JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 53, No. 3, June 2018, pp. 1163–1194
COPYRIGHT 2018, MICHAEL G. FOSTER SCHOOL OF BUSINESS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195
doi:10.1017/S0022109018000029

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Do Antitakeover Provisions Spur Corporate Innovation? A Regression Discontinuity Analysis

Thomas J. Chemmanur and Xuan Tian*

Abstract

We study the effect of antitakeover provisions (ATPs) on innovation. To establish causality, we use a regression discontinuity approach that relies on locally exogenous variation generated by shareholder proposal votes. We find a positive, causal effect of ATPs on innovation. This positive effect is more pronounced in firms that are subject to a larger degree of information asymmetry and operate in more competitive product markets. The evidence suggests that ATPs help nurture innovation by insulating managers from short-term pressures arising from equity markets. Finally, the number of ATPs contributes positively to firm value for firms involved in intensive innovation activities.

I. Introduction

Innovation is a key corporate strategy for the long-run growth and competitive advantage of firms (Porter (1992)). Unlike routine tasks, such as mass production and marketing, innovation activities involve a long process that is full of uncertainties and with a high probability of failure (Holmstrom (1989)). Therefore, motivating innovation needs significant tolerance for failure in the short run

*Chemmanur (corresponding author), chemmanu@bc.edu, Boston College Carroll School of Management; Tian, tianx@pbcsf.tsinghua.edu.cn, Tsinghua University PBC School of Finance. We thank Jarrad Harford (the editor) and an anonymous referee for several valuable comments that helped to greatly improve this paper. We are grateful for comments and suggestions from Julian Atanassov, Ramin Baghai, Martijn Cremers, Jack He, Josh Lerner, Megan MacGarvie, Amit Seru, Jun Yang, seminar participants at Boston College and Indiana University, and conference participants at the 2012 American Finance Association annual meetings, the 2012 European Finance Association annual meetings, the 2011 Financial Intermediation Research Society conference, the 2011 conference on Economics of Entrepreneurship and Innovation, and the 2011 North American Econometric Society Winter Meetings. We thank Michael Flores and Fangzhou Liu for their research assistance. We remain responsible for all errors and omissions. Chemmanur acknowledges summer research support from Boston College and additional support from a Hillenbrand Distinguished Fellowship. Tian acknowledges financial support from the National Natural Science Foundation of China (Grant No. 71790591) and a Tsinghua University Research Grant (Grant No. 20151080451).

and reward for success in the long run (Manso (2011)).[1] However, equity markets are often accused of creating excessive pressure on managers and exacerbating managerial myopia. In this paper, we empirically analyze whether corporate governance mechanisms, such as antitakeover provisions (ATPs) that insulate managers from threats of hostile takeovers, help to solve the conflicts between pressures from short-term public markets and innovation activities which are in the long-run interest of the firm.

We develop our testable hypothesis based on two strands in the theoretical literature that produce contradictory predictions regarding how ATPs affect firm innovation. A stream of the theoretical literature suggests that firms innovate more when managers are insulated from the pressures from short-term equity market investors. Stein (1988) argues that shareholders cannot properly evaluate a manager's investment in long-run innovation due to information asymmetry. Therefore, outside shareholders tend to undervalue the stocks of firms investing in long-term (innovative) projects, which leads these firms to have a greater exposure to hostile takeovers. To protect themselves and current shareholders against such expropriation, managers tend to invest less in innovation and put more effort in routine tasks that offer quicker and more certain returns (corporate myopia). Chemmanur and Jiao (2012) develop a formal theoretical analysis demonstrating that ATPs such as dual-class share structures allow high-talent firm managers to undertake long-term (innovative) projects rather than short-term projects, thus mitigating the above corporate myopia problem.[2] This theoretical literature, thus, implies that ATPs enhance corporate innovation. We will refer to this hypothesis as the "long-term value creation" hypothesis.

Another strand in the theoretical literature produces the opposite implication. Moral hazard models suggest that managers who are not properly monitored shirk or tend to invest suboptimally in routine tasks with quicker and more certain returns to enjoy private benefits, thus reducing firm value. The existence of hostile takeovers serves as an effective disciplining mechanism to mitigate this moral hazard problem and encourage innovation, thus increasing firm value. In this setting, the theoretical literature emerging from the seminal works of Grossman and Hart (1988) and Harris and Raviv (1988), (1989) implies that ATPs serve mainly to entrench incumbent firm management, thus reducing the effectiveness of equity market discipline on them and, consequently, reducing

[1]See also Ferreira, Manso, and Silva (2014), who argue that insiders of private firms (which are less transparent to outside investors compared to public firms) are more tolerant of failures and, thus, more inclined to choose innovative projects, unlike the insiders of public firms, who choose more conventional (safer) projects that allow them to cash in early following the arrival of good news.

[2]See also Shleifer and Summers (1988), who suggest that incumbent managers have less power over shareholders when takeover threats are high, so they invest less effort and human capital in activities that pay off only in the long run (e.g., innovation). In other words, incumbent managers invest less in innovation or other long-term projects because they fear that a hostile bidder will dismiss them after a successful takeover (when the innovation is completed), so they are unable to enjoy any of the profits resulting from their investment in innovation.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

corporate innovation (lowering firm value).[3] We refer to this as the "management entrenchment" hypothesis.

We test the above two competing hypotheses by examining whether ATPs nurture or impede firm innovation. We use firm patent information to construct our main measures of firm innovation. Our ordinary least squares (OLS) regressions show that firms with a greater number of ATPs are significantly more innovative: Firms with a larger number of ATPs not only generate more patents but also generate patents with higher quality. Our OLS results are robust to alternative subsamples, alternative ATP measures, and alternative econometric models.

While our OLS results suggest a positive association between the number of ATPs and innovation output, identifying the causal effect of ATPs on firm innovation is difficult because the number of ATPs a firm has is endogenous.[4] To establish causality, we use a regression discontinuity (RD) design that relies on "locally" exogenous variation in the number of ATPs a firm has generated by governance proposal votes that pass or fail to pass by a small margin of votes in annual meetings. This RD approach is a powerful and appealing identification strategy: For these close-call votes, passing is very close to a random, independent event and, hence, is unlikely to be correlated with unobservable firm characteristics. This nearly randomized variation in the number of ATPs is generated by voters' inability to precisely control the assignment variable (votes) near the known cutoff that determines the vote outcomes (Lee and Lemieux (2010)).

After performing diagnostic tests to ensure that the key identifying assumption of the RD approach (i.e., there is no precise manipulation of votes around the known threshold) is not violated, we show a positive, causal effect of ATPs on firm innovation. According to our baseline RD regression estimation, passing a governance proposal that intends to decrease the number of ATPs results in a decrease in patent counts and patent citations in the first 3 years after the vote. This result is generally robust to alternative bandwidths and alternative RD regression specifications and is absent at artificially chosen thresholds (other than the true threshold) that determine the governance proposal vote outcomes.

Next, we look for further evidence for identification in the cross section by exploring how cross-sectional heterogeneity in firm characteristics alters the marginal effect of ATPs on innovation in the RD framework. Specifically, we find that the effect of ATPs on firm innovation is more pronounced for firms that are subject to a larger degree of information asymmetry and that operate in more competitive product markets. These findings are consistent with the implication of the

[3]The previous models consider a setting where the incumbent management of a firm (large shareholder) obtains not only cash flow or "security" benefits (arising from their equity ownership in the firm) but also private benefits from being in control; outside shareholders receive only security benefits. These models conclude that ATPs are value reducing, since they reduce the chance of takeovers by rival management teams who can increase the cash flows to current shareholders by managing the firm better than the incumbent. Thus, under these theories, ATPs are inefficient, and the only role of such provisions is to entrench existing management and, thereby, reduce the chance of their losing control. See also Cary (1969) and Williamson (1975), who made earlier, more informal arguments that ATPs act primarily to entrench incumbent management.

[4]For example, the number of ATPs and innovation could be jointly determined by unobservable firm characteristics, making correct inferences difficult to draw. There could also be a reverse causality in which the level of expected future innovation affects the current number of ATPs.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

long-term value creation hypothesis that ATPs provide insulation for managers against pressures from short-term public market investors and allow managers to focus on investing in innovative activities. Therefore, ATPs are more valuable when information asymmetry is more severe and when short-term competitive pressures are greater.

In the final part of the paper, we explore the valuation effect of ATPs through their effect on innovation. Starting with Gompers, Ishii, and Metrick (2003), there has been considerable debate on the effect of ATPs on various aspects of firm performance and, thereby, on shareholder value. For example, while Gompers et al. show that firms with a larger number of ATPs have lower long-term stock returns, Core, Guay, and Rusticus (2006) question the previous findings, arguing that there is no conclusive evidence indicating that ATPs affect actual firm performance. Johnson, Karpoff, and Yi (2015) find that ATPs help to improve IPO firms' valuation and subsequent operating performance. One possibility is that ATPs destroy shareholder value in a subset of firms, while they are value-neutral or even value-enhancing in others. Consistent with this conjecture, we find that ATPs contribute positively to firm value, but only if the firm is involved in intensive innovation activities and has high innovation output. If, however, ATPs are adopted while firms are not conducting a significant extent of innovation activities, ATPs reduce firm value, which is consistent with the findings of Gompers et al. Our findings suggest that the role of ATPs may be more nuanced than that indicated by the bulk of the literature so far: While adopting these ATPs is optimal (value-enhancing) for innovative firms, it is suboptimal (value-reducing) for firms that are not engaged in a significant extent of innovation.

The rest of the paper is organized as follows: Section II discusses the related literature. Section III describes sample selection and reports summary statistics. Section IV provides our main results relating ATPs and firm innovation. Section V presents our results regarding the impact of ATPs on firm value through innovation. Section VI discusses some limitations of our analysis and presents some policy implications. Section VII concludes.

II. Relation to the Existing Literature

Our paper contributes to two strands in the existing literature. The first strand is the one on corporate innovation. Recent empirical research testing the implications of Manso (2011) includes Ederer and Manso (2012), who conduct a controlled laboratory experiment; Azoulay, Graff Zivin, and Manso (2011), who exploit key differences across funding streams within the academic life sciences; and Tian and Wang (2014), who show that IPO firms financed by more failure-tolerant venture capital investors are more innovative. All studies provide supporting evidence for the predictions of Manso.

The existing empirical literature shows that various firm and industry characteristics affect managerial incentives for investing in innovation. A larger institutional ownership (Aghion, Van Reenen, and Zingales (2013)), lower stock liquidity (Fang, Tian, and Tice (2014)), hedge fund activism (Brav, Jiang, Ma, and Tian (2018)), private instead of public equity ownership (Lerner, Sorensen, and Stromberg (2011)), and corporate rather than traditional venture capitalists

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

(Chemmanur, Loutskina, and Tian (2014)) alter managerial incentives and, hence, help nurture innovation. Other studies show that product market competition, market conditions, firm boundaries, financial market development, banking competition, chief executive officer overconfidence, labor unions, bank interventions, customer feedback, and financial analysts all affect firm innovation (Aghion, Bloom, Blundell, Griffith, and Howitt (2005), Hirshleifer, Low, and Teoh (2012), Cornaggia, Mao, Tian, and Wolfe (2015), He and Tian (2013), Nanda and Rhodes-Kropf (2013), Hsu, Tian, and Xu (2014), Seru (2014), Bradley, Kim, and Tian (2017), Chu, Tian, and Wang (2018), and Gu, Mao, and Tian (2018)).[5] Our paper contributes to this empirical literature by documenting that ATPs provide some insulation from the takeover market and spur innovation in public firms.

A few recent papers examine research questions somewhat more closely related to ours. Atanassov (2013) uses the enactment of Business Combination laws as a proxy for the decrease in the threat of hostile takeovers and finds that state antitakeover laws stifle innovation. Sapra, Subramanian, and Subramanian (2014) show a U-shaped relationship between innovation and takeover pressures. They also use the state-level antitakeover laws as the proxy for takeover. However, Karpoff and Wittry (2015) point out that using state ATPs to identify empirical tests is problematic, since, in many cases, these state ATPs did not raise the barriers to takeover for firms incorporated in these states. Rather, changes in state ATPs first reduced such barriers to takeover in prior years before increasing them. Therefore, comparing innovations before and after the passage of these state antitakeover laws to identify empirical tests on the effect of ATPs on innovation may yield misleading results. Different from Atanassov (2013) and Sapra et al. (2014), we use firm-level ATP information and the RD approach to tackle the research question. Our paper, thus, is the first one that studies the causal link between the number of ATPs at the firm level and corporate innovation. It is also worth noting that the broader conclusions that can be drawn from our paper regarding the relation between ATPs and innovation are opposite to those implied by the above two papers using state-level antitakeover laws.[6]

Our paper also contributes to the debate on the effect of ATPs on firm value and on long-term stock returns. Gompers et al. (2003) show that firms with a greater number of ATPs have lower stock returns. However, Core et al. (2006) question the findings of Gompers et al. and argue that there is no conclusive evidence suggesting that having a larger number of ATPs leads to poorer stock returns. Bebchuk and Cohen (2005) find that staggered boards are associated with a reduction in firm value. Bebchuk, Cohen, and Ferrell (2009) show that 6 ATPs deserve the most attention, and they are associated with reductions in firm valuation and stock returns. Cremers and Ferrell (2014) use a new hand-collected data set and find a robustly negative association between the number of ATPs and firm value over the 1978–2006 period. In contrast to the previous papers, Chemmanur,

[5]He and Tian (2017) provide a survey of this literature.

[6]A contemporaneous work by Becker-Blease (2011) finds a positive association between the presence of governance provisions and corporate innovation for a sample of firms between 1984 and 1997. He further shows that coverage by state-level antitakeover legislations is typically unassociated or negatively associated with innovation. However, it is hard to draw causal inferences from his analysis, and he does not explore the valuation effect of ATPs through corporate innovation.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Paeglis, and Simonyan (2011) show that firms with higher quality management teams tend to adopt a larger number of ATPs and that firms with a combination of higher quality management teams and a larger number of ATPs outperform the other groups of firms in their sample. Johnson et al. (2015) find that ATPs help to bond the IPO firm's commitments to its business partners, such as customers, suppliers, or strategic partners, which improves the IPO firm's valuation and subsequent operating performance. Our paper contributes to the previous debate by showing a new channel, namely, corporate innovation, through which ATPs positively contribute to firm value. However, we also show that, if ATPs are adopted by noninnovative firms (the vast majority of firms fall into this category), they are associated with a reduction in firm value. Thus, our evidence is consistent with the findings of Gompers et al. that ATPs lower subsequent stock returns, but it sheds further light on their findings by distinguishing between the kinds of firms in which ATPs reduce firm value and those in which they increase firm value.

III. Sample Selection and Summary Statistics

The sample examined in this paper includes 3,474 publicly traded firms with ATP information available during the period 1990–2006. We combine innovation data from the National Bureau of Economic Research (NBER) Patent Citation database, ATPs and shareholder proposal vote information from RiskMetrics, firms' balance sheet data from Compustat, analyst coverage data from Thomson Reuters Institutional Brokers' Estimate System (IBES), and institutional ownership data from the Thomson Financial 13F institutional holdings database.

A. Measuring Innovation

The innovation variables are constructed from the latest version of the NBER Patent Citation database created by Hall, Jaffe, and Trajtenberg (2001), which contains updated patent and citation information from 1976 to 2006. The NBER Patent Citation database provides annual information regarding patent assignee names, the number of patents, the number of citations received by each patent, the technology class of the patent, the patent application year, and the patent grant year. As suggested by the innovation literature (e.g., Griliches, Hall, and Pakes (1987)), the application year is more important than the grant year since it is closer to the time of the actual innovation. We then construct innovation variables based on the patent application year.

The NBER patent database is subject to two types of truncation problems. We follow the innovation literature to correct for the truncation problems in the NBER Patent Citation database. The first type of truncation problem regards patent counts, because the patents appear in the database only after they are granted and the lag between patent applications and patent grants is significant (approximately 2 years on average). As we approach the last few years for which there are patent data available, we observe a smaller number of patent applications that are eventually granted. This is because many patent applications filed during these years were still under review and had not been granted by 2006. Following Hall et al. (2001), we correct for the truncation bias in patent counts by first estimating the application-grant lag distribution for the patents filed and granted

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

between 1995 and 2000. This estimation is done by calculating the time interval in years between a patent's application year and its grant year. We define W_s, the application-grant lag distribution, as the percentage of patents applied for in a given year that are granted in s years. We then compute the truncation-adjusted patent counts, P_{adj}, as $P_{\text{adj}} = P_{\text{raw}} / \sum_{s=0}^{2006-t} W_s$, where P_{raw} is the unadjusted number of patent applications at year t and t is between 2001 and 2006.

The second type of truncation problem regards the citation counts: While patents keep receiving citations over a long period of time, we observe, at best, only the citations received up to 2006. Following Hall et al. (2001), we correct the truncation in citation counts by scaling up the citation counts using the variable "hjtwt" provided by the NBER database, which relies on the shape of the citation lag distribution.

The NBER patent database is unlikely to be subject to survivorship bias. An eventually granted patent application is counted and attributed to the applying firm at the time when the patent application is submitted even if the firm later gets acquired or goes bankrupt. In addition, patent citations attribute to a patent but not a firm. Hence, a patent assigned to an acquired or bankrupt firm can continue to receive citations for many years even after the firm goes out of existence.

We construct two measures for a firm's innovation output. The first measure is the truncation-adjusted patent count for a firm each year. Specifically, this variable counts the number of patent applications filed in a year that are eventually granted. One concern is that a simple count of patents may not distinguish breakthrough innovations from incremental technological discoveries. Griliches et al. (1987) show that the distribution of patents' value is very skewed; that is, most of the value is concentrated in a small number of patents. Therefore, we construct the second measure that intends to capture the importance of patents by counting the number of citations each patent receives in the subsequent years. To precisely capture the impact of patents, we exclude self-citations when we compute citations per patent. Given a firm's size and its innovation inputs, the number of patents captures its innovation quantity and the number of citations per patent captures its innovation quality.

One issue regarding using patent counts to capture a firm's innovation output has to do with the timing of patent application and research and development (R&D) expenditures. While there is significant variation in the time interval between R&D expenditures and patent application across different industries, the average lag is relatively short (e.g., typically less than 1 year). As a result, the existing economics literature (e.g., Hall, Griliches, and Hausman (1986), Hausman, Hall, and Griliches (1984), and Lerner and Wulf (2007)) argues that patent applications are generated nearly contemporaneously with R&D expenditures and uses 1-year-ahead patent applications to capture innovation output. Following the existing literature, we use 1-year-ahead patent counts and future citations received by these patents as our main dependent variable. To reflect the long-term nature of investment in innovation, we also measure both innovation proxies 2 and 3 years into the future.

Panel A of Table 1 presents the firm-year summary statistics of the innovation variables. On average, a firm in our sample has 7.4 granted patents per year and each patent receives 2.4 non-self-citations. Since the distributions of patent counts

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

TABLE 1
Summary Statistics

Table 1 reports the summary statistics for variables constructed based on the sample of U.S. public firms from 1990 to 2006. Panel A reports the summary statistics of innovation variables. Panel B reports the summary statistics of the percentage of votes that are for shareholder proposals. Panel C presents the year distribution of shareholder proposal voting. Panel D reports the industry distribution of firms with shareholder proposals. Panel E reports the summary statistics of antitakeover provision (ATP) variables as well as other control variables. SIC stands for Standard Industrial Classification. Variables are defined in the Appendix.

Panel A. Innovation Variables

Variable	Mean	Median	Std. Dev.	No. of Obs.
PATENTS	7.40	0	36.74	36,159
CPP	2.42	0	10.43	36,159

Panel B. Percentage of Votes That Are for Shareholder Proposals

Proposals	Mean	25th Percentile	Median	75th Percentile	Passage Rate	No. of Obs.
All	26.7%	7%	18%	43%	18.7%	4,787
ATP-related	50.9%	36%	52%	65%	52.4%	1,380
Close-call ATP	50.2%	45%	51%	55%	51.5%	233

Panel C. Year Distribution of Shareholder Proposal Voting

Year	All Proposals	ATP-Related Proposals	Close-Call ATP Proposals
1997	386	114	22
1998	373	119	19
1999	405	143	9
2000	391	126	31
2001	412	127	34
2002	439	147	29
2003	607	183	43
2004	592	139	23
2005	554	131	23
2006	628	151	0
Total	4,787	1,380	233

Panel D. Industry Distribution of Firms with Shareholder Proposals

SIC	Description	All Proposals	ATP-Related Proposals	Close-Call ATP Proposals
0	Agriculture	11	1	1
1	Mining	169	43	6
2	Light manufacturing	852	157	49
3	Heavy manufacturing	752	206	49
4	Transportation	682	236	58
5	Wholesale trade	437	109	36
6	Finance	516	140	11
7	Services	214	69	20
8	Health services	28	9	0
9	Public administration	90	12	3

Panel E. Independent Variables

Variable	Mean	Median	Std. Dev.	No. of Obs.
GINDEX	8.93	9.00	2.75	36,159
ASSETS (in US $billions)	8.77	1.20	46.20	33,233
ROA (%)	12.14	6.22	10.13	32,300
R&D/ASSETS (%)	1.65	0.00	7.11	36,159
PPE/ASSETS (%)	58.82	51.81	40.47	29,890
LEVERAGE (%)	25.59	23.51	20.34	32,956
CAPEXP/ASSETS (%)	5.88	4.53	5.67	30,436
HHI	0.21	0.09	0.27	36,159
TOBINS_Q	1.76	1.35	1.37	32,545
KZ_INDEX	0.90	0.84	6.96	27,080
INST_OWNERSHIP (%)	35.29	35.23	33.45	36,159
DISPERSION (%)	8.89	4.50	13.58	22,129

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

and citations per patent are highly right skewed, we use the natural logarithm of patent counts and citations per patent as the main innovation measures in our analysis. To avoid losing firm-year observations with 0 patents or citations, we add 1 to the actual value when calculating the natural logarithm.

B. Measuring ATPs and Control Variables

The number of ATPs a firm has in its corporate charter and bylaws is collected from RiskMetrics, which publishes an index compiled by Gompers et al. (2003). There have been eight publications of this index (1990, 1993, 1995, 1998, 2000, 2002, 2004, and 2006). They include detailed information on the ATPs of approximately 1,500 firms in each of the 8 publication years, with more firms added in recent publications. The index (defined as the GINDEX hereafter) counts how many ATPs the firm has and consists of 24 different provisions in five categories.[7] Following Gompers et al. (2003), we assume that, during the years between two consecutive publications, firms have the same ATPs as in the previous publication year. Therefore, we replace the missing values of the GINDEX with the values reported in the previous publication year. Panel E of Table 1 reports the descriptive statistics of the ATP measures. On average, a firm has 8.9 ATPs each year. The median GINDEX is 9.

To implement the RD analysis, we also collect data from RiskMetrics for 9,082 shareholder proposals from 1997 to 2006.[8] Since the percentage of votes that are for or against the proposal is critical in our RD analysis, we exclude proposals with this information missing. As a result, we end up with 4,787 shareholder proposals. Among them, we identify 1,380 proposals that are related to changing the number of ATPs a firm has. RiskMetrics provides information on the meeting date, the percentage of votes in favor of the proposal, and the proponent. Because all proposals from RiskMetrics are shareholder sponsored, these ATP-related proposals are intended to decrease the number of ATPs a firm has in its corporate charter, thus increasing shareholder power. For example, these shareholder proposals involve issues such as repealing classified boards, redeeming a poison pill, or eliminating a supermajority provision.

In the RD analysis, we mainly focus on proposals with shareholder votes falling in a narrow band around the vote threshold. The choice of bandwidth reflects a trade-off between precision and bias. Using a wider bandwidth includes more observations and yields more precise estimates. However, a wider bandwidth can introduce more noise and, hence, bias the estimates. Following the existing literature (e.g., Cuñat, Gine, and Guadalupe (2012)), we define a "close-call" sample as one that includes ATP-related shareholder proposals with a percentage of votes for proposals falling in a 10-percentage-point margin around the threshold. There are 233 proposals in this close-call sample.

[7] See Gompers et al. (2003) for a detailed description of the individual provisions.

[8] Rule 14a-8 allows shareholders to submit proposals that request certain corporate matters to be put to a vote at the firm's next annual shareholder meeting. To be eligible to submit a shareholder proposal, a shareholder needs to be a beneficial owner of at least 1% or $2,000 in market value of securities that are entitled to vote, have owned these securities for at least 1 year, and continue to own these securities through the date of the meeting.

It is worth noting that most of the vote outcomes are nonbinding because shareholder proposals are presented as a recommendation to the board of directors. However, even such nonbinding shareholder proposals, if passed, will impose greater pressure on firm management to decrease the number of ATPs in a firm's corporate charter compared to situations in which such shareholder proposals are not passed. This greater pressure on management for the reduction of ATPs upon passage of shareholder proposals recommending reduction of the number of ATPs to the board allows us to use them for the identification of the causal effect of ATPs on firm innovation in our RD framework, provided that the key identifying assumption of the RD analysis is satisfied. According to Lee and Lemieux (2010), the variation in a firm's number of ATPs is as good as that from a randomized experiment regardless of whether these proposals are binding or not if no one can precisely manipulate the forcing variable (i.e., the number of votes) near the known cutoff, the key RD identifying assumption. This is because, if this assumption is satisfied, the ex ante propensity of firms that barely pass and that of firms that barely fail to pass the vote to implement a reduction in ATPs should be identical. As a result, even when these shareholder proposals are not binding, it would not affect our identification of a causal effect here. We later do a number of diagnostic tests to ensure the satisfaction of the previous key RD identifying assumption in our setting.

Panel B of Table 1 reports the summary statistics for the percentage of shares that vote for various shareholder proposals. In the full sample, an average proposal receives 26.7% of shareholder support. The median percentage vote for these proposals is 18%. As a result, 18.7% of proposals are passed. In the ATP-related proposals, the passage rate is much higher: The passage rate is 52.4% and the median percentage vote for a proposal is 52%. In the close-call sample, the passage rate is 51.5%, with the mean percentage vote for the proposals being 50.2%.

Panel C of Table 1 presents the frequency of all shareholder proposals, ATP-related proposals, and proposals in the close-call sample. Consistent with Cuñat et al. (2012), there is a significant increase in the number of shareholder proposals from 2003 onward (over 550 per year). Because we use 1-year-ahead patent variables as our dependent variables and the patent data end in 2006, our close-call ATP sample ends in 2005.

Panel D of Table 1 displays the industry distribution (based on 1-digit Standard Industrial Classification (SIC) codes) of firms with shareholder proposal votes in the full sample, the ATP-related proposal sample, and the close-call sample. The bulk of proposal votes are concentrated in the manufacturing industry (1-digit SIC codes of 2 and 3, light and heavy manufacturing, respectively), followed by the transportation industry (1-digit SIC code of 4) and the wholesale trade industry (1-digit SIC code of 5).

Following the innovation literature, we control for a vector of firm and industry characteristics that may affect a firm's innovation productivity. In OLS regressions, our control variables include firm size (measured by the natural logarithm of total assets), profitability (measured by return on assets (ROA)), investments in innovation (measured by R&D expenditure over total assets), asset tangibility (measured by net property, plant, and equipment (PPE) scaled by total assets), leverage, capital expenditure,

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

product market competition (measured by the Herfindahl–Hirschman index (HHI) based on sales), growth opportunity (measured by Tobin's Q), financial constraints (measured by the Kaplan and Zingales (1997) 5-variable KZ index), and institutional ownership. To capture the possible nonlinear relation between product market competition and innovation (Aghion et al. (2005)), we also include the HHI^2 in the OLS regressions. We provide detailed variable definitions in the Appendix.

Panel E of Table 1 reports the descriptive statistics of our control variables. An average firm has total assets of \$8.8 billion, ROA of 12.1%, leverage of 25.6%, Tobin's Q of 1.76, net PPE ratio of 58.8, and approximately 35% of equity shares held by institutional investors.

IV. ATPs and Innovation

In Section IV.A, we report the results from our OLS regressions and present a rich set of robustness checks. We then discuss the main results from our identification strategy, the RD approach, in Section IV.B. Section IV.C explores the cross-sectional heterogeneity in the effect of ATPs on innovation. Specifically, we examine how a firm's information environment and product market competition alter the effect of ATPs on innovation.

A. OLS Results

To examine how ATPs affect firm innovation, we estimate the following model using OLS regressions:

$$\text{(1)} \qquad \ln(\text{INNOVATION}_{i,t+n}) = \alpha + \beta \text{GINDEX}_{i,t} + \delta' Z_{i,t} + \text{YEAR}_t + \text{FIRM}_i + u_{i,t},$$

where i indexes firm, t indexes time, and n equals 1, 2, and 3. ln(INNOVATION) is the dependent variable and can be one of the following two measures: the natural logarithm of the number of patents filed by the firm, ln(PATENT), or the natural logarithm of the number of citations per patent, ln(CPP). Z is a vector of firm and industry characteristics that may affect a firm's innovation productivity, as discussed in Section III.B. YEAR_t captures fiscal year fixed effects. We control for time-invariant unobservable firm characteristics by including firm fixed effects, FIRM_i. We cluster standard errors at the firm level.[9]

We include firm fixed effects in our OLS regressions to mitigate the endogeneity concern that innovation output and the adoption of ATPs could be jointly determined by certain unobservable firm characteristics. As a result, unobservables may bias our coefficient estimates and make correct inferences difficult to draw. Including firm fixed effects helps to mitigate this concern if the unobservable firm characteristics correlated with both ATPs and innovation output are constant over time.

Table 2 reports our OLS regression results. We estimate equation (1) with ln(PATENT) as the dependent variable and report the regression results in

[9]Besides pooled OLS regressions, we use a Tobit model that takes into consideration the nonnegative nature of patent and citation data. We also run a Poisson regression when the dependent variable is the number of patents to take care of the discrete nature of patent counts. Our OLS results are robust to using the previous alternative econometric models.

TABLE 2

OLS Regressions

Table 2 reports the OLS regressions estimating equation (1). The dependent variables are the natural logarithm of the number of patents in a year (columns 1–3) and the natural logarithm of the number of citations per patent in a year (columns 4–6). The variable of interest is the GINDEX. Other independent variables include the natural logarithm of firm assets, return on assets, R&D scaled by firm assets, PPE scaled by firm assets, firm leverage, capital expenditure scaled by firm assets, the Herfindahl–Hirschman index (HHI), the HHI^2, Tobin's Q, the KZ index, and institutional ownership. Variables are defined in the Appendix. Coefficient estimates and standard errors clustered by firm are reported. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	ln(PATENT$_{t+1}$)	ln(PATENT$_{t+2}$)	ln(PATENT$_{t+3}$)	ln(CPP$_{t+1}$)	ln(CPP$_{t+2}$)	ln(CPP$_{t+3}$)
Variable	1	2	3	4	5	6
GINDEX	0.090***	0.087***	0.090***	0.044***	0.038***	0.035**
	(0.025)	(0.025)	(0.028)	(0.014)	(0.015)	(0.016)
ln(ASSETS)	0.296***	0.181***	0.058	0.053	−0.001	−0.036
	(0.065)	(0.067)	(0.071)	(0.037)	(0.038)	(0.041)
ROA	0.599***	1.062***	1.356***	0.525***	0.752***	0.676***
	(0.183)	(0.224)	(0.283)	(0.134)	(0.147)	(0.163)
R&D/ASSETS	1.276***	1.763***	2.147***	0.887***	1.110***	1.052***
	(0.376)	(0.370)	(0.401)	(0.206)	(0.194)	(0.213)
PPE/ASSETS	0.351**	0.250	0.323*	0.089	0.094	0.170
	(0.143)	(0.160)	(0.177)	(0.090)	(0.100)	(0.110)
LEVERAGE	−0.399***	−0.319**	−0.297**	−0.196***	−0.186**	−0.180**
	(0.114)	(0.127)	(0.144)	(0.071)	(0.079)	(0.088)
CAPEXP/ASSETS	1.134***	1.464***	1.367***	0.729***	0.784***	0.689***
	(0.381)	(0.422)	(0.483)	(0.231)	(0.253)	(0.257)
HHI	0.102	−0.019	−0.126	0.045	−0.076	−0.116
	(0.174)	(0.188)	(0.219)	(0.108)	(0.116)	(0.128)
HHI^2	−0.201	−0.074	0.006	−0.127	−0.025	0.016
	(0.247)	(0.267)	(0.304)	(0.150)	(0.162)	(0.182)
TOBINS_Q	0.075***	0.049**	−0.009	0.004	−0.016	−0.036***
	(0.023)	(0.023)	(0.023)	(0.012)	(0.012)	(0.012)
KZ_INDEX	0.002***	0.002***	0.003***	0.001***	0.001***	0.001***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
INST_OWNERSHIP	0.284***	0.340***	0.389***	0.208***	0.210***	0.143*
	(0.102)	(0.123)	(0.148)	(0.059)	(0.070)	(0.084)
Constant	−1.490***	−2.926***	−2.043***	0.320	−0.688**	−0.358
	(0.491)	(0.592)	(0.620)	(0.274)	(0.333)	(0.360)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	20,204	18,663	17,018	20,204	18,663	17,018
R^2	0.800	0.787	0.778	0.671	0.657	0.645

columns 1–3 of Table 2. The coefficient estimate of the GINDEX in column 1 of Table 2 is positive and significant at the 1% level, suggesting that having a larger number of ATPs is associated with a larger number of patents generated by the firm in the following year.

Because the innovation process generally takes longer than 1 year to come to fruition, we examine the impact of a firm's number of ATPs on its patenting in the subsequent years in columns 2 and 3 of Table 2, where the dependent variables are the natural logarithm of patent counts measured in 2 years ($t+2$) and in 3 years ($t+3$), respectively.[10] The coefficient estimates of the GINDEX are positive and significant at the 1% level in both columns, suggesting that the number of ATPs is positively related to the firm's innovation output 2 and 3 years in the

[10]In an untabulated analysis, we examine the impact of ATPs on innovation productivity in the contemporaneous year (t) and in 4 years ($t+4$), and the results are both quantitatively and qualitatively similar.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

future. More importantly, the magnitudes of the coefficient estimates are very much comparable to that in column 1 of Table 2 and remain stable over time (across various columns).

We control for a comprehensive set of firm characteristics that may affect a firm's innovation output. We find that firms that are larger, more profitable, and have lower leverage are more innovative. A larger R&D spending, which indicates a larger innovation input, is associated with more innovation output. Larger capital expenditures are also associated with higher innovation productivity. Further, consistent with the findings of Aghion et al. (2013), institutional ownership is positively associated with a firm's innovation output. However, industry competition does not appear to significantly affect firm innovation.

Columns 4–6 of Table 2 report the regression results estimating equation (1) with the dependent variable replaced with ln(CPP) to examine the effect of ATPs on innovation quality. The coefficient estimates of the GINDEX are positive and significant at either the 1% or 5% level, suggesting a positive association between ATPs and firm innovation output. Once again, we find that firms that are more profitable, have lower financial constraints, have lower leverage, invest more in R&D, and have higher institutional ownership are more likely to generate high impact patents.

To check the robustness of our OLS regression results, we conduct a rich set of robustness tests. First, we focus on the subsample of firms that has at least 1 patent in the sample period. This test is to address the concern that our results may be driven by the large number of firm-year observations with 0 patents. In an untabulated analysis, we continue to observe positive and significant coefficient estimates of the GINDEX in this subsample. For example, in this untabulated analysis, the coefficient estimate for the GINDEX is 0.129 (p-value < 0.001) in column 1 of Table 2, where $\ln(\text{PATENT}_{t+1})$ is the dependent variable, and 0.047 (p-value $= 0.015$) in column 4, where $\ln(\text{CPP}_{t+1})$ is the dependent variable. In other words, the magnitudes of the coefficient estimates of the GINDEX are larger in the subsample of firms with at least 1 patent than those estimated from the full sample, since innovation is more relevant in this subsample.

Second, while we cluster standard errors at the firm level in the OLS regressions, we cluster standard errors by both firm and year to check the robustness of these findings. In an untabulated analysis, the positive coefficient estimates of the GINDEX continue to be statistically significant, although the significance level drops to the 5% level.

Third, in addition to using the GINDEX as a continuous variable, we construct a dummy variable that equals 1 if the number of ATPs a firm has is higher than the sample median (i.e., 9), and 0 otherwise. Our OLS results are robust to this alternative empirical specification. For example, in this untabulated analysis, the coefficient estimate of the dummy variable is 0.313 (p-value < 0.001) in column 1 of Table 2 (analysis of the number of patents) and 0.172 (p-value < 0.001) in column 4 (analysis of the citations per patent).

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Finally, Bebchuk et al. (2009) suggest that 6 ATPs are most effective and are not influenced by the noise produced by the other ATPs.[11] Therefore, we follow the existing literature and define the index consisting of the above 6 provisions as the E-index and reestimate equation (1) with the main independent variable replaced with the E-index. In untabulated tests, the coefficient estimates of the E-index are positive and significant at the 1% level in all specifications.

Overall, the results reported in this section suggest that having a larger number of ATPs is positively associated with a firm's subsequent innovation output; that is, the firm not only generates more patents but also produces patents with higher impact. These findings are consistent with the long-term value creation hypothesis.

B. Identification

While our OLS results are consistent with the long-term value creation hypothesis, as discussed earlier, our empirical design may be subject to endogeneity concerns due to omitted variables. For example, management talent could be an example of an omitted variable. High-quality managers may tend to have a larger number of ATPs relative to low-quality managers (Chemmanur et al. (2011)). High-quality managers would also be more likely to engage in innovation activities that result in a larger innovation output. Thus, the number of ATPs will be positively correlated with the firm's innovation. However, having a larger number of ATPs, per se, may not lead to higher innovation output. There could also be a reverse causality concern; that is, the level of expected future innovation predicts the firm's current number of ATPs. Our strategy of including firm fixed effects in the OLS regressions partially mitigates the omitted variables concern if the unobservable firm characteristics biasing the results are constant over time. However, if the unobservable characteristics are time varying, including firm fixed effects is not adequate to control for the endogeneity problem. To establish causality, we use an RD approach that relies on the simple majority passing rule (50%) and explore the passage of ATP-decreasing governance proposals in annual shareholder meetings.

The RD approach relies on locally exogenous variation in ATPs generated by ATP-related shareholder proposal votes that pass or fail by a small margin of votes around the 50% threshold. Conceptually speaking, this empirical approach compares firms' innovation output subsequent to shareholder proposal votes that pass by a small margin to those shareholder proposal votes that pass by a small margin. It is a powerful and appealing identification strategy, because, for these close-call votes, passing is very close to a random, independent event that is unrelated to firm unobservable characteristics, and it provides a randomized variation in a firm's number of ATPs as a consequence of the RD design, which helps us to identify the causal effect of ATPs on firm innovation (Lee and Lemieux (2010)).[12] The fact that these shareholder proposals are not binding but

[11]The 6 provisions are staggered boards, limits to bylaw amendments, limits to charter amendments, supermajority requirements for mergers, poison pills, and golden parachutes.

[12]Another advantage of the RD design is that we do not have to include observable covariates, Z, in the analysis because the inclusion of covariates is unnecessary for identification (Lee and Lemieux (2010)).

are presented as a recommendation to the board of directors does not present a problem for ouridentification. This is because even such nonbinding shareholder proposals, if passed, will impose greater pressure on firm management to decrease the number of ATPs in a firm's corporate charter compared to the situation where such shareholder proposals are not passed.[13]

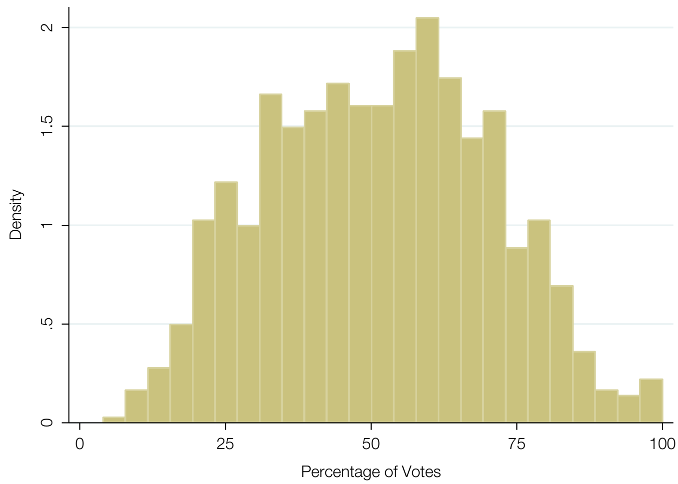
A key identifying assumption of the RD approach is that agents (both shareholders and managers in our setting) cannot precisely manipulate the forcing variable (i.e., the number of votes) near the known cutoff (Lee and Lemieux (2010)).[14] If this identifying assumption is not violated, the variation in the number of ATPs is as good as that from a randomized experiment. To check the validity of this assumption, we perform two diagnostic tests.

First, Figure 1 shows a histogram of the sample distribution of proposal vote shares in 25 equal-spaced vote share bins (with a bin width of 4%), and the

FIGURE 1

Density of Governance Proposal Vote Shares

Figure 1 plots a histogram of the distribution of the number of elections with the percentage of votes for unionizing in our sample across 25 equal-spaced bins (with a 4% bin width). The *x*-axis is the percentage of votes favoring an ATP-decreasing proposal. Governance proposals are obtained from RiskMetrics.



[13]One possible concern here is that, within the set of firms passing such nonbinding proposals by a small margin, the probability of implementing these proposals (leading to a reduction in the number of ATPs) may be correlated with management and firm factors, which, in turn, could also be correlated with their potential for innovation. However, it should be noted that, even if this were the case, this would not present a problem for identifying the causal effect, since, under our RD design, we are essentially comparing two ex ante identical groups of firms, one passing these proposals by a close-call vote and the other failing to pass these ATP-decreasing proposals by a close-call vote, and establishing that the former group of firms are more innovative on average. Of course, under the circumstances postulated here, there may be a selection effect as well as a treatment effect. We thank an anonymous referee for raising this concern.

[14]Lee (2008) shows that, even in the presence of manipulation, as long as firms do not have precise control over the forcing variable, an exogenous discontinuity still allows for random assignment to the treatment.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

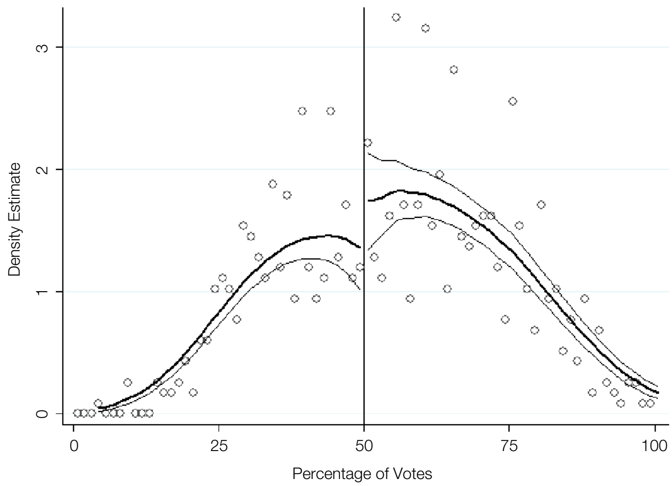
x-axis represents the percentage of votes favoring passing a governance proposal that intends to decrease the number of ATPs. If there is systematic sorting of firms within close proximity of the threshold, this sorting would be observed by a discontinuity in the vote share distribution at the 50% vote threshold. Figure 1 shows that the vote share distribution is continuous within close proximity of the cutoff; thus, no evidence of precise manipulation is observed at the cutoff point.

Second, we follow McCrary (2008) and provide a formal test of a discontinuity in the density. Using the 2-step procedure developed in McCrary (2008), Figure 2 plots the density of governance proposal vote shares.[15] The x-axis represents the percentage of votes favoring passing an ATP-decreasing governance proposal. The dots depict the density and the solid line represents the fitted density function of the forcing variable (i.e., the number of votes) with a 95% confidence interval around the fitted line.

FIGURE 2

Density of Governance Proposal Vote Shares

Figure 2 plots the density of governance proposal vote shares following the procedure in McCrary (2008). The x-axis is the percentage of votes favoring an ATP-decreasing proposal. The small circles depict the density estimate. The solid line represents the fitted density function of the forcing variable (the number of votes) with a 95% confidence interval around the fitted line. Governance proposals are obtained from RiskMetrics.



The density appears generally smooth and the estimated curve gives little indication of a strong discontinuity near the 50% threshold. The discontinuity estimate (the difference in height) is 0.247, with a standard error of 0.193. Therefore, we cannot reject the null hypothesis that the difference in density at the threshold is 0. As a result, it appears that the identifying assumption of the RD design that there is no precise manipulation by voters or managers at the known threshold is not violated.

Before we undertake rigorous RD regressions, we first plot the raw data in Figure 3 to visually check if there is a discontinuity in innovation output around

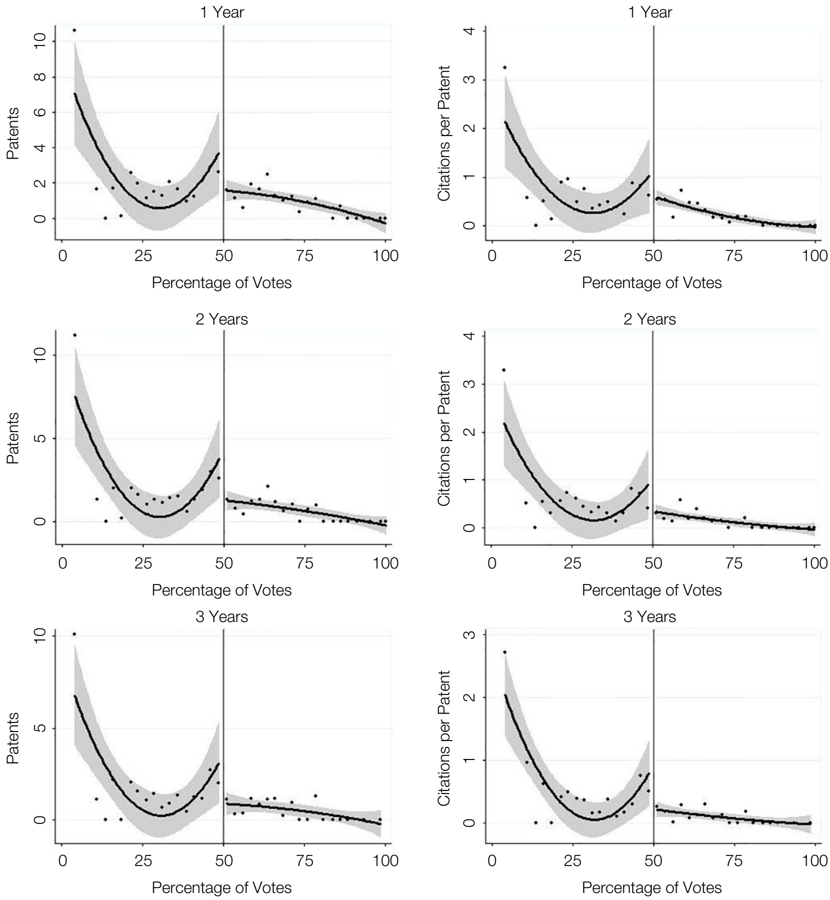
[15]See http://emlab.berkeley.edu/~jmccrary/DCdensity for a detailed discussion of the algorithm.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

FIGURE 3

Regression Discontinuity Plots

Figure 3 presents regression discontinuity (RD) plots using a fitted quadratic polynomial estimate with a 95% confidence interval around the fitted value. The x-axis is the percentage of votes favoring the ATP-decreasing proposals. Patent data are from the NBER Patent Citation database and Governance proposals are obtained from RiskMetrics.



the cutoff. The left-hand-side graphs present plots for the number of patents, and the right-hand-side plots present the number of citations per patent. The x-axis represents the percentage of votes for passing the governance proposals. In all plots displayed, firms that fail to pass the proposals are to the left of the 50% threshold and firms that succeed in passing the proposals are to the right of the threshold. The dots depict the average value of innovation outcome variables in the bins. The solid line represents the fitted quadratic polynomial estimate with a 95% confidence interval around the fitted value.

Figure 3 shows a discontinuity in both patent counts and patent citations at the threshold in each of the 3 years after the governance proposal votes. Specifically, within close proximity of the threshold, patent counts and citations drop significantly once the percentage of votes in favor of proposals that intend to

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

decrease the number of ATPs crosses the 50% cutoff point. This observation points to a positive, causal effect of ATPs on firm innovation.

We next move on to more rigorous RD tests. Specifically, we estimate the following model in the close-call sample and report the results in Table 3:

$$(2) \quad \ln(\text{INNOVATION}_{t+n}) = \alpha + \beta \text{PASS}_t + \delta' Z_t + \text{YEAR}_t + \text{IND}_i + u_{i,t},$$

where i indexes industry, t indexes time, and n equals 1, 2, and 3. ln(INNOVATION) can be one of the following two measures: the natural logarithm of the number of patents filed by the firm, ln(PATENT), or the natural logarithm of the number of citations received by each patent, ln(CPP). The key variable of interest is PASS, which equals 1 if the governance proposal passes, and, therefore, the number of ATPs increases, and 0 if the governance proposal fails to pass. Z is the same vector of firm and industry characteristics as that in equation (1). YEAR_t captures fiscal year fixed effects, and IND_i captures industry fixed effects.

In columns 1–3 of Table 3, where the dependent variable is patent counts, the coefficient estimates of PASS are all negative and statistically significant. The evidence suggests that passing a governance proposal that intends to reduce the number of ATPs leads to a decrease in innovation quantity. In columns 4–6 of Table 3, we replace the dependent variable with patent citations to capture innovation quality. We find negative coefficient estimates of PASS in all three columns and significant coefficients in the last two columns, in which future citations received by patents generated 2 and 3 years after the vote are considered, suggesting that a decrease in ATPs leads to a decrease in the quality for patents filed 2 and 3 years post the proposal vote.

After our baseline RD analysis, we perform three robustness checks to examine the sensitivity of our RD results. First, we repeat the baseline RD analysis in alternative "class-call" samples with smaller bandwidths, that is, a 5-percentage-point margin around the winning threshold, a 2-percentage-point margin around the winning threshold, and a 1-percentage-point margin around the winning threshold. The choice of bandwidths reflects a trade-off between bias and precision. On the one hand, a wider bandwidth makes use of more observations within the local neighborhood of the cutoff and, thus, yields more precise

TABLE 3

Regression Discontinuity Analysis

Table 3 presents the regression discontinuity (RD) results estimating equation (2). The dependent variables are innovation measures, and the variable of interest is a PASS dummy. The dependent variable in columns 1–3 is the natural logarithm of 1 plus patent counts, which measures innovation quantity. In columns 4–6, the dependent variable is the natural logarithm of 1 plus citation counts scaled by patents, which measures the quality of innovation. Patent data are from the NBER Patent Citation database, and governance proposals are obtained from RiskMetrics. Variables are defined in the Appendix. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	$\ln(\text{PATENT}_{t+1})$	$\ln(\text{PATENT}_{t+2})$	$\ln(\text{PATENT}_{t+3})$	$\ln(\text{CPP}_{t+1})$	$\ln(\text{CPP}_{t+2})$	$\ln(\text{CPP}_{t+3})$
Variable	1	2	3	4	5	6
PASS	−0.627***	−0.680**	−0.585*	−0.048	−0.164**	−0.164*
	(0.194)	(0.301)	(0.328)	(0.080)	(0.083)	(0.083)
No. of obs.	233	210	187	233	210	187
R^2	0.454	0.440	0.419	0.263	0.255	0.282

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

estimates. However, wider bandwidths may introduce more noise and bias into the estimation since tests with wider bandwidths use more "nonlocal" observations away from the cutoff. The converse problem arises for narrower bandwidths: on the negative side, the number of observations falling into the bandwidth becomes smaller, which reduces the test power; however, on the positive side, fewer nonlocal observations are used.

We report the results in Table 4. In Panel A, in which we focus on observations within the 5-percentage-point margin around the winning threshold, we continue to observe negative coefficient estimates of PASS. Here, 4 out of 6 coefficients are statistically significant. In Panels B and C, in which 2- and 1-percentage-point margins are used, respectively, we continue to observe generally negative coefficient estimates. However, these coefficients are not statistically significant. This observation is probably due to the fact that fewer observations fall into these narrower bandwidths, which reduces the power of these tests.

Second, we explore an alternative RD specification, a global polynomial series model (see, e.g., Cuñat et al. (2012)), that uses the entire support of all governance proposal votes in our sample. Specifically, we estimate the following model:

$$\text{(3)} \qquad \ln(\text{INNOVATION}_{t+n}) \quad = \quad \alpha + \beta \text{PASS}_t + P_l(v,c) + P_r(v,c) + \varepsilon_t,$$

where t indexes time and n equals 1, 2, and 3. The dependent variables are the same as those in equation (2). $P_l(v,c)$ is a flexible polynomial function for

TABLE 4

RD Analysis with Alternative Ways of Defining "Close-Call" Sample

Table 4 presents the RD regression results estimating equation (2). Panel A reports the results for observations falling in a 5-percentage-point margin around the winning threshold. Panel B reports the results for observations falling in a 2-percentage-point margin around the winning threshold. Panel C reports the results for observations falling in a 1-percentage-point margin around the winning threshold. The dependent variables are innovation measures, and the variable of interest is a PASS dummy. The dependent variable in columns 1–3 is the natural logarithm of 1 plus patent counts, which measures innovation quantity. In columns 4–6, the dependent variable is the natural logarithm of 1 plus citation counts scaled by patents, which measures the quality of innovation. Patent data are from the NBER Patent Citation database, and governance proposals are obtained from RiskMetrics. Variables are defined in the Appendix. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	$\ln(\text{PATENT}_{t+1})$	$\ln(\text{PATENT}_{t+2})$	$\ln(\text{PATENT}_{t+3})$	$\ln(\text{CPP}_{t+1})$	$\ln(\text{CPP}_{t+2})$	$\ln(\text{CPP}_{t+3})$
Variable	1	2	3	4	5	6
Panel A. 5-Percentage-Point Margin around the Winning Threshold						
PASS	−1.041**	−0.733*	−0.677	−0.330**	−0.350*	−0.180
	(0.374)	(0.339)	(0.436)	(0.142)	(0.169)	(0.138)
No. of obs.	145	132	122	145	132	122
R^2	0.539	0.499	0.495	0.338	0.356	0.356
Panel B. 2-Percentage-Point Margin around the Winning Threshold						
PASS	−0.528	−0.435	−0.570	−0.105	−0.185	−0.121
	(0.807)	(0.885)	(0.518)	(0.263)	(0.261)	(0.140)
No. of obs.	58	54	49	58	54	49
R^2	0.610	0.572	0.643	0.671	0.613	0.655
Panel C. 1-Percentage-Point Margin around the Winning Threshold						
PASS	0.700	−0.215	−0.270	−0.170	−0.012	−0.068
	(1.000)	(1.212)	(0.951)	(0.336)	(0.362)	(0.219)
No. of obs.	40	38	33	40	38	33
R^2	0.661	0.612	0.742	0.693	0.631	0.748

observations on the left-hand side of the threshold c with different polynomial orders; $P_r(v,c)$ is a flexible polynomial function for observations on the right-hand side of the threshold c with different polynomial orders; and v is a total vote share (percentage of votes in favor). Because governance proposal votes win with a simple majority of support among the voters, c equals 50% in our setting.

In this estimation, β is the key variable of interest and its magnitude is estimated by the difference in these two smoothed functions at the cutoff, which captures the causal effect of passing an ATP-decreasing shareholder proposal on firm innovation output 1, 2, and 3 years later. Note, however, that because RD estimates are essentially weighted average treatment effects, where the weights are the ex ante probability that the value of an individual proposal vote falls in the neighborhood of the win threshold (Lee and Lemieux (2010)), this coefficient should be interpreted locally in the immediate vicinity of the win cutoff.

We present the results estimating equation (3) in Table 5. We report the results with polynomials of order 2, but our results are qualitatively similar if we use other polynomial orders. The coefficient estimates of PASS are negative in all years and statistically significant 2 and 3 years after proposal votes for patent counts and 3 years after proposal votes for patent citations. These findings, once again, suggest a positive, causal effect of ATPs on innovation output. Hence, our baseline results are robust to an alternative RD specification.

TABLE 5

Alternative Regression Discontinuity Specification

Table 5 presents our results estimating the global polynomial model specified in equation (3). The dependent variables are innovation measures, and the variable of interest is a PASS dummy. The dependent variable in columns 1–3 is the natural logarithm of 1 plus patent counts, which measures innovation quantity. In columns 4–6, the dependent variable is the natural logarithm of 1 plus citation counts scaled by patents, which measures the quality of innovation. Patent data are from the NBER Patent Citation database, and governance proposals are obtained from RiskMetrics. Variables are defined in the Appendix. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	$\ln(\text{PATENT}_{t+1})$	$\ln(\text{PATENT}_{t+2})$	$\ln(\text{PATENT}_{t+3})$	$\ln(\text{CPP}_{t+1})$	$\ln(\text{CPP}_{t+2})$	$\ln(\text{CPP}_{t+3})$
Variable	1	2	3	4	5	6
PASS	−0.719	−0.784*	−0.909**	−0.065	−0.160	−0.309***
	(0.541)	(0.411)	(0.425)	(0.186)	(0.154)	(0.093)
No. of obs.	755	669	572	755	669	572
R^2	0.128	0.142	0.153	0.210	0.211	0.198

Third, we do placebo tests in which we artificially create alternative thresholds (other than the true threshold of 50%) to examine whether the "passage" of governance proposals around these artificial thresholds is related to a firm's post-vote innovation output. If the results reported in Table 3 are spurious, we should observe similar findings around these artificial thresholds as well. We do the placebo tests in the close-call sample, assuming 30% and 70% are the "win" thresholds, and report the results in Table 6. For consistency with our baseline results presented in Table 3, we use equation (2) to estimate these placebo tests as well.

In these placebo tests, the coefficient estimates of PASS have mixed signs. More importantly, none of them are statistically significant. We observe a similar finding in other artificially chosen thresholds. The evidence is consistent with

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

TABLE 6
Placebo Test: Artificially Assumed Thresholds

Table 6 presents the results of our baseline RD analysis but with artificially assumed thresholds. Panel A reports the results when 30% is assumed to be the threshold; Panel B reports the results when 70% is assumed to be the threshold. The dependent variables are innovation measures, and the variable of interest is a PASS dummy. The dependent variable in columns 1–3 is the natural logarithm of 1 plus patent counts, which measures innovation quantity. In columns 4–6, the dependent variable is the natural logarithm of 1 plus citation counts scaled by patents, which measures the quality of innovation. Patent data are from the NBER Patent Citation database, and governance proposals are obtained from RiskMetrics. Variables are defined in the Appendix. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	$\ln(PATENT_{t+1})$	$\ln(PATENT_{t+2})$	$\ln(PATENT_{t+3})$	$\ln(CPP_{t+1})$	$\ln(CPP_{t+2})$	$\ln(CPP_{t+3})$
Variable	1	2	3	4	5	6
Panel A. Artificial Threshold = 30%						
PASS	0.412	0.291	−0.269	−0.313	−0.297	−0.190
	(0.491)	(0.476)	(0.561)	(0.240)	(0.190)	(0.152)
No. of obs.	169	159	135	169	159	135
R^2	0.448	0.441	0.431	0.183	0.218	0.263
Panel B. Artificial Threshold = 70%						
PASS	−0.213	0.056	0.428	−0.003	0.037	0.083
	(0.528)	(0.488)	(0.412)	(0.101)	(0.104)	(0.081)
No. of obs.	168	144	119	168	144	119
R^2	0.250	0.215	0.131	0.194	0.124	0.080

the conjecture that the treatment effect of ATPs on firm innovation is absent at artificially chosen vote thresholds. It also suggests that the positive effects of ATPs on firm innovation that we documented earlier are unlikely to be driven by chance; therefore, our RD estimates are unlikely to be spurious.

Overall, in this section, we use an RD approach that explores locally exogenous variation in the number of ATPs a firm has incorporated into its corporate charter through ATP-related governance proposal votes to establish a causal link between ATPs and firm innovation output. We show a positive, causal effect of ATPs on innovation.[16] Our evidence is consistent with the implications of the long-term value creation hypothesis.

C. Cross-Sectional Analysis

In this section, we look for further evidence of identification in the cross section using the cross-sectional variation in firms' information environment, product market competition, and innovation difficulty to further explore the effect of ATPs on innovation. We first examine how heterogeneity in firms' information asymmetry affects the positive effect of ATPs on innovation in Section IV.C.1. We then examine how a firm's product market competition alters the effect of ATPs on innovation in Section IV.C.2. In this subsection, we undertake all the analyses in the RD framework estimating equation (2) in the close-call sample.

1. Information Asymmetry

A key assumption in the models of Stein (1988) and Chemmanur and Jiao (2012), on which the long-term value creation hypothesis is based, is the information asymmetry between firm insiders and outside investors. Due to information asymmetry, outside investors in a firm are unable to properly evaluate the firm

[16]Note, however, some of the limitations of our RD analysis that we discuss in Section VI.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

manager's investment in long-term innovation and tend to undervalue the equity of companies undertaking innovative projects (and, consequently, underinvest in such innovative projects). Therefore, a natural implication of the long-term value creation hypothesis is that the value of ATPs that insulate managers from short-term investor pressures will be larger if there is a larger information gap between firm insiders and outsiders. Therefore, we expect that the positive effect of ATPs on innovation is more pronounced for firms with a larger degree of information asymmetry.

Following previous literature, we construct a measure for a firm's information asymmetry, DISPERSION, which equals the average standard deviation of analyst earnings forecasts in a year. Firms with a higher standard deviation of analyst earnings forecasts are expected to have a larger extent of information asymmetry. We obtain analyst forecast information from the IBES database. We test the cross-sectional implication of the long-term value creation hypothesis by splitting the sample based on the median value of DISPERSION: a subsample of firms with more information asymmetry, with DISPERSION being above the sample median, and a subsample of firms with less information asymmetry, with DISPERSION being below the sample median. If the previous conjecture is supported, we expect to observe a much stronger effect of ATPs on innovation in the subsample of firms with more information asymmetry.

Table 7 reports these regression results. We report the results for larger information asymmetry firms in Panel A and the results for smaller information asymmetry firms in Panel B. In Panel A, the coefficient estimates of PASS are positive and statistically significant in all columns except for column 4. However, the coefficient estimates of PASS are mixed and statistically significant only in column 4 in Panel B. The magnitudes of coefficient estimates on PASS from Panel A are

TABLE 7

Information Asymmetry

Table 7 presents our subsample analysis (in the RD framework) of how information asymmetry affects the relation between ATPs and innovation. The dependent variables are innovation measures, and the variable of interest is a PASS dummy. The dependent variable in columns 1–3 is the natural logarithm of 1 plus patent counts, which measures innovation quantity. In columns 4–6, the dependent variable is the natural logarithm of 1 plus citation counts scaled by patents, which measures the quality of innovation. Panel A reports the results for the subsample for which the standard error of forecasts is above the sample median (more information asymmetry), and Panel B reports the results for the subsample for which the standard error of forecasts is below the sample median (less information asymmetry). Patent data are from the NBER Patent Citation database, and governance proposals are obtained from RiskMetrics. Variables are defined in the Appendix. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	$\ln(\text{PATENT}_{t+1})$	$\ln(\text{PATENT}_{t+2})$	$\ln(\text{PATENT}_{t+3})$	$\ln(\text{CPP}_{t+1})$	$\ln(\text{CPP}_{t+2})$	$\ln(\text{CPP}_{t+3})$
Variable	1	2	3	4	5	6
Panel A. More Information Asymmetry						
PASS	−1.579*** (0.453)	−1.868*** (0.515)	−1.310** (0.456)	−0.411 (0.253)	−0.390*** (0.106)	−0.271* (0.123)
No. of obs.	111	98	87	111	98	87
R^2	0.640	0.624	0.578	0.438	0.404	0.446
Panel B. Less Information Asymmetry						
PASS	0.324 (0.408)	0.434 (0.458)	0.342 (0.517)	0.324* (0.150)	0.102 (0.142)	0.096 (0.152)
No. of obs.	122	112	100	122	112	100
R^2	0.334	0.326	0.338	0.232	0.351	0.260

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

also much larger than those in Panel B. This evidence appears to suggest that the positive effect of ATPs on firm innovation output is more pronounced for firms that are subject to a larger degree of information asymmetry.

To check the robustness of this finding, we construct two alternative proxies for firm information asymmetry. The first one is a firm's analyst coverage, which equals the number of analysts following the firm in a year. Firms with larger analyst coverage are considered to have a smaller degree of information asymmetry. The second measure is a firm's analyst earnings forecast error, which equals the ratio of the absolute difference between the forecasted and actual earnings per share over the absolute actual earnings per share in a year. Firms with a higher analyst forecast error can be expected to have a greater extent of information asymmetry. In an untabulated analysis, we split the sample using these two alternative information asymmetry proxies and obtain qualitatively similar results.

Overall, our findings are consistent with the theoretical predictions of Stein (1988) and Chemmanur and Jiao (2012) and support the implications of the long-term value creation hypothesis: The effect of ATPs on firm innovation is more pronounced for firms with a larger degree of information asymmetry.

2. Product Market Competition

Product market competition increases a firm's pressure to keep competitive advantages over its rivals and make progress in generating profits in the short run to satisfy its equity market investors (see Aghion et al. (2013) for a similar argument). If ATPs indeed provide a shield for managers against short-term investors such that they can focus on innovation, we then expect that the positive effect of ATPs on innovation to be more pronounced when the firm is operating in a more competitive product market, namely, the pressures are higher and the insulation provided by ATPs is more needed and valued.

We test the cross-sectional implication of the long-term value creation hypothesis by splitting the sample based on the median value of the HHI: a subsample of firms in more competitive product markets with HHI below the sample median and a subsample of firms in less competitive product markets with HHI above the sample median. The economics and finance literature has widely used the HHI as a proxy for product market competition. The HHI is calculated by summing the square of each firm's market share (in sales) at the 4-digit SIC level. The HHI ranges between 0 and 1, with an increase in its value indicating a decrease in product market competition.

Table 8 reports our regression results. We report the results for more competitive product markets in Panel A and the results for less competitive product markets in Panel B. In Panel A, the coefficient estimates of PASS are negative and statistically significant in all columns except for column 6. However, in Panel B, the coefficient estimates of PASS have mixed signs and are statistically insignificant in all columns. The magnitudes of coefficient estimates on PASS from Panel A are also much larger than those in Panel B. This evidence appears to suggest that the positive effect of ATPs on firm innovation output is more pronounced for firms operating in more competitive product markets and is consistent with our conjecture that the effect of ATPs on firm innovation is larger when firms face tougher product market competition.

TABLE 8

Product Market Competition

Table 8 presents our subsample analysis (in the RD framework) of how product market competition affects the relation between ATPs and innovation. The dependent variables are innovation measures, and the variable of interest is a PASS dummy. The dependent variable in columns 1–3 is the natural logarithm of 1 plus patent counts, which measures innovation quantity. In columns 4–6, the dependent variable is the natural logarithm of 1 plus citation counts scaled by patents, which measures the quality of innovation. Panel A reports the results for the subsample in which the Herfindahl–Hirschman index (HHI) is below the sample median (more competitive product markets), and Panel B reports the results for the subsample in which the HHI is above the sample median (less competitive product markets). Patent data are from the NBER Patent Citation database, and governance proposals are obtained from RiskMetrics. Variables are defined in the Appendix. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable					
	ln(PATENT$_{t+1}$)	ln(PATENT$_{t+2}$)	ln(PATENT$_{t+3}$)	ln(CPP$_{t+1}$)	ln(CPP$_{t+2}$)	ln(CPP$_{t+3}$)
Variable	1	2	3	4	5	6
Panel A. More Competitive Product Markets						
PASS	−1.008**	−1.224**	−0.878*	−0.333**	−0.295*	−0.030
	(0.334)	(0.435)	(0.404)	(0.133)	(0.152)	(0.095)
No. of obs.	124	112	100	124	112	100
R^2	0.450	0.374	0.349	0.198	0.231	0.223
Panel B. Less Competitive Product Markets						
PASS	−0.147	0.099	0.042	−0.068	−0.026	−0.010
	(0.235)	(0.295)	(0.428)	(0.148)	(0.080)	(0.097)
No. of obs.	109	98	87	109	98	87
R^2	0.615	0.650	0.592	0.411	0.378	0.413

V. ATPs and Firm Value

Innovation enhances the long-run competitive advantages of firms. Ultimately, the objective of undertaking innovation is to increase firm value. As we have discussed in this paper, adopting ATPs is one of the strategies that a firm can use to insulate managers from short-term market pressures and enhance innovation. In this section, we examine the impact of innovation on firm value attributable to the firm's number of ATPs.

The existing literature has argued that ATPs reduce firm value largely because of management entrenchment (e.g., Gompers et al. (2003), Bebchuk and Cohen (2005), Bebchuk et al. (2009), and Cremers and Ferrell (2014)). However, another stream of literature has suggested that ATPs do not necessarily reduce firm value (e.g., Core et al. (2006)). We approach this debate by estimating variations of the following model in this section:

$$(4)\quad Q_{i,t+n} = \alpha + \beta_1 \times \text{GINDEX}_{i,t} + \beta_2 \times \text{GINDEX}_{i,t} \times \ln(\text{INNOVATION}_{i,t}) + \beta_3 \times \ln(\text{INNOVATION}_{i,t}) + \delta' Z_{i,t} + \text{YEAR}_t + \text{FIRM}_i + u_{i,t},$$

where i indexes firms, t indexes time, and n equals 1, 2, and 3. Tobin's Q is the dependent variable. ln(INNOVATION) can be one of the following two measures: the natural logarithm of the number of patents filed by the firm, ln(PATENT), or the natural logarithm of the number of citations received by each patent, ln(CPP). Similar to equation (1), we control for year and firm fixed effects to absorb any unobservable characteristics that vary only by year or by firm but cannot explain our results.

Panel A of Table 9 reports the regression results estimating equation (3) with patent counts as the proxy for innovation. In column 1 of Panel A, we estimate

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

a variation of equation (3) by excluding ln(PATENT) and the interaction term ln(PATENT) × GINDEX from the regression and control for industry instead of firm fixed effects. The purpose of this regression is to examine the direct effect of ATPs on firm value, which serves as a benchmark to which we are going to compare. Since industry fixed effects instead of firm fixed effects are included in the regression, the variation mainly comes from the cross section. Consistent with

TABLE 9

Regressions for Firm Value

Panel A of Table 9 reports the results of our OLS regressions estimating equation (4). The dependent variable is Tobin's *Q*. The main independent variable is the interaction between the GINDEX and the natural logarithm of the number of patents. Other independent variables include the natural logarithm of the number of patents, the GINDEX, the natural logarithm of firm assets, return on assets, R&D scaled by firm assets, PPE scaled by firm assets, firm leverage, capital expenditure scaled by firm assets, the Herfindahl–Hirschman index (HHI), the HHI^2, the KZ index, and institutional ownership. Panel B of Table 9 reports the results of our OLS regressions estimating equation (4). The dependent variable is Tobin's *Q*. The main independent variable is the interaction between the GINDEX and the natural logarithm of the number of citations per patent. Other independent variables include the natural logarithm of the number of citations per patent, the GINDEX, the natural logarithm of firm assets, return on assets, R&D scaled by firm assets, PPE scaled by firm assets, firm leverage, capital expenditure scaled by firm assets, the HHI, the HHI^2, the KZ index, and institutional ownership. Variables are defined in the Appendix. Standard errors are clustered by firm. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable				
	Q_{t+1}	Q_{t+1}	Q_{t+1}	Q_{t+2}	Q_{t+3}
Variable	1	2	3	4	5
Panel A. ln(PATENT)					
GINDEX	−0.017**	−0.024***	−0.013*	−0.009	−0.010
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
GINDEX × ln(PATENT)		0.005***	0.004***	0.004***	0.003***
		(0.001)	(0.001)	(0.001)	(0.001)
ln(PATENT)		−0.000	0.000***	0.000***	0.000**
		(0.000)	(0.000)	(0.000)	(0.000)
ln(ASSETS)	−0.047***	−0.078***	−0.447***	−0.430***	−0.350***
	(0.016)	(0.019)	(0.036)	(0.032)	(0.033)
ROA	4.091***	4.013***	1.774***	0.972***	0.355
	(0.821)	(0.823)	(0.493)	(0.257)	(0.224)
R&D/ASSETS	6.765***	6.559***	2.010***	0.629	−0.272
	(1.135)	(1.120)	(0.756)	(0.538)	(0.433)
PPE/ASSETS	−0.411***	−0.424***	0.007	0.097	−0.085
	(0.071)	(0.071)	(0.083)	(0.069)	(0.071)
LEVERAGE	0.012	0.051	0.365**	0.433***	0.504***
	(0.226)	(0.227)	(0.156)	(0.126)	(0.117)
CAPEXP/ASSETS	1.677**	1.602*	−0.235	−0.593*	−0.436
	(0.846)	(0.832)	(0.301)	(0.327)	(0.314)
HHI	0.117	0.095	0.062	0.011	0.079
	(0.132)	(0.132)	(0.100)	(0.100)	(0.099)
HHI^2	−0.139	−0.120	−0.044	−0.015	−0.084
	(0.143)	(0.142)	(0.105)	(0.110)	(0.111)
KZ_INDEX	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
INST_OWNERSHIP	0.104	0.097	0.162***	0.106***	0.065
	(0.066)	(0.066)	(0.038)	(0.039)	(0.046)
Constant	1.802***	2.032***	4.361***	4.326***	4.070***
	(0.376)	(0.361)	(0.271)	(0.244)	(0.241)
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	No	No	No
Firm fixed effects	No	No	Yes	Yes	Yes
No. of obs.	19,064	19,064	19,064	16,737	14,529
R^2	0.276	0.280	0.689	0.707	0.720

(continued on next page)

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

TABLE 9 (continued)
Regressions for Firm Value

	Dependent Variable			
	Q_{t+1}	Q_{t+1}	Q_{t+2}	Q_{t+3}
Variable	1	2	3	4
Panel B. ln(CPP)				
GINDEX	−0.021***	−0.009	−0.003	−0.006
	(0.007)	(0.007)	(0.007)	(0.012)
GINDEX × ln(CPP) (e-05)	0.283	0.563***	0.580***	0.475***
	(0.198)	(0.265)	(0.244)	(0.202)
ln(CPP)	0.020	0.020**	0.003	0.003
	(0.024)	(0.009)	(0.011)	(0.009)
ln(ASSETS)	−0.053***	−0.443***	−0.425***	−0.425***
	(0.020)	(0.036)	(0.055)	(0.031)
ROA	4.229***	1.768***	0.964**	0.964***
	(0.781)	(0.492)	(0.440)	(0.257)
R&D/ASSETS	7.039***	2.010***	0.634	0.634
	(1.062)	(0.756)	(0.813)	(0.539)
PPE/ASSETS	−0.458***	0.011	0.099	0.099
	(0.054)	(0.082)	(0.105)	(0.069)
LEVERAGE	0.032	0.358**	0.420*	0.420***
	(0.208)	(0.156)	(0.238)	(0.125)
CAPEXP/ASSETS	1.873**	−0.257	−0.610**	−0.610*
	(0.749)	(0.302)	(0.257)	(0.327)
HHI	0.047	0.069	0.020	0.020
	(0.144)	(0.100)	(0.122)	(0.100)
HHI^2	−0.115	−0.063	−0.035	−0.035
	(0.156)	(0.105)	(0.158)	(0.110)
KZ_INDEX	−0.001	−0.001	−0.001	−0.001
	(0.002)	(0.001)	(0.001)	(0.001)
INST_OWNERSHIP	0.024	0.162***	0.105*	0.105***
	(0.067)	(0.038)	(0.059)	(0.039)
Constant	1.813***	4.319***	4.292***	4.292***
	(0.197)	(0.270)	(0.414)	(0.243)
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	No	No	No
Firm fixed effects	No	Yes	Yes	Yes
No. of obs.	19,064	19,064	16,737	14,529
R^2	0.240	0.689	0.707	0.720

the findings of the existing literature, the coefficient estimate of ATPs in column 1 is negative and significant, suggesting that having a larger number of ATPs overall reduces firm value.

In column 2 of Panel A of Table 9, we add ln(PATENT) and ln(PATENT) × GINDEX to the regression. Once again, industry and year fixed effects are included in the regression. The coefficient estimate of the GINDEX is negative and significant at the 1% level, suggesting that the cross-sectional negative effect of ATPs on firm value in the following year is still present if a firm produces no patent. The coefficient estimate of ln(PATENT) is negative but not significant, suggesting that cross-sectional variation in innovation is not related to cross-sectional variation in firm value. However, the coefficient estimate of GINDEX × ln(PATENT) is positive and significant at the 1% level, suggesting that the negative effect of ATPs on firm value is mitigated by innovation. This finding is consistent with the prediction of the long-term value creation hypothesis that if ATPs are adopted to insulate managers against capital market pressures so

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

that managers can focus on long-run value creation activities, such as innovation, then ATPs contribute positively to firm value.

In column 3 of Panel A of Table 9, we repeat the regression in column 2 but control for firm fixed effects instead of industry fixed effects and examine how the time-series variation in the number of ATPs within a firm affects firm value over time. The coefficient estimate of the GINDEX is negative and significant, suggesting that a firm's adopting more ATPs is negatively related to the firm's market value in the subsequent year if the firm is not innovative and generates no patent. The coefficient estimate of ln(PATENT) is positive and significant at the 1% level, suggesting that innovation increases firm value in the following year, although the economic impact of patents is quite small. This finding is consistent with the existing literature that innovation increases firm value (e.g., Hall, Jaffe, and Trajtenberg (2005)). More importantly, the coefficient estimate of GINDEX $\times$ ln(PATENT) is positive and significant at the 1% level, suggesting that the negative effect of ATPs on firm value is largely mitigated by the firm's innovativeness. To be more concrete, the marginal effect of ATPs on firm value is initially negative but becomes positive when the value of ln(PATENT) is greater than 3.8 ($=0.013/0.004$). This is equivalent to the number of patents being greater than 48, which is at the 93rd percentile of the full sample but slightly below the median for the subsample for firms with at least 1 patent. The evidence is consistent with the implications of the long-term value creation hypothesis that ATPs positively contribute to firm value through innovation (for the top 7% firms that are most innovative). However, for the vast majority of firms that are not innovative (i.e., the remaining 93% of firms), ATPs reduce firm value.

In columns 4 and 5 of Panel A of Table 9, we estimate equation (3) with firm value in 2 years ($t+2$) and in 3 years ($t+3$) as the dependent variable, respectively. We continue to observe negative coefficient estimates of the GINDEX and positive and significant coefficients of the interaction term in columns 4 and 5.

Panel B of Table 9 reports the regression results with the number of citations each patent receives as the proxy for innovation quality. The regressions are parallel to those reported in Panel A, with model 1 of Panel A omitted. In column 1 of Panel B, in which industry and year fixed effects are included, we find a similar negative effect of GINDEX on firm value. In columns 2–4 of Panel B, we control for firm and year fixed effects and find positive and significant coefficient estimates of GINDEX $\times$ ln(CPP).

One concern is that our results could be driven by a large number of firm-year observations with 0 patents. We then rerun our firm value regressions estimating equation (3) in a subsample of firms that have at least 1 patent in the sample period. In an untabulated analysis, we find even stronger results. For example, the coefficient estimate of GINDEX in model 1 of Panel A of Table 9 is -0.028 (p-value $= 0.006$), suggesting that having a larger number of ATPs is associated with a reduction in firm value, even for innovative firms (recall that this analysis is for firms that generate at least 1 patent in the sample period). The coefficient estimates of GINDEX and GINDEX $\times$ ln(PATENT) in model (3) of Panel A are -0.029 (p-value $= 0.010$) and 0.004 (p-value < 0.001), respectively. Once again, the evidence suggests that the GINDEX positively contributes to firm value for very innovative firms, but it destroys firm value for less innovative firms.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

VI. Limitations of Our Study and Policy Implications

In the previous sections, we have shown, using both OLS regressions and our RD analysis, that ATPs have a positive effect on corporate innovation, and we have argued that this relation between ATPs and innovation is causal. Further, the analyses we present in Table 5 (alternative RD specification) and Table 6 (placebo tests using alternative thresholds) show that our RD analysis is fairly robust and unlikely to be driven by chance. However, our study is not without some limitations. In particular, we note that our RD results are weak in some specifications. Thus, while our RD results are strong when we follow the existing literature (e.g., Cuñat et al. (2012)) and define a close-call sample as consisting of those firms with a percentage of votes for ATP-reducing shareholder proposals falling in the 10-percentage-point margin around the winning threshold, our results become weaker when we narrow the close-call sample to only include firms with a percentage of votes for ATP-reducing shareholder proposals falling in a 5-percentage-point margin around the winning threshold (Table 4).[17] Our RD results weaken more (with the coefficient of the PASS variable becoming insignificant, though with the appropriate sign) when we further narrow the above bandwidth around the winning threshold to 2 percentage points or 1 percentage point (possibly because the number of observations falling into these narrower bands is very small, with the corresponding loss of test power).

Nevertheless, our paper makes an important contribution by presenting evidence opposing the result shown in papers, such as Atanassov (2013), based on somewhat problematic methodologies that ATPs dampen innovation (the underlying methodology in these papers is problematic for the reasons we discuss in Section II).[18] Our analysis implies that one has to take a more nuanced approach to the effect of ATPs on innovation: In particular, one has to distinguish between state-level and firm-level ATPs. Our empirical analysis showing that having a larger number of (and stronger) ATPs in a firm's corporate charter spurs innovation in many settings is especially important in the context of the fact that many innovative firms (especially in the technology sector) going public in the last decade have tended to incorporate a large number of (and stronger) ATPs in their corporate charter.[19] It is reassuring to find that the management of such firms may not have adopted such strong ATPs in their corporate charter purely with

[17]Our results are also somewhat weaker if we follow the alternative RD methodology of Imbens and Kalyanaraman (2012).

[18]The focus of Atanassov (2013) is on state ATPs. Karpoff and Wittry (2015) point out that using state ATPs to identify empirical tests is likely to yield misleading results (for the reasons we discuss in detail in Section II).

[19]Some prominent examples of such firms going public within the last decade with dual-class share structures are Google (now known formally as Alphabet Inc.), Facebook, and LinkedIn (these firms currently maintain their dual-class share structure). A dual-class share structure is considered to be a strong ATP, though several technology firms had other strong provisions, such as staggered boards, in their corporate charter (e.g., LinkedIn). While this is only anecdotal, Google Inc. is arguably one of the most innovative firms listed on the New York Stock Exchange (NYSE); further, having a dual-class share structure does not seem to have dampened Google's innovativeness over the past decade. In a letter to shareholders at the time their firm went public, the founders of Google, Sergei Brin and Larry Page, explicitly mentioned their desire to focus on the long run and a desire to manage the firm "without regard to the next quarter's profit numbers" as one justification for adopting a dual-class share structure (consistent with the long-run value creation hypothesis for which we find empirical support

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

the objective of engaging in shareholder-value-destroying activities. Thus, along with two IPO-related papers, Chemmanur et al. (2011), who show that, on average, firms with a combination of high management quality and a larger number of ATPs outperform all other sample firms in terms of IPO valuation and post-IPO operating performance, and Johnson et al. (2015), who show that firms with a larger number of ATPs have greater IPO valuations and post-IPO operating performance, our paper contributes to reopening the debate on the relation between the number of ATPs in a firm's corporate charter and the long-run performance of the firm.

In particular, our paper shows that, even for established public firms, ATPs improve long-run performance when performance is measured by productivity in real activities such as corporate innovation. Our analysis further implies that, by allowing firm managers to focus on long-run value creation without being too concerned about losing control of the firm in the short run, ATPs in a firm's corporate charter may benefit shareholders, even if, in some firms, they have an entrenchment effect as well. In particular, our analysis presented in Table 9 suggests that, far from destroying value, incorporating ATPs into their corporate charters may, in fact, enhance shareholder value in the case of highly innovative firms.

VII. Conclusion

We have studied the relation between ATPs and corporate innovation. We test two different hypotheses: i) the long-term value creation hypothesis, which implies that ATPs spur corporate innovation by insulating managers from short-term pressures arising from the equity market (or other related sources), allowing them to focus on long-term value creation; ii) the management entrenchment hypothesis, which implies that ATPs reduce innovation by mitigating the disciplining effect of the market for corporate control on firm managers.

To establish causality, we use an RD design relying on locally exogenous variation in the number of ATPs a firm has generated by governance proposals that pass or fail by a small margin of votes. Our identification strategies suggest a positive, causal effect of ATPs on firm innovation. The positive effect of ATPs on innovation is more pronounced in firms that are subject to a larger degree of information asymmetry and operate in more competitive product markets. The evidence is consistent with the hypothesis that ATPs spur innovation by insulating managers from short-term pressures arising from the equity market, allowing them to focus on long-term value creation. Finally, we show that the number of ATPs contributes positively to firm value for firms involved in intensive innovation activities but reduces firm value in other firms.

Appendix. Variable Definitions and Data Sources

Innovation Variables (Source: NBER Patent Citation Database)

PATENT_{it}: Number of patents firm i applied for in year t.

CPP_{it}: Number of citations per patent firm i applied for in year t.

here). Facebook founder Mark Zuckerberg gave a similar justification for adopting a dual-class share structure prior to Facebook's IPO.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

$\ln(\text{PATENT}_{it})$: Natural logarithm of 1 plus the number of patents firm i applied for in year t.

$\ln(\text{CPP}_{it})$: Natural logarithm of 1 plus the number of citations per patent firm i applied for in year t.

ATP Variables (Source: RiskMetrics)

GINDEX_{it}: Firm i's governance index based on 24 ATPs, taken from Gompers et al. (2003) in year t.

PASS_{i}: A dummy that equals 1 if an ATP-decreasing governance proposal passes, and 0 otherwise.

Firm Characteristics (Source: Compustat)

ASSETS_{it}: Total assets of firm i in year t (in $millions).

ROA_{it}: Operating income before depreciation to total assets ratio of firm i in year t.

$\text{R\&D/ASSETS}_{it}$: Research and development expenditure to total assets ratio of firm i in year t.

PPE/ASSETS_{it}: Net property, plant, and equipment to assets ratio of firm i in year t.

LEVERAGE_{it}: Total debt of firm i in year t divided by its total assets.

$\text{CAPEXP/ASSETS}_{it}$: Capital expenditure to total assets ratio of firm i in year t.

HHI_{it}: Herfindahl–Hirschman index of firm i's industry in year t constructed based on sales at 4-digit SIC industries.

TOBINS_Q_{it}: Market-to-book ratio of firm i in year t (total assets $+$ year end closing price $\times$ year end outstanding shares $-$ book equity)/total assets.

KZ_INDEX_{it}: Firm i's KZ index at year t is calculated as $-1.002 \times$ Cash flow $+ 0.28 \times$ Q $+ 3.18 \times$ Leverage $- 39.368 \times$ Dividends $- 1.315 \times$ Cash holdings.

$\text{INST\_OWNERSHIP}_{it}$: Total percentage of firm i's equity held by institutional investors in year t. *Source:* Thomson 13F institutional holdings database.

DISPERSION_{it}: Average standard deviation of analyst earnings forecast for firm i in year t. *Source*: IBES.

References

Aghion, P.; N. Bloom; R. Blundell; R. Griffith; and P. Howitt. "Competition and Innovation: An Inverted U Relationship." *Quarterly Journal of Economics*, 120 (2005), 701–728.

Aghion, P.; J. Van Reenen; and L. Zingales. "Innovation and Institutional Ownership." *American Economics Review*, 103 (2013), 277–304.

Atanassov, J. "Do Hostile Takeovers Stifle Innovation? Evidence from Antitakeover Legislation and Corporate Patenting." *Journal of Finance*, 68 (2013), 1097–1131.

Azoulay, P.; J. Graff Zivin; and G. Manso. "Incentives and Creativity: Evidence from the Academic Life Sciences." *RAND Journal of Economics*, 42 (2011), 527–554.

Bebchuk, L., and A. Cohen. "The Costs of Entrenched Boards." *Journal of Financial Economics*, 78 (2005), 409–433.

Bebchuk, L.; A. Cohen; and A. Ferrell. "What Matters in Corporate Governance?" *Review of Financial Studies*, 22 (2009), 783–827.

Becker-Blease, J. "Governance and Innovation." *Journal of Corporate Finance*, 17 (2011), 947–958.

Bradley, D.; I. Kim; and X. Tian. "Do Unions Affect Innovation?" *Management Science*, 63 (2017), 2251–2271.

Brav, A.; W. Jiang; S. Ma; and X. Tian. "How Does Hedge Fund Activism Reshape Corporate Innovation?" *Journal of Financial Economics*, forthcoming (2018).

Cary, W. "Corporate Devices Used to Insulate Management from Attack." *Antitrust Law Journal*, 39 (1969), 318–324.

Chemmanur, T. J., and Y. Jiao. "Dual Class IPOs: A Theoretical Analysis." *Journal of Banking and Finance*, 36 (2012), 305–319.

Chemmanur, T. J.; E. Loutskina; and X. Tian. "Corporate Venture Capital, Value Creation, and Innovation." *Review of Financial Studies*, 27 (2014), 2434–2473.
Chemmanur, T. J.; I. Paeglis; and K. Simonyan. "Management Quality and Antitakeover Provisions." *Journal of Law and Economics*, 54 (2011), 651–692.
Chu, Y.; X. Tian; and W. Wang. "Corporate Innovation along the Supply Chain." *Management Science*, forthcoming (2018).
Core, J.; W. Guay; and T. Rusticus. "Does Weak Governance Cause Weak Stock Returns? An Examination of Firm Operating Performance and Investors' Expectations." *Journal of Finance*, 61 (2006), 655–687.
Cornaggia, J.; Y. Mao; X. Tian; and B. Wolfe. "Does Banking Competition Affect Innovation?" *Journal of Financial Economics*, 115 (2015), 189–209.
Cremers, M., and A. Ferrell. "Thirty Years of Shareholder Rights and Firm Valuation." *Journal of Finance*, 69 (2014), 1167–1196.
Cuñat, V.; M. Gine; and M. Guadalupe. "The Vote Is Cast: The Effect of Corporate Governance on Shareholder Value." *Journal of Finance*, 67 (2012), 1943–1977.
Ederer, F., and G. Manso. "Is Pay-for-Performance Detrimental to Innovation?" *Management Science*, 59 (2012), 1498–1513.
Fang, V.; X. Tian; and S. Tice. "Does Stock Liquidity Enhance or Impede Firm Innovation?" *Journal of Finance*, 69 (2014), 2085–2125.
Ferreira, D.; G. Manso; and A. Silva. "Incentives to Innovate and the Decision To Go Public or Private." *Review of Financial Studies*, 27 (2014), 256–300.
Gompers, P.; J. Ishii; and A. Metrick. "Corporate Governance and Equity Prices." *Quarterly Journal of Economics*, 118 (2003), 107–155.
Griliches, Z.; B. Hall; and A. Pakes. "The Value of Patents as Indicators of Inventive Activity." In *Economic Policy and Technological Performance*, P. Dasgupta and P. Stoneman, eds. Cambridge, UK: Cambridge University Press (1987).
Grossman, S., and O. Hart. "One Share–One Vote and the Market for Corporate Control." *Journal of Financial Economics*, 20 (1988), 175–202.
Gu, Y.; C. Mao; and X. Tian. "Bank Interventions and Firm Innovation: Evidence from Debt Covenant Violations." *Journal of Financial Economics*, forthcoming (2018).
Hall, B. H.; Z. Griliches; and J. Hausman. "Patents and R and D: Is There a Lag?" *International Economic Review*, 27 (1986), 265–283.
Hall, B.; A. Jaffe; and M. Trajtenberg. "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools." Working Paper, National Bureau of Economic Research (2001).
Hall, B.; A. Jaffe; and M. Trajtenberg. "Market Value and Patent Citations." *RAND Journal of Economics*, 36 (2005), 16–38.
Harris, M., and A. Raviv. "Corporate Governance: Voting Rights and Majority Rules." *Journal of Financial Economics*, 20 (1988), 203–235.
Harris, M., and A. Raviv. "The Design of Securities." *Journal of Financial Economics*, 24 (1989), 255–287.
Hausman, J.; B. H. Hall; and Z. Griliches. "Econometric Models for Count Data with an Application to the Patents-R & D Relationship." *Econometrica*, 52 (1984), 909–938.
He, J., and X. Tian. "The Dark Side of Analyst Coverage: The Case of Innovation." *Journal of Financial Economics*, 109 (2013), 856–878.
He, J., and X. Tian. "Finance and Corporate Innovation: A Survey." Working Paper, University of Georgia (2017).
Hirshleifer, D.; A. Low; and S. Teoh. "Are Overconfident CEOs Better Innovators?" *Journal of Finance*, 67 (2012), 1457–1498.
Holmstrom, B. "Agency Costs and Innovation." *Journal of Economic Behavior and Organization*, 12 (1989), 305–327.
Hsu, P.; X. Tian; and Y. Xu. "Financial Market Development and Innovation: Cross-Country Evidence." *Journal of Financial Economics*, 112 (2014), 116–135.
Imbens, G., and K. Kalyanaraman. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies*, 79 (2012), 933–959.
Johnson, W.; J. Karpoff; and S. Yi. "The Bonding Hypothesis of Takeover Defenses: Evidence from IPO Firms." *Journal of Financial Economics*, 117 (2015), 307–332.
Kaplan, S., and L. Zingales. "Do Financial Constraints Explain Why Investment Is Correlated with Cash Flow?" *Quarterly Journal of Economics*, 112 (1997), 169–215.
Karpoff, J., and M. Wittry. "Test Identification with Legal Changes: The Case of State Antitakeover Laws." Working Paper, University of Washington (2015).
Lee, D. S. "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics*, 142 (2008), 675–697.

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029

Lee, D., and T. Lemieux. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48 (2010), 281–355.

Lerner, J.; M. Sorensen; and P. Stromberg. "Private Equity and Long-Run Investment: The Case of Innovation." *Journal of Finance*, 66 (2011), 445–477.

Lerner, J., and J. Wulf. "Innovation and Incentives: Evidence from Corporate R&D." *Review of Economics and Statistics*, 89 (2007), 634–644.

Manso, G. "Motivating Innovation." *Journal of Finance*, 66 (2011), 1823–1860.

McCrary, J. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142 (2008), 698–714.

Nanda, R., and M. Rhodes-Kropf. "Investment Cycles and Startup Innovation." *Journal of Financial Economics*, 110 (2013), 403–418.

Porter, M. "Capital Disadvantage: America's Failing Capital Investment System." *Harvard Business Review*, 70 (1992), 65–82.

Sapra, H.; A. Subramanian; and K. Subramanian. "Corporate Governance and Innovation: Theory and Evidence." *Journal of Financial and Quantitative Analysis*, 49 (2014), 953–1007.

Seru, A. "Firm Boundaries Matter: Evidence from Conglomerates and R&D Activity." *Journal of Financial Economics*, 111 (2014), 381–405.

Shleifer, A., and L. H. Summers. "Breach of Trust in Hostile Takeovers." In *Corporate Takeovers: Causes and Consequences*, A. J. Auerbach, ed. Chicago, IL: University of Chicago Press (1988).

Stein, J. "Takeover Threats and Managerial Myopia." *Journal of Political Economy*, 96 (1988), 61–80.

Tian, X., and T. Wang. "Tolerance for Failure and Corporation Innovation." *Review of Financial Studies*, 27 (2014), 211–255.

Williamson, O. "Markets and Hierarchies." In *Analysis and Antitrust Implications: A Study in the Economics of Internal Organization*, New York, NY: Free Press (1975).

Downloaded from https://www.cambridge.org/core. Tsinghua University, on 09 May 2019 at 03:02:15, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0022109018000029