.

Table A5: Summary Statistics

Panel A: Full sample							
Variable	Mean	SD	P5	P25	P50	P75	P95
$PatCite$	74.373	321.398	0	0	0	0	320.396
$PatApp$	4.078	18.000	0	0	0	0	16
$KPSS$	0.027	0.122	0	0	0	0	0.135
$ConHold^{broad}(\%)$	0.401	1.743	0	0	0	0	2.509
$ConHold^{narrow}(\%)$	0.020	0.251	0	0	0	0	0
Inst holding(%)	18.556	23.822	0	0	6.238	32.549	69.019
DED inst holding(%)	2.663	5.624	0	0	0.000	2.905	13.671
QIX inst holding(%)	11.304	15.281	0	0	2.972	19.227	44.016
TRA inst holding(%)	4.119	7.355	0	0	0.092	5.511	19.840
ROA	0.049	23.593	-0.425	0.015	0.108	0.171	0.278
Vol	0.608	1.964	0.044	0.124	0.279	0.630	2.037
R&D stock	44.448	619.595	0	0	0	0	47.201
PPE/Assets	0.306	0.231	0.032	0.121	0.248	0.440	0.782
Leverage	0.240	0.213	0	0.049	0.207	0.369	0.646
Q	1.990	1.813	0.725	1.019	1.365	2.147	5.582
R&D/Asset	0.011	0.046	0	0	0	0	0.070
CAPX/Asset	0.074	0.078	0.004	0.024	0.049	0.093	0.237
Observations	113,503						

Panel B: Samples Sorted by Connected Holdings				
	$ConHold^{narrow} > 0$		$ConHold^{narrow} = 0$	
Variable	Mean	SD	Mean	SD
$PatCite$	403.563	758.187	62.164	285.736
$PatApp$	22.344	42.523	3.401	16.004
$KPSS$	0.132	0.293	0.023	0.109
$ConHold^{broad}(\%)$	5.320	4.672	0.219	1.188
$ConHold^{narrow}(\%)$	0.566	1.208	0.000	0.000
Inst holding(%)	60.895	19.023	16.985	22.498
DED inst holding(%)	8.475	7.019	2.448	5.448
QIX inst holding(%)	37.327	13.062	10.339	14.484
TRA inst holding(%)	14.689	9.798	3.727	6.946
ROA	0.142	0.124	0.046	0.238
Vol KPSS	0.265	0.306	0.621	1.998
R&D stock	508.795	2070.772	27.226	480.464
PPE/Assets	0.287	0.205	0.307	0.231
Lev	0.223	0.185	0.241	0.214
Q	2.553	1.952	1.970	1.804
R&D/Asset	0.035	0.062	0.010	0.045
CAPX/Asset	0.065	0.057	0.074	0.078
Observations	4,059		109,444	

Panel A reports the summary statistics for the full sample of variables constructed based on the sample of U.S. public firms from 1980 to 2003. Panel B separately shows statistics for those with positive CONNECTED(narrow) holdings (left) and those with no CONNECTED(narrow) holdings (right). Table A1 provides variable definitions.

Table A6: Age distribution

	Mean	SD	P10	P25	P50	P75	P90
Fund manager	47.04	10.16	35	40	45	54	62
Board member	54.16	9.90	42	47	54	61	67

This table displays the distribution of fund manager and board member age in my dataset as of year 2003.

Table A7: Connected Holdings and Innovation (Poisson Model)

	(1)	(2)
Dependent Variable	$PatApp$	$PatCite$
$ConHold$	362.32***	465.44***
	(74.50)	(79.76)
Year FEs	yes	yes
Firm FEs	yes	yes
Control variables	yes	yes
Observations	113,503	113,503

Table A1 lists variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. The estimated model is the Poisson model with firm and year fixed effects, and robust standard errors are calculated and provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A8: Connected Holdings and Innovation Controling for CEO FE

	(1)	(2)	(3)
Dependent variable	$Ln\widetilde{PatApp}$	$Ln\widetilde{PatCite}$	$Ln\widetilde{KPSS}$
$ConHold$	37.41***	66.91**	9.13***
	(14.49)	(27.84)	(2.86)
Year FE	yes	yes	yes
Firm FE	yes	yes	yes
University FE	yes	yes	yes
CEO FE	yes	yes	yes
Control variables	yes	yes	yes
R-squared	0.89	0.83	0.68
Observations	29,219	29,219	29,219

This table reports the results of regressions of the innovation outcomes on connected holdings and other control variables. The difference between this table and Table is that in this table, I also control for CEO fixed effects. Table A1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 3.2.2. The estimated model is two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate that the coefficient is significant at the 1%, 5%, and 10% level, respectively.

Table A9: Connected Holdings (Management Only) and Innovation

	(1)	(2)	(3)	(4)	(5)	(6)
	Senior management connection			Board member connection (baseline)		
Dependent variable	$\widetilde{LnPatApp}$	$\widetilde{LnPatCite}$	$\widetilde{LnKPSS}$	$\widetilde{LnPatApp}$	$\widetilde{LnPatCite}$	$\widetilde{LnKPSS}$
$ConHold$	50.32**	82.50**	10.88***	42.26***	59.47***	8.83***
	(20.56)	(36.80)	(3.52)	(12.83)	(22.60)	(2.10)
Year FEs	yes	yes	yes	yes	yes	yes
Firm FEs	yes	yes	yes	yes	yes	yes
University FEs	yes	yes	yes	yes	yes	yes
Control variables	yes	yes	yes	yes	yes	yes
R-squared	0.84	0.77	0.50	0.84	0.77	0.59
Observations	110,663	110,663	110,663	110,663	110,663	110,663

This table reports the results of regressions of innovation outcome on connected holdings and other control variables for two types of connection definitions. Table A1 lists the variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FEs are defined in Section 3.2.2.The estimated model is a two-stage least-squares model with firm and year fixed effects, and robust standard errors are calculated and provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A10: Connected Holdings (Broad) and Innovation

	(1)	(2)	(3)
Dependent variable	$\widetilde{LnPatApp}$	$\widetilde{LnPatCite}$	$\widetilde{LnKPSS}$
$ConHold^{Broad}$	5.45***	4.95**	0.53***
	(1.25)	(2.08)	(0.15)
Year FE	yes	yes	yes
Firm FE	yes	yes	yes
University FE	yes	yes	yes
CEO FE	yes	yes	yes
Control variables	yes	yes	yes
R-squared	0.85	0.77	0.63
Observations	110,663	110,663	110,663

This table reports the results of regressions of the innovation outcomes on connected holdings and other control variables. The difference between this table and Table is that in this table, I use broad connections to compute connected holdings. Broad connection means board members and mutual fund managers went to the same school. Table A1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 3.2.2. The estimated model is two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate that the coefficient is significant at the 1%, 5%, and 10% level, respectively.

Table A11: Connected Holdings versus Institutional Holdings

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	$\widetilde{lnPatApp}$	$\widetilde{lnPatCite}$	$\widetilde{lnKPSS}$	$\widetilde{lnPatApp}$	$\widetilde{LnPatCite}$	$\widetilde{lnKPSS}$
$ConHold$	42.26***	59.47***	8.83***			
	(12.83)	(22.60)	(2.10)			
Inst holdings	0.031	0.133**	-0.018***	0.079***	0.201***	-0.008***
	(0.031)	(0.062)	(0.004)	(0.029)	(0.058)	(0.003)
Year FEs	yes	yes	yes	yes	yes	yes
Firm FEs	yes	yes	yes	yes	yes	yes
University FEs	yes	yes	yes	yes	yes	yes
Control variables	yes	yes	yes	yes	yes	yes
R-squared	0.84	0.77	0.59	0.82	0.74	0.58
Observations	110,663	110,663	110,663	110,663	110,663	110,663

Table A12: Citation adjustment factors

Age \ Tech class	1	2	3	4	5	6
0	0.0083	0.0059	0.0047	0.0071	0.0061	0.0058
1	0.0442	0.0303	0.0259	0.0384	0.0350	0.0324
2	0.1049	0.0735	0.0609	0.0931	0.0871	0.0798
3	0.1725	0.1252	0.1051	0.1559	0.1480	0.1345
4	0.2403	0.1832	0.1540	0.2201	0.2097	0.1911
5	0.3083	0.2450	0.2067	0.2834	0.2724	0.2492
6	0.3733	0.3092	0.2627	0.3447	0.3326	0.3067
7	0.4365	0.3749	0.3212	0.4035	0.3912	0.3642
8	0.4954	0.4394	0.3810	0.4596	0.4465	0.4199
9	0.5521	0.5012	0.4414	0.5127	0.5006	0.4740
10	0.6072	0.5611	0.5033	0.5639	0.5548	0.5285
11	0.6589	0.6191	0.5641	0.6151	0.6076	0.5835
12	0.7094	0.6747	0.6241	0.6645	0.6592	0.6380
13	0.7577	0.7278	0.6873	0.7144	0.7114	0.6908
14	0.8042	0.7801	0.7482	0.7633	0.7612	0.7439
15	0.8479	0.8297	0.8042	0.8110	0.8108	0.7964
16	0.8864	0.8728	0.8553	0.8560	0.8571	0.8455
17	0.9213	0.9117	0.8988	0.8986	0.8984	0.8922
18	0.9517	0.9455	0.9364	0.9370	0.9355	0.9319
20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

This table provides the citation adjustment factors. Technology classes are defined in Hall *et al.* (2001). More specifically, technology class: 1, Chemical (excluding Drugs); 2, Computers and Communications (C&C); 3, Drugs and Medical (D&M); 4, Electrical and Electronics (E&E); 5, Mechanical; and 6, Others. See appendix 1 in Hall *et al.* (2001) for more details. Age is the number of years the patent has been filed. I make the assumption that a patent accumulates all its citations after 20 years.

Table A13: Daily Trading Volume and Connected Holdings

	Daily Trading Volume			$ConHold^{narrow}$	$ConHold^{broad}$	Inst Holding	
	Mean	P25	P75	Mean	Mean	Mean	Observations
1990	0.0050	0.0026	0.0061	0.0041	0.0147	0.5242	38
2000	0.0099	0.0053	0.0116	0.0054	0.0579	0.5993	469
2003	0.0097	0.0112	0.0055	0.0064	0.0672	0.6897	593

This table reports daily trading volume for stocks with positive $ConHold^{narrow}$. Daily trading volume is defined as the daily stock sold divided by The total shares outstanding. For each stock in each year, I compute the annual average. I also report the 25th and 75th percentile of the daily trading volume for a given stock in a given year. Table A1 provides all variable definitions.

Table A14: Summary Statistics for Termination and Connectedness

Variable	Mean	SD	Median
Termination	0.105	0.307	0
Promotion	0.069	0.255	0
Demotion	0.036	0.186	0
No. connected firm	22.329	28.707	12
No. connected director	18.307	21.920	10
No. fund managed	1.576	1.049	1
Performance	0.019	0.115	0.006
Mutual fund manager age	45.633	8.874	45
Observations	3,416		

The data are at the fund-year frequency. I only include single manager- managed fund. Termination is a dummy equaling 1 when either the fund manager stops managing the fund or when the fund exits. Promotion is a dummy equaling 1 if a manager manages more funds in the next year than in this year. Demotion is a dummy equaling 1 if a manager manages fewer funds in the next year than in this year. According to the definition of CONNECTED(narrow) , the number of connected firms is the number of connected firms to funds in that year. A similar definition for the number of connected directors. There are cases in which one manager could manage more than one fund, so I count the number of funds managed. Alpha is calculated as the difference between a fund's gross return and its corresponding index return.

Table A15: Relationship between Termination and Connectedness

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Termination	Promotion	Demotion	Termination	Promotion	Demotion
No. connected firms	-0.000120	-0.000235	-0.00210			
	(0.00125)	(0.00136)	(0.00191)			
No. connected directors				-0.000516	-0.000359	-0.00247
				(0.00164)	(0.00178)	(0.00247)
$Alpha_t$	-0.585	0.376	-0.0934	-0.583	0.376	-0.0929
	(0.384)	(0.416)	(0.581)	(0.384)	(0.416)	(0.581)
$Alpha_t * Age$	0.0357	-0.0547	0.0687	0.0361	-0.0546	0.0679
	(0.0353)	(0.0390)	(0.0499)	(0.0353)	(0.0389)	(0.0500)
$Alpha_{t-1}$	-0.477	0.529	-0.870	-0.478	0.528	-0.875
	(0.375)	(0.390)	(0.579)	(0.375)	(0.390)	(0.578)
$Alpha_{t-2}$	-0.987**	0.667*	-0.210	-0.986**	0.667*	-0.215
	(0.387)	(0.403)	(0.542)	(0.387)	(0.403)	(0.542)
Age	-0.0110**	0.00631	0.00123	-0.0110**	0.00628	0.000870
	(0.00447)	(0.00487)	(0.00616)	(0.00445)	(0.00486)	(0.00614)
Year FEs	yes	yes	yes	yes	yes	yes
Investment category FEs	yes	yes	yes	yes	yes	yes

This table reports the probit model estimations of equation 6. Standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table A16: Connected Mutual Fund Votes on "Say-on-Pay" Issue

Dependent variable = 1 if vote with board recommendation	(1)	(2)	(3)	(4)
$Connected(broad)$	0.0366	0.0386	0.1085**	0.1083**
	(0.0338)	(0.0342)	(0.0435)	(0.0438)
Control variables	no	no	yes	yes
Industry FEs	no	yes	no	yes
Pseudo R-squared	0.0000	0.0175	0.2885	0.2918
Observations	249,205	249,205	249,205	249,205
Marginal effect on $Connected(broad)$	0.0061	0.0063	0.0124	0.0122

In this table, the sample is restricted to "say-on-pay" proposals. Each observation represents the vote of a mutual fund on a proposal made at a company's shareholder meeting. The dependent variable equals 1 if the mutual fund's vote follows the board's recommendation. $Connected(broad)$ is a dummy variable that equals 1 if the mutual fund manager went to the same university as at least one of the firm's board members . Control variables include ISS recommendation, firm size, book-to-market ratio, the previous year's total stock return, fund size, fund turn ratio, fund expense ratio, and fund family size. Industry refers to the Fama-French 12-industry classifications. Standard errors clustered at firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A17: Mutual Fund Trading and Hedge Fund Activism

	(1)	(2)	(3)	(4)	(5)
	$ConHold$	$Non-ConHold$	$Diff$	$Diff$	$Diff$
Hedge fund activism	-0.3539	-105.8***	105.49***	79.94***	99.69***
	(0.9923)	(24.10)	(24.13)	(27.46)	(31.17)
Controls	no	no	no	yes	yes
Fixed effects	no	no	no	no	yes
Observations	274,510	274,510	274,510	164,755	164,755

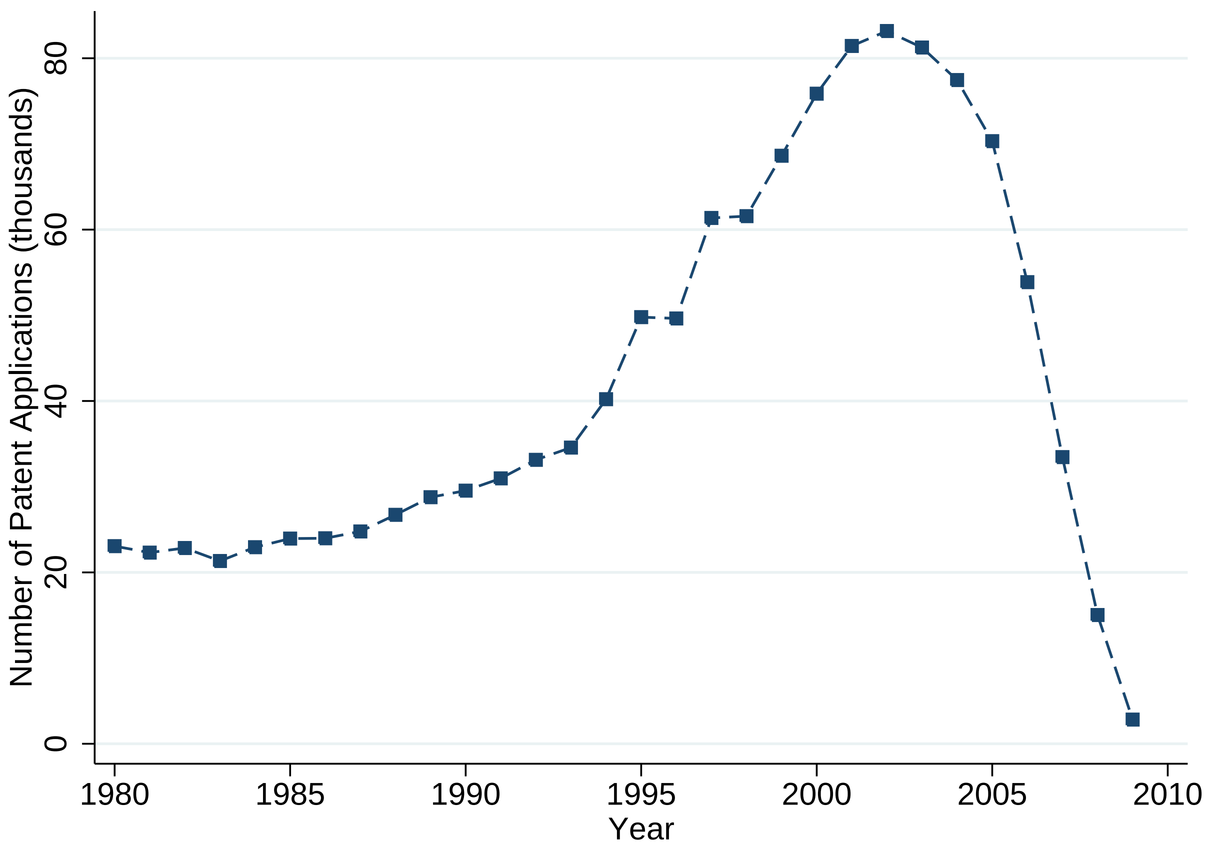
The dependent variable in column (1) is the change in $ConHold$ from quarter $t-3$ to quarter t, and, in column (2), it is the change in $Non-ConHold$ from quarter $t-3$ to quarter t. $Non-ConHold$ is the difference between institutional holdings and $ConHold$. The dependent variable in columns (3) to (5) is the difference between the above two variables. All the dependent variables are in basis points. Hedge fund activism is a dummy variable that equals 1 if the firm in quarter t has a hedge fund activism. Control variables include firm size, book-to-market ratio, and the previous year's total stock return. Industry refers to the Fama-French 12-industry classifications. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A18: Active Funds versus Passive Funds

	Number of funds			Total AUM		
	Active	Passive	Passive/Total (%)	Active	Passive	Passive/Total (%)
1980	279	1	0.35	49.42	0.018	0.03
1985	440	3	0.67	130.5	0.863	0.65
1990	801	14	1.71	259.4	1.480	0.56
1995	1,449	37	2.48	1,155	6.212	0.53
2000	2,589	76	2.85	3,186	27.43	0.85
2005	3,204	95	2.87	4,376	71.65	1.61
2010	3,031	93	2.97	5,117	115.8	2.21

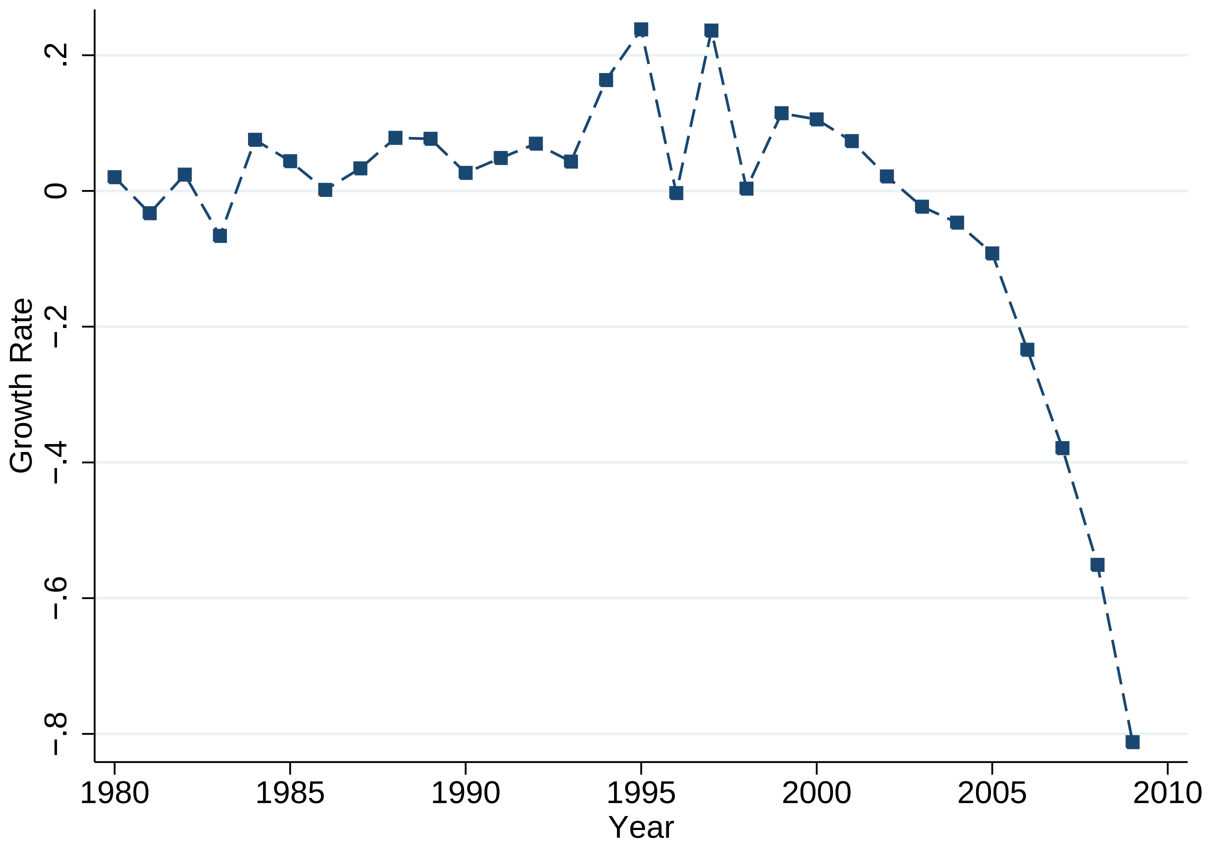
This table shows the number of funds and total AUM (in billions) for active funds and passive funds in my sample. Passive funds are defined as funds that contain "Index" or "index" in their fund name.

Figure A1: Total Number of Patent Applications



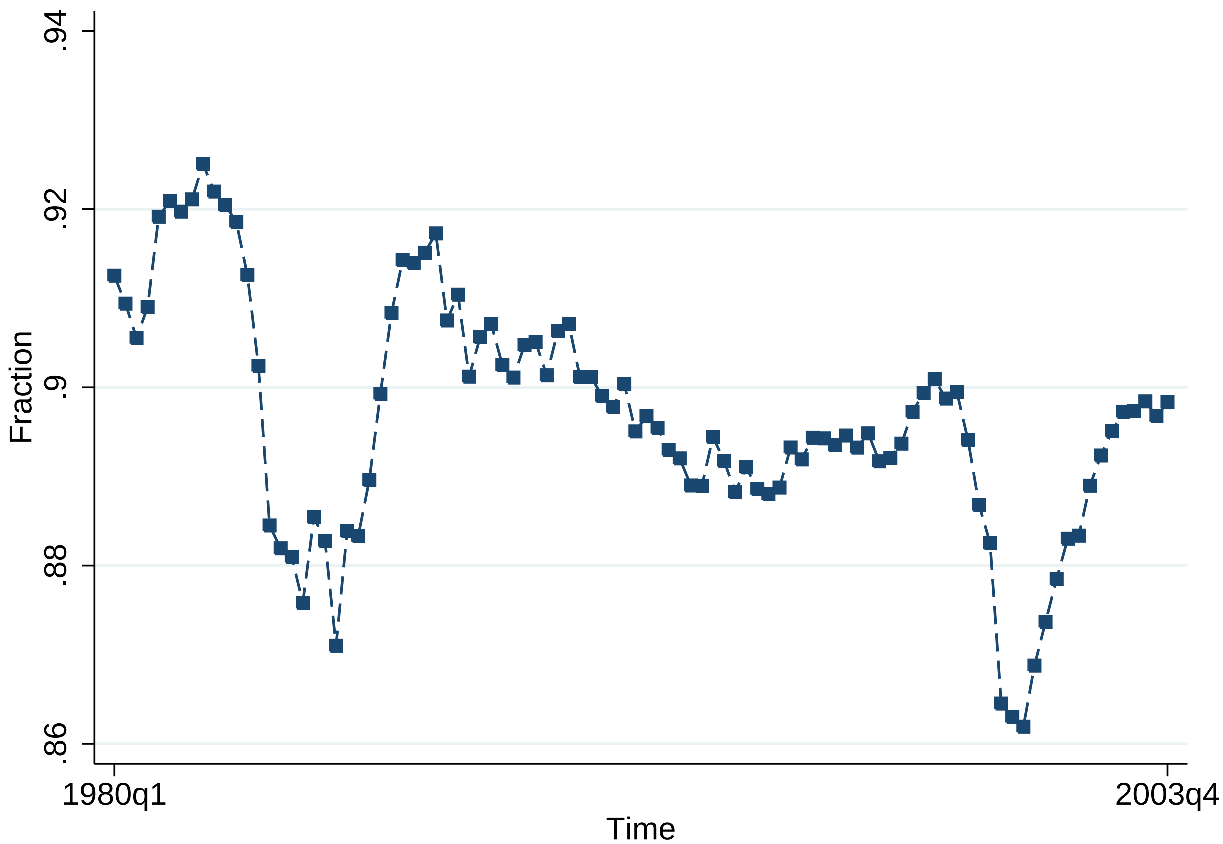
This figure displays the total number of patent applications (that are eventually granted) in the dataset by years. A censor problem shows up from year 2002 or 2003 until the end of the sample.

Figure A2: Growth Rate of the Total Number of Patent Applications



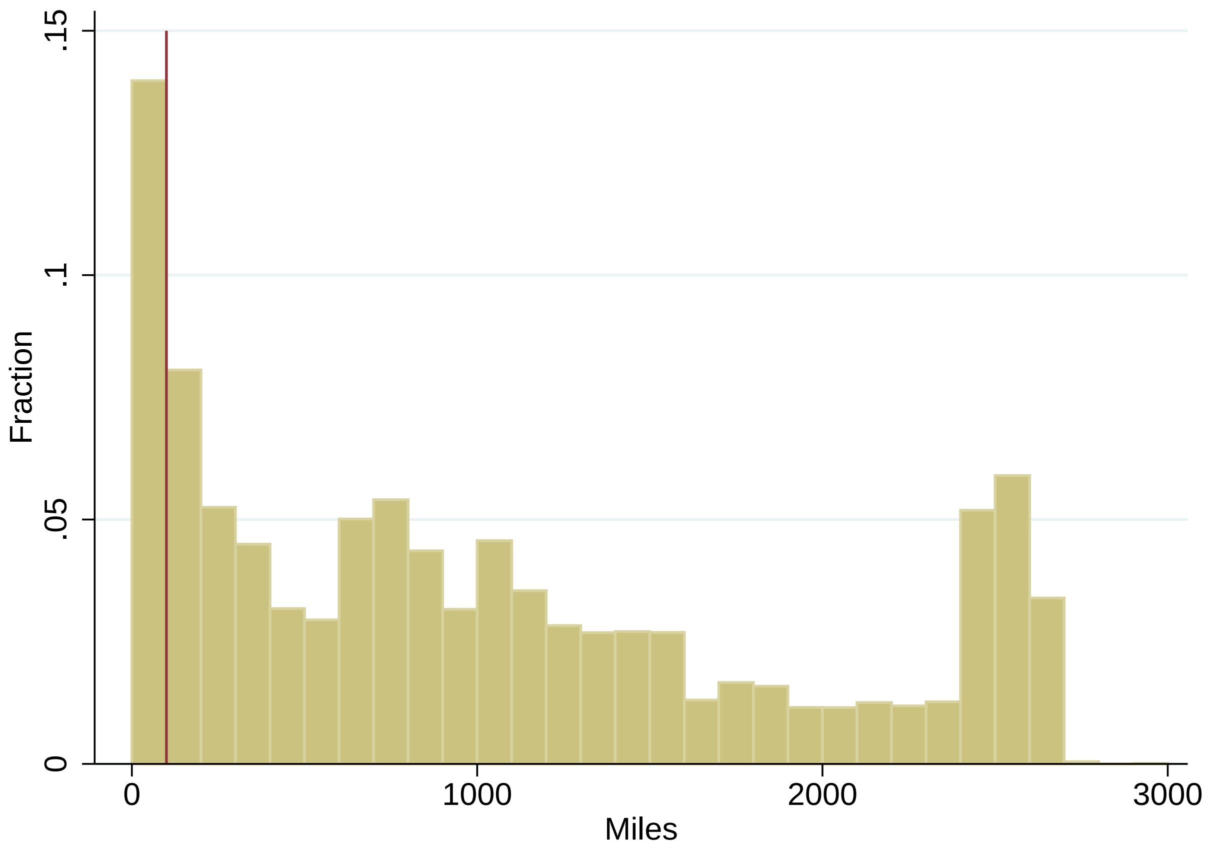
This figure displays the annual growth rate of the total number of patent applications (defined as the current year's total number of patents divided by last year's total number of patents) in the dataset from 1980 to 2010.

Figure A3: Total AUM of Domestic Active Funds as a Fraction of All Funds



The graph displays the total AUM of mutual funds that exclusively hold domestic stocks as a fraction of the total AUM of all matched mutual funds. The data series covers quarter 1 of 1980 to quarter 4 of 2003.

Figure A4: Histogram of Distance between Funds and Firms



This histogram plots the distance in miles between the headquarters of mutual fund companies and the headquarters of public firms. The vertical red line indicates 100 miles. According to Coval & Moskowitz (2001), if the distance between the headquarters of mutual fund companies and that of public firms is smaller than 100 km (which is less than 100 miles), the holdings will be defined as local holdings.