Patent Success, Patent Holdup, and the Structure of Property Rights

Heng Geng*

Victoria University of Wellington

Harald Hau**

University of Geneva and Swiss Finance Institute

Sandy Lai***

National Taiwan University

January 21, 2018

Abstract

Innovation processes under patent protection generate holdup problems if complementary patents are owned by different firms (Hart, 1995). We show that shareholder ownership overlap across firms with patent complementarities helps mitigate such holdup problems and correlates significantly with higher patent investment and more patent success. The positive innovation effect is strongest for concentrated overlapping ownership and for the cases in which overlapping shareholders are dedicated investors, with long investment horizons and underdiversified portfolios.

JEL Classification: L22, G31, G32

Keywords: Patents, Holdup Problems, Innovation, Institutional Ownership

*School of Economics and Finance, Victoria University of Wellington, New Zealand. Tel.: (+64) 4 463 5523. E-mail: griffin.geng@vuw.ac.nz.

**University of Geneva, 40 Bd du Pont d'Arve, 1211 Genève 4, Switzerland. Tel.: (++41) 22 379 9581. E-mail: prof@haraldhau.com. Web page: http://www.haraldhau.com.

***College of Management, National Taiwan University, No.1, Sec. 4, Roosevelt Rd., Taipei City 106. Tel.: (+886) 2 33661087. E-mail: sandylai@ntu.edu.tw.

1 Introduction

Technological progress has been recognized as the main source of long-run economic growth (see, e.g., Solow, 1957; Hall, Jaffe, and Trajtenberg, 2005; Kogan, Papanikolaou, Seru, and Stoffman, 2017). However, the question of how corporate ownership structure and property rights in patents affect technological innovation remains relatively unexplored. This paper gives a new empirical perspective on the role of equity ownership structure in attenuating holdup problems induced by patent protection in the corporate innovation process.

Patent protection provides inventors with exclusive rights to the commercial use of their discoveries.[1] But such discoveries are often part of a larger technological process of interdependent innovations, and the full economic value of a patent might only be unlocked if the innovating firm can simultaneously secure access to many complementary patents. Therefore, patent processes generate a holdup problem whenever such complementary patents are owned by different firms and ex-ante contracting is incomplete.

When an ex-ante complete contract cannot be written, ex-post negotiation on the division of patent revenue surplus is required and such ex-post bargaining imposes two types of costs on an innovating firm, as emphasized by the transaction cost literature (Coase, 1937; Williamson, 1975, 1985). First, time and efforts spent in negotiating ex-post division of surplus create ex-post inefficiencies for the innovating firm because some of the resources are not put to productive use. Moreover, asymmetric information can lead to negotiation failure and subject the innovating firm to the risk of forgoing all its prior investment in the project. Second, because the downstream innovating firm fears that it will not recover its investment costs due to a potential holdup (in the form of either complete negotiation failure or excessive royalty fees) by upstream firms owning complementary patents, it underinvests in equilibrium, creating ex-ante inefficiencies for the firm.[2]

[1] Blundell, Griffith, and Reenen (1999) document that at least 4500 technologically important innovations were commercialized by British firms in the period 1945–1983.

[2] See the detailed discussion in Hart (1995). In particular, even if ex-post negotiation is efficient (i.e., no haggling or asymmetric information), the innovating firm might still underinvest relative to the first-best scenario. Consider a simple example as follows: A downstream innovating firm A needs a patent license from its upstream firm B for commercialization of its own patent. Assume that firms A's gross revenue from the patent project is $R(i)$, which is concave and increasing in its ex-ante investment i, and that the total cost of producing the upstream firm's patent is C, which was incurred prior to the start of firm A's patent project and is independent of i. Further assume that without the patent license from firm B, firm A would realize zero gross revenue. In the first-best world, the optimal investment i^* solves the problem of $R'(i) = 1$. Now, suppose firm A expects ex-post bargaining to result in a 50:50 split of ex-post gains between the two firms (by Nash bargaining). Firm A would optimally choose an investment level i^{**} that solves the problem of $\frac{1}{2}R'(i) = 1$; that is, underinvestment occurs ($i^{**} < i^*$).

The property rights literature (Grossman and Hart, 1986; Hart and Moore, 1990; Hart, 1995) suggests that *joint asset ownership* attenuates holdup problems under conditions of *asset specificity* and *ex-ante incomplete contracting.* The first condition (asset specificity) is fulfilled for many new downstream patents because the full economic and commercial value of these patents can often be realized only in conjunction with their upstream complementary patents. From a technological perspective, we can regard these upstream complementary patents as an essential input for generating the commercial success of the downstream patents. The second condition (ex-ante incomplete contracting) is also fulfilled. An ex-ante complete contract cannot be written because it is often unclear at the beginning of R&D investment in the new patent which bundle of upstream patents is required in the commercialization phase of a new downstream patent. Even if it is possible to identify the bundle of required upstream patent licenses, the difficulty of planning for various contingencies that may arise during an innovation process and the difficulty contracting parties experience in negotiating these plans make it impossible to write an ex-ante complete contract. The need for ex-post negotiation thus creates a patent holdup problem.[3,4]

A case study by Williams (2013) estimates that patent holdup has reduced downstream research and product development by 20% to 30% when an upstream company called Celera owned essential intellectual property (IP) rights on the sequencing of genes between 2001 and 2003. Celera not only charged hefty fees for the use of its IP protected gene-level data but also demanded that firms negotiate licensing agreements with the company for any resulting commercial applications. Celera's licensing agreements were mostly negotiated ex post rather than ex ante. Ex-post licensing negotiation exposes a downstream innovating firm to rent extraction because much of its own research costs are already sunk at the time of negotiation (Bessen, 2004). Moreover, asymmetric information can cause a breakdown of ex-post negotiations—reducing ex-ante investment incentives of downstream firms even further.

Applying the insight from the property rights literature to the patent process, we conjecture

[3] Our empirical analyses focus on the holdup problems encountered by a downstream innovating firm. We assume, for simplicity, that the ex-ante investment in upstream patents is exogenous and not altered by expected royalties from any downstream innovating firm. This differs slightly from the symmetric case in which both contracting parties make ex-ante 'relationship-specific' investments.

[4] Klein, Crawford, and Alchian (1978) provide a classic example of asset-specificity represented by a piece of critical infrastructure such as a pipeline. Construction of a refinery connected to the pipeline represents an asset-specific investment. The value of the refinery depends on access to the pipeline. Under separate ownership, the refinery owner may be held up by the pipeline owner in the sense that the latter can raise the price of crude oil (as the refinery input) to a very high level. The holdup problem can be overcome by joint ownership if ex-ante contracting (on pipeline access) is incomplete or difficult.

that shareholder overlap (which amounts to partial integration) between an innovating (downstream) firm and other (upstream) firms controlling complementary patents can attenuate the holdup problem and contribute to the patent success of the innovating firm.[5] Two separate channels can promote the internalization of such patent holdup. First, a *transfer internalization channel* implies that investors with joint ownership in the downstream innovating firm and upstream firms holding complementary patents could influence management of the downstream firm to internalize future patent rent transfers to the upstream firms (for the portion of the transfer payments received by the overlapping shareholders) and reduce underinvestment in downstream patents. Hansen and Lott (1996) and He and Huang (2017) provide empirical evidence consistent with this argument. The former document that investors help their portfolio firms internalize externalities imposed on one another, and the latter show that investors try to influence the product market strategies of their (same-industry) portfolio firms in a way that would maximize their overall portfolio values.

Second, a *transfer reduction channel* suggests that if such patent rent transfer can only be obtained at an efficiency loss (for example, due to potential patent litigations that retard the commercial adoption of the patent), overlapping investors can contribute to swift conflict resolution, reducing the overall patent transfer payments and increasing ex-ante investment incentives by the downstream firm. For example, Albert J. Wilson, Vice President and Secretary of TIAA-CREF, noted in a public speech that his pension fund had applied pressure on both sides in the litigation cases of Pennzoil vs. Texaco and Apple vs. Microsoft to forward the resolution of conflicts sooner than would have happened otherwise because of the fund's ownership in both litigants (Hansen and Lott, 1996).

To subject this property-rights perspective of patent success to a systematic empirical examination, we combine a large sample of U.S. patent data from the United States Patent and Trademark Office (USPTO) with institutional ownership data from Thomson Reuters for the period 1991–2007. In particular, we track stock ownership not only for the innovating firms, but also for firms owning complementary patents. The complementarities are identified directly from patent filings that explicitly list important upstream patents owned by other firms. By law, each newly filed patent must list prior art references (i.e., precursory or upstream patents) that are technologically

[5] In our empirical analysis, we identify upstream firms as those cited by a downstream innovating firm in its patent filings.

related and material to the patentability of the new application. Although inventors have a duty of candor to disclose all material prior art, patent examiners in USPTO are officially responsible for constructing the list of references. According to Alcácer, Gittelman, and Sampat (2009), examiners insert at least one citation in 92% of patent applications, and examiner citations account for about 63% of all citations made by an average patent. Our analysis identifies potential patent holdup based on this list of prior art references and assumes that the list is exogenously determined by the technology to be patented. Indeed, the frequent addition of precursory patents by patent examiners suggests that the patent filing firms have limited scope in manipulating the reference list.

Prior research (Ziedonis, 2004; Galasso and Schankerman, 2010; Noel and Schankerman, 2013) suggests that owners of upstream cited patents are reasonable proxies for the potential licensors of downstream citing patents. So-called patent-consultants occasionally disclosed that they screened the list of companies that cited their clients' patents to identify potential licensees (Ziedonis, 2004).[6] In fact, two U.S. inventors, Stephen K. Boyer and Alex Miller, were granted a patent (US6879990) in 2005 for proposing a systematic approach to identifying potential licensees from patent citation references.[7] Following this line of the literature and industry practice, our analysis uses patent citation links to upstream firms to proxy for asset complementarity and potential holdup problems faced by a downstream firm. Figure 1 provides supportive evidence for such a proxy: Firms with citation links are on average 15 times as likely to engage in patent-related lawsuits against each other as those without any citation links. The relative patent litigation risk related to citation links is even higher in R&D-intensive industries such as pharmaceuticals and computer hardware.[8] Notwithstanding the imperfect nature of the proxy, it allows us to identify asset complementarity for a large sample of firms, particularly among firms at the forefront of the innovation process.

[6]Ziedonis (2004) discussed three cases in her paper (Mogee Associates, InteCap, and Delphion). Ambercite, another intellectual property consulting company, advocated a similar approach in a recent internet posting (www.ambercite.com, 2014).

[7]They suggest creating a pool of associated patents from citation references of the target patents. Certain weighting scheme and ranking criteria are then applied to rank the owners of these associated patents to identify companies that are most likely to need a patent license from the target firms.

[8]The figure is based on the Audit Analytics Litigation database collected primarily from corporate disclosures to the Securities and Exchange Commission (SEC). Reported are 604 patent lawsuits over the period 2000–2007. Although these lawsuits may represent only a subset of all patent lawsuits, we are not aware of any reporting bias toward firm pairs with or without citation links. Previous literature, such as Schmidt (2012), has also employed this database to carry out litigation-related analysis.

Our main hypothesis is the *holdup attenuation hypothesis*, which argues that joint equity ownership between the downstream innovator and the upstream firms controlling complementary patents attenuates the holdup problem, increases investment in R&D, and contributes to the long-run patent success of the innovating firm. We further explore two refinements of this basic hypothesis: We examine whether shareholders' investment horizon and ownership concentration matter for the holdup attenuation effect.

To test these hypotheses, we first construct a new explanatory variable, *firm-level shareholder overlap* (SOL), which aggregates minimum ownership share that investors own jointly in both the innovating firm and the firms controlling the complementary assets. Consider a patent p owned by a downstream firm $O(p)$ that cites a precursory patent p_u owned by an upstream firm $O(p_u)$. If two investors A and B, respectively, own 3% and 5% in the downstream firm $O(p)$, and 2% and 6% in the upstream firm $O(p_u)$, their combined shareholder overlap for the patent pair (p, p_u) amounts to 7% $[= min(3\%, 2\%)+min(5\%, 6\%)]$. The *patent-level shareholder overlap* (sol) follows by averaging over all upstream patents cited in the patent filing of patent p, and the *firm-level shareholder overlap* (SOL) is obtained by jointly averaging over all patents of the downstream innovating firm and their respective upstream patents.

Following the literature, we only examine patents that are eventually granted by USPTO. We measure patent success by the cumulative citation count $cites_{p,t}$ of each patent p that is filed in year t and subsequently granted. Overall firm-level patent success is denoted as $CITES_{s,t}$, which aggregates all future patent citations of the entire cohort of patents filed by firm s in year t. Our measure of patent success can also be interpreted as R&D productivity because we control for firm-level R&D stock in all regression specifications. Our choice of proxy for patent success is widely used in the existing literature (e.g., Aghion, Van Reenen, and Zingales, 2013) and is in line with the studies that show a positive correlation of future citation count with the economic value of a patent (e.g., Harhoff, Narin, Scherer, and Vopel, 1999; Kogan et al., 2017) and with firm value (e.g., Hall et al., 2005).

Main Findings

Consistent with the *holdup attenuation hypothesis* of shareholder overlap, we find strong evidence that joint (overlapping) equity ownership in complementary patents fosters patent success through the attenuation of holdup. Overall, an increase by one standard deviation in firm-level shareholder overlap $SOL_{s,t-1}$ with firms owning complementary patents enhances patent success as measured by a firm's (log) patent citations ($ln[1+CITES_{s,t}]$) by 11.3% of its standard deviation. It also increases the extensive margin of patent production (i.e., number of patents successfully filed) by 18%. The results are qualitatively robust to the inclusion of various firm controls and industry or firm fixed effects, as well as to the alternative measurement of SOL with ownership data lagged by two to four years.

In addition, we show a stronger effect of shareholder overlap on patent success when such overlap originates from *dedicated investors*, characterized by concentrated portfolio positions and a long-term investment horizon, and much less so when the overlap is from other investor types.[9] This finding suggests that long-term, dedicated overlapping shareholders have stronger incentives to resolve patent holdup conflicts. While recent research has highlighted the governance influence of long-term, concentrated investors (Van Nieuwerburgh and Veldkamp, 2010; Asker, Farre-Mensa, and Ljungqvist, 2015; McCahery, Sautner, and Starks, 2016), our evidence differs in its focus on *inter-firm conflict* (rather than intra-firm conflict) in which dedicated overlapping shareholders play a special role.

How can long-term, dedicated investors influence corporate decisions? In a survey of institutional investors, McCahery, Sautner, and Starks (2016) document that long-term, dedicated investors intervene more frequently than short-term investors. They do so mainly through private, behind-the-scene discussions with management and private meetings with corporate board members. In addition, they discipline management with threats of exit, which they view as a complement to direct intervention.

We also find that the concentration of overlapping shareholder ownership matters for patent success. We argue that coordinated action might be easier to organize, and shareholders have stronger incentives to resolve a potential holdup, if the downstream innovating firm and upstream firms are jointly owned by only a few relatively large shareholders. Large overlapping shareholders of innovating firms are more likely to simultaneously serve on the boards of both upstream and

[9] We do not classify institutions based on the conventional approach because there is substantial heterogeneity even within the same class of institutions. For example, about 68% of hedge funds are among the top one-third of institutions with the highest portfolio turnover, but a significant proportion (about 16%) of them appears to pursue a long-term investment strategy, with low turnover. In our empirical analysis, we sort all institutions separately by their portfolio turnover and portfolio diversification every year. Dedicated investors are those among the top tercile of institutions with the highest portfolio concentration and lowest turnover. Overall, 21% of hedge funds, 30% of pension funds, 48% of bank trust and insurance companies, and 32% of investment companies are classified as dedicated investors.

downstream firms. In particular, 11% of the downstream firms in our sample have on average one or more board members who also sit on the boards of some of their upstream firms.[10] Our result complements the finding by Chemmanur, Shen, and Xie (2017) that overlapping equity blockholders facilitate the formation of R&D-related strategic alliances between firms in the same industry and that such alliance contributes to patent success.

Furthermore, the holdup attenuation hypothesis implies that shareholder overlap should attenuate the negative effect of holdup on firm investment. Empirically, we find an economically strong positive relation between shareholder overlap and R&D expenditure.

To the best of our knowledge, the role of joint stock ownership structure in mitigating holdup problems in patent processes has not been subject to any systematic analysis. Ex-ante complete contracting about access to auxiliary patents is difficult before the feasibility and commercial potential of a new patent are established, and ex-post contract negotiation typically occurs only after large proportions of the patent investments have been sunk. Holdup expectations reduce ex-ante investment incentives (resulting in ex-ante inefficiency) unless overlapping shareholders internalize such rent extraction through simultaneous ownership in upstream and downstream firms. Costly patent rent extraction (resulting in ex-post efficiency losses) might also be reduced through the power of overlapping shareholders vis-à-vis upstream firms.

Our paper continues as follows. In section 2.1, we present three patent-level and four firm-level strategies to address the endogeneity issues. Section 2.2 surveys the related literature. In section 3, we discuss the data, variable construction, and summary statistics. Section 4 presents the empirical evidence, and section 5 concludes. Appendix A proposes a simple model of patent holdup (from a property rights perspective) to illustrate the mechanism through which shareholder overlap increases ex-ante patent investment. Appendix B provides detailed variable definitions.

[10] We obtain board data from the BoardEx database. The database has limited coverage prior to 2000, and it covers about 66% of CRSP stocks in 2000 and 74% in 2007 (Engelberg, Gao, and Parsons, 2013). We are able to find board information for 1,755 downstream firms and 1,532 upstream firms in our sample during the period 2000–2006. For the 11% of the downstream firms that share one or more common board members with their upstream firms, their average shareholder overlap SOL is 12.25%, much higher than the average SOL (5.36%) for the rest of the firms.

2 Endogeneity Issues and Literature Review

2.1 Empirical Strategies

We pursue three patent-level and four firm-level strategies to address the endogeneity issues in the empirical relation between shareholder overlap and patent success. The three patent-level strategies are as follows.

First, we reproduce our firm-level regressions at the patent level while controlling for interacted firm and year fixed effects. These fixed effects control for all unobservable omitted variables at the level of the downstream firm. Effectively, we compare the success of any two patents filed by the same firm in the same year as a function of their patent-level shareholder overlap *sol* with the respective upstream firms. We find that this *within-firm* patent success is again positively correlated with patent-level variations in shareholder overlap at a high level of statistical significance.

Second, given the patent-level result with firm-year fixed effects, any potential omitted variable effect still remaining can only arise from the ownership structure of the patent-specific upstream firms. To address this endogeneity concern, we instrument the patent-level shareholder overlap *sol* with the average market capitalization of patent-specific upstream firms. The average size of the patent-specific upstream firms correlates positively with *sol* and so influences the patent holdup intensity, but it should otherwise be irrelevant for the success of the downstream patent, satisfying both the relevance and exogeneity conditions required of an instrument. Using a two-stage least squares approach, we again confirm that the *within-firm* variation of patent success covaries strongly with the patent-specific shareholder overlap.

Third, we use a quasi-natural experiment of financial institution mergers to identify exogenous variation in patent-level shareholder overlap *sol*. If substantial shares of a downstream firm and its upstream firm owning a complementary patent are held by two separate financial institutions, merger of the two institutions can create an exogenous increase in patent-level shareholder overlap. We find that such merger events indeed significantly increase patent-level shareholder overlap (*sol*) of the treatment patents, and that these treatment patents receive substantially more future citations than a group of otherwise similar control patents.

The four firm-level strategies are as follows. First, to further probe omitted variables operating at the firm level, we design two placebo tests. We replace the actual firm-level shareholder overlap

(SOL) with a placebo shareholder overlap. The latter replaces each cited upstream firm with a "similar" firm not cited by the downstream firm for the given year. "Similarity" is defined either as belonging to the same industry and sharing the same firm characteristics ($SOL\_Placebo1$) or by closeness in terms of technological proximity ($SOL\_Placebo2$). In both cases, the placebo shareholder overlap has no statistically significant effect on holdup mitigation and patent success.

Second, we address the issue of reverse causality by examining the evolution of shareholder overlap around patent filing events. The corresponding evolution of the two placebo measures of shareholder overlap provides a natural benchmark for the null hypothesis of no reverse causality. If investors anticipate a positive effect of shareholder overlap on future patent success and strategically acquire overlapping ownership shares prior to the public disclosure of potentially more valuable patent filings to benefit from such patent rents, then future patent success (at time $t+1$) can cause shareholder overlap (at time t), resulting in a reverse causality problem in our regression setup. Our event study evidence for the evolution of shareholder overlap around the patent filing year shows that the true shareholder overlap evolves similarly to the two placebo measures of shareholder overlap, with no discernible effect of future patent filings on true SOL. This finding is not surprising because patent developments are generally kept secret and trading on insider information is sanctioned by law.

Third, some investors may specialize in acquiring stakes in innovative firms that have a disproportionate share of patents. These technology-savvy shareholders may bring particular knowledge to the innovation process, allowing for better governance of the innovating firm. The existence of such shareholders might explain our finding of the positive SOL effect. To address this concern, we create a measure of *shareholder innovation focus* ($SIF_{s,t-1}$), which calculates the investment bias of each institutional investor toward patent filing firms and then aggregates this measure over all institutional shareholders of each downstream firm. Unsurprisingly, we find that the general innovation focus of a firm's shareholders fosters the patent success of the firm. However, the SOL effect remains strong even after controlling for this effect.

Fourth, we decompose institutional ownership into a component that contributes to shareholder overlap SOL and a component that consists of non-overlapping (or standalone) institutional ownership. We argue that non-overlapping institutional investors face a shareholder conflict with the overlapping institutional investors. Generally, the latter would like management of the downstream firm to internalize patent rents to the upstream firms. Non-overlapping institutional investors, who do not have an investment interest in the upstream firms, should view the resulting R&D investment level from patent rent internalization as *overinvestment*. Therefore, overlapping and non-overlapping institutional ownership are expected to feature opposite signs in R&D investment regressions. We do indeed find that a larger share of non-overlapping institutional investors correlates with lower R&D investment, consistent with the governance influence of overlapping shareholders as one of the causes for higher R&D expenditure. We discuss these results in more detail in Sections 4.4–4.9 and additional robustness tests in Section 4.10.

2.2 Related Literature

Notwithstanding its prominence in economic theory, the property rights view of the boundaries of the firm has seen few empirical applications. A variety of empirical problems explain the scarcity of evidence. First, non-contractible holdup problems are often difficult to identify in a complicated business environment. Second, underinvestment at the project level, as implied by the theory, requires a level of data disaggregation typically not available from corporate investment data. Any firm-level analysis is clouded by the fact that a firm can shift investments to those projects for which holdup problems are less severe. Third, investments may involve intangible resources (such as managerial attention), which pose additional measurement problems for empirical analyses. In this study, we overcome various empirical difficulties. First, we identify the potential holdup problem in patent success directly through the explicit citation of precursory patents in patent filings. Our approach is in line with the existing literature and industry practices.[11] Second, we infer (latent) project underinvestment at the patent level indirectly from the diminished success of a patent. Aggregate firm-level underinvestment is inferred either indirectly from the diminished success of all patents in a firm or directly from the reported R&D expenditure. Third, we measure the success of a patent using future patent citation count, following the evidence provided by Harhoff et al. (1999) and Kogan et al. (2017) on the positive relation between future citation count and the economic value of a patent.

How can firms avoid patent conflicts? Given the cumulative and sequential nature of technological development, it is not always possible to invent around a patented technology. In practice, licensing agreements are often used (e.g., Shapiro, 2001; Ziedonis, 2004; Hall and Ziedonis, 2007)—yet these typically require ex-post negotiation and such negotiation might not be a frictionless

[11] See, e.g., Ziedonis, 2004; and a U.S. patent, US6879990, on methodologies of identifying patent licensees.

process, resulting in efficiency losses for the innovating firms. Firms might also seek outright ownership integration via mergers to resolve patent disputes. However, firm mergers involve high transaction costs and might be challenged in court for anti-competitive reasons (Creighton and Sher, 2009). Our evidence suggests that in liquid equity markets, partial ownership integration via shareholder overlap might be achieved at lower costs.

Recent empirical work on the determinants of patent success focuses on the role of institutional shareholders. Aghion, Van Reenen, and Zingales (2013) argue that institutional shareholders are conducive to patent investment and innovation success as these shareholders provide reassurance to managers who are concerned about the risk involved in innovation projects. Bena, Ferreira, Matos, and Pires (2017) and Harford, Kecskés, and Mansi (2017) argue that long-term institutional shareholders have stronger incentives to monitor managers and therefore contribute to innovation success. By contrast, our paper examines the role of institutional investors in a world of patent complementarities from a property rights perspective. Here, institutional investors can have conflicting shareholder interest with respect to patent investments depending on their ownership overlap with upstream firms holding complementary patents.

Our work is also related to a nascent literature on the coordination role of common shareholders in corporate policies. Azar, Schmalz, and Tecu (2017) show that cross-holdings of institutional shareholders soften product market competition. He and Huang (2017) show that firms sharing common equity blockholders are more likely to engage in joint ventures, strategic alliances, and acquisitions with each other, resulting in higher profitability and market share growth. Chemmanur, Shen, and Xie (2017) focus specifically on R&D-related strategic alliances among same-industry firms backed by common equity blockholders. They show that such alliances have a positive effect on corporate innovation in that the benefits (such as knowledge spillover and human capital redeployment) outweigh the costs (such as moral hazard) of alliances. We note that in our empirical analysis, we control for two countervailing R&D spillover effects (i.e., technology spillover and product market rivalry effects) and show that our result is robust (as reported in section 4.10).

Recent empirical work has also highlighted the complementarity between equity market development and the degree of patent innovation (Brown, Martinsson, and Petersen, 2013, 2017; Hsu, Tian, and Xu, 2014). Insofar as equity market development allows for better internalization of holdup problems (through enhanced and adjustable *shareholder overlap*), this paper offers a deeper microeconomic interpretation rooted in the theory of the firm for the documented findings.

3 Data

3.1 Patent Information

We collect patent and citation information from the data set provided by Kogan et al. (2017). The data set contains annual patent and citation information for patents granted over the period 1926–2010.[12] Patent applications that have not been approved are not included in the data set. Following the existing literature (e.g., Griliches, Pakes, and Hall, 1988), we use the total number of a patent p's future citations ($cites_{p,t}$) from the patent filing year t to 2010 as our proxy for patent success. Generally, a patent is not known to the public during its application stage until USPTO publishes it, typically 18 months after the filing date. For earlier patents (filed before November 29, 2000), patent applications are not published until after they are granted. According to Hall, Jaffe, and Trajtenberg (2001), it takes on average 18 months for a patent's application to be approved and about 95% of successful patent applications are granted within three years of application, so the lag between patent filing and the first citation can range from zero to three years in most cases.

We examine the firm-level patent citations by summing up the patent-level citations by patent filing year instead of grant year because the former is closer to the date of invention. We aggregate the count statistic $cites_{p,t}$ to the total number of future patent citations generated by the cohort of patents filed by firm s in year t, denoted by $CITES_{s,t}$. Self-citations are excluded. Patent and citation counts are set to zero whenever there is no patent or citation information provided in the data. We also examine the extensive margin of patent production $N_{s,t}$, defined as the number of patent filings by firm s in year t. The corresponding intensive margin is measured by the average citations per patent $\overline{cites}_{s,t}$ (which equals the ratio of $CITES_{s,t}$ to $N_{s,t}$). Because most of these patent-related measures feature highly right-skewed distributions, we generally apply a log transformation $ln(1+X)$ to obtain more normally distributed variables for regression analyses.

We follow standard procedures to adjust for patent and citation truncation biases. First, because the patent data set only includes those patents that are eventually granted, we use only patent applications up to 2007 in our empirical analysis to allow for a three-year window of future citations up to 2010. Second, we control for time fixed effects in all our regressions to account

[12] The data set is available at https://iu.app.box.com/patents. We thank Professor Noah Stoffman for making the data set available to us.

for the fact that earlier cohorts of patents have more time to be cited than later cohorts. Third, we adjust for patent citation count based on the shape of the citation-lag distribution suggested by Hall, Jaffe, and Trajtenberg (2001, 2005).[13] Fourth, as a robustness check, we count only the citations received during the calendar year of the patent grant and three subsequent years (Lerner, Sørensen, Strömberg, 2011). Note also that because expired patents would not create any holdup problems, we ignore upstream cited patents that have expired by the time the shareholder overlap measure is constructed.[14]

3.2 Ownership Data

We obtain the ownership data from the Thomson Reuters 13F database. The SEC requires all institutional organizations, companies, universities, etc., that exercise discretionary management of investment portfolios over $100 million in equity assets to report their holdings on a quarterly basis. All common stock positions greater than 10,000 shares or $200,000 must be reported. Aghion, Van Reenen, and Zingales (2013) show reporting inconsistencies in ownership data prior to 1991, so we use ownership data only from 1991 onwards.

We then combine the patent and citation data with institutional ownership data for publicly listed firms in the United States. Our final sample includes all U.S. publicly listed firms that have more than one patent application over the sample period 1992–2007. We require each firm to have at least two valid observations because we control for firm fixed effects in our main regression specifications. Our final sample includes $2,964$ firms. We exclude all firm-year observations with missing values for the explanatory variables or control variables. The control variables, including the (log) stock market capitalization $ln(MktCap_{s,t-1})$, cumulative R&D investment $ln(1 + R\&D\ Stock_{s,t-1})$, capital intensity $ln(K/L_{s,t-1})$, and sales $ln(Sales_{s,t-1})$, are drawn from the Compustat database. The sample features $19,020$ firm-years of patent production, involving a total of $581,240$ patents. On average, a firm produces 31 patents per year.

[13] For example, for a chemical patent filed in 2000, we observe only 10 years of citations. According to Table 5 of Hall et al. (2011), for a typical chemical patent about 52.9% of the estimated total citations occur during the first 10 years. Therefore, we would divide the observed total by 0.529 to yield the truncation-adjusted total citations.

[14] According to USPTO, the 20-year protection period for utility patents starts from the grant date and ends 20 years after the patent application was first filed. The only exception applies to those patents that are filed before June 8, 1995; these patents have a protection period that is the greater of either the 20-year term discussed earlier or 17 years from the grant date. (See http://www.uspto.gov/web/offices/pac/mpep/mpep-2700.pdf.)

3.3 Variable Construction

A key explanatory variable in our analysis is *shareholder overlap*, which we define as follows: Let $O(p)$ designate the downstream innovating firm owning patent p and $O(p_u)$ represent the upstream firm owning patent p_u. The *pairwise (institutional) shareholder overlap* between the downstream patent p and an upstream patent p_u is defined as

$$PSOL(p, p_u) = \sum_i \min[w_{i,O(p)}, w_{i,O(p_u)}], \tag{1}$$

where $w_{i,O(p)}$ and $w_{i,O(p_u)}$ are the ownership share (relative to the total institutional ownership of the respective firm) of institutional investor i in firms $O(p)$ and $O(p_u)$, respectively. We lag the ownership measure by one year relative to the application year of patent p. The *patent-level shareholder overlap* (sol) follows as the (importance) weighted average of $PSOL(p, p_u)$ over the N_p upstream patents of patent p, given by

$$sol_p = \sum_{u=1}^{N_p} w(p_u) PSOL(p, p_u). \tag{2}$$

The *firm-level shareholder overlap* (SOL) is obtained as the (importance) weighted average sol_p over all N_s patents filed by firm s in a given year, given by

$$SOL_s = \sum_{p=1}^{N_s} w(p) sol_p = \sum_{p=1}^{N_s} \sum_{u=1}^{N_p} w(p) w(p_u) PSOL(p, p_u). \tag{3}$$

A measurement issue concerns the choice of the weights, $w(p)$ and $w(p_u)$, which reflect the importance of patents p and p_u, relative to other patents. In the context of our model (presented in Appendix A), a higher weight is assigned to a more important upstream patent, reflecting the fact that its owner is likely to have stronger bargaining power in terms of future rent extraction. A higher weight is also assigned to a more important downstream patent, reflecting the fact that any percentage holdup loss from such a patent amounts to more value loss for the firm.

In our main empirical tests, we measure relative importance by the relative (log) citation count as follows:

$$w(p) = \frac{ln[1 + cites_s(p)]}{\sum_{p=1}^{N_s} ln[1 + cites_s(p)]} \quad \text{and} \quad w(p_u) = \frac{ln[1 + cites(p_u)]}{\sum_{u=1}^{N_p} ln[1 + cites(p_u)]}. \tag{4}$$

In the robustness section, 4.10, we report additional results using two alternative weighting schemes: The first uses a non-parametric rank measure of future citations to calculate the relative importance weight, and the second simply uses equal weights. The results are qualitatively similar.

A limitation of our analysis is that due to data constraint we can measure ownership for only publicly listed firms, not private firms. Data on the portfolio holdings of private investors are generally not publicly available either. As a result, we may underestimate the extent of shareholder overlap, especially when the proportion of privately owned upstream patents is large. This imprecise measure of shareholder overlap creates an attenuation bias in the OLS estimate of SOL. To mitigate this effect, we track the average share of privately owned upstream patents for each downstream firm s and include it as a control variable, denoted by $Private\ Patent\ Share_s$. Because this variable captures potential "underestimation" of the true SOL, we expect it to have a positive sign.

3.4 Summary Statistics

Institutional ownership in U.S. listed stocks has grown rapidly, from an average of 25% in 1991 to 49% in 2006. The corresponding share is considerably larger for patent filing firms and rises from 41% in 1991 to 71% in 2006. Patent filing firms tend to be larger, and institutional investors typically prefer large firms. Graphs A and B in Figure 2 depict the distributions of institutional ownership and firm-level shareholder overlap, respectively, for the period 1991–2006. Parallel to the rise in institutional ownership, the average firm-level shareholder overlap increases from 5.6% in 1991 to 7.4% in 2006. In our analysis, time fixed effects are included in all regressions to ensure that the documented shareholder overlap effect does not capture any parallel time trend in patent success. Cross-sectionally, shareholder overlap is positively related to institutional ownership in the downstream firm and even more strongly with its market capitalization, as shown in Figure 2, Graphs C and D. Shareholder overlap also varies substantially across firms with similar levels of institutional ownership and market capitalization. Such large heterogeneity in a firm's indirect control over complementary upstream patents via overlapping shareholders could plausibly condition patent holdup and determine a firm's long-run patent success.

Table 1 reports the summary statistics of key variables used in our analysis. Patent-level shareholder overlap (sol) shows an average value of 14.4% with a standard deviation of 14.2%, much larger than the corresponding statistics of 6.2% and 6.3% for firm-level shareholder overlap (SOL). The higher mean and standard deviation for the former are explained by the fact that firms with many patent filings are usually larger and feature a higher level of shareholder overlap. Detailed definitions of all variables are provided in Appendix B.

4 Evidence of Patent Success

Patent is about the extension of ownership rights to new ideas, products, and processes. The element of novelty implies that the scope for ex-ante contracting prior to patent investment is limited. The property rights view of a firm is therefore a natural starting point for thinking about patent investment and development. In Appendix A we develop a simple model of holdup attenuation through shareholder overlap, from a property rights perspective. In this section, we examine several testable hypotheses implied by the model.

4.1 Baseline Specification

Our main hypothesis (*the holdup attenuation hypothesis*) argues that joint equity ownership between the downstream innovator and the upstream firms controlling complementary patents attenuates the holdup problem and contributes to the long-run patent success of the innovating firm. We measure patent success in log terms as $ln[1+CITES]$.[15] The baseline regression linking patent success to shareholder overlap is

$$ln[1 + CITES_{s,t}] = \beta_0 + \beta_1 SOL_{s,t-1} + \beta_2 Controls_{s,t-1} + \epsilon_s + \mu_t + \eta_{s,t}, \quad (5)$$

where the coefficient of interest is $\beta_1 \geq 0$. (In particular, the model developed in Appendix A implies $\beta_1 = (\frac{1}{b} + \frac{\gamma}{b} + \gamma)\delta \geq 0$.) More shareholder overlap with firms holding upstream patents should boost the downstream innovating firm's patent success because holdup problems are attenuated. In the above specification, β_0 represents the overall constant for all observations, β_1 is the coefficient for SOL, β_2 denotes the vector of coefficients for control variables, ϵ_s and μ_t denote, respectively, firm and year fixed effects, and $\eta_{s,t}$ is the error term.

We estimate Eq. (5) over the period 1992–2007. The citation count $CITES_{s,t}$ for patents filed

[15] As discussed in the robustness section (4.10), using $ln[CITES]$ as the dependent variable yields qualitatively similar results.

by firm s in year t includes all future citations up to year 2010. *Shareholder overlap* ($SOL_{s,t-1}$) measures the ownership overlap at the end of year $t-1$ between the innovating firm and all other firms controlling complementary patents. For the choice of control variables, we follow Aghion, Van Reenen, and Zingales (2013) and include the cumulative R&D investment $ln(1+R\&D\ Stock_{s,t-1})$, a measure of relative capital intensity $ln(K/L_{s,t-1})$, and firm sales $ln(Sales_{s,t-1})$. We also control for firm market capitalization value $ln(MktCap_{s,t-1})$ and the (weighted) share of private firms in the cited upstream firms, $Private\ Patent\ Share_{s,t-1}$.

In Table 2, Columns 1–2 present the results for all firms and Columns 3–4 for firms in the top three R&D-intensive sectors (pharmaceuticals, computer hardware, and telecommunications equipment).[16] Robust standard errors clustered at the firm level are reported in parentheses. All regressions control for a full set of year dummies and industry dummies based on four-digit SIC codes. Columns 2 and 4 additionally control for firm fixed effects, using the Blundell, Griffith, and Van Reenen (1999) pre-sample mean scaling estimator.

The ordinary fixed effect estimator with firm dummies is consistent only if the independent variables are strictly exogenous with respect to the error term $\eta_{s,t}$. Theoretically, such strict exogeneity can be relaxed to predetermined regressors if a first-difference estimator is used together with lagged regressors as instruments. However, slowing moving regressors (as in our case) prevent instrumentation by their lagged values. Blundell et al. (1999) therefore propose a "pre-sample mean scaling" method. Specifically, they suggest replacing firm dummies with the pre-sample mean of the dependent variable (measured at the firm level). They show that this estimator is consistent when the pre-sample size is large, even if the independent variables are only weakly exogenous (or predetermined). To obtain consistent regression estimates, we following this procedure and construct a 25-year pre-sample mean of $CITES_{s,t}$. The same procedure is also adopted by Blundell et al. (1999) to examine the relation between innovations and market shares, by Aghion et al. (2013) to examine the relation between innovations and institutional ownership, and by Blanco and Wehrheim (2017) to examine the relation between innovations and option trading.

The baseline regression in Column 1 shows that *shareholder overlap SOL* represents a statistically and economically significant explanatory variable. The point estimate of 3.692 in Column

[16] We identify the three R&D-intensive sectors following the approach suggested by Bloom, Schankerman, and Van Reenan (2013). Specifically, based on the Fama-French 49-industry classification, the pharmaceutical sector corresponds to industry 13 (drug: pharmaceutical product), the computer hardware sector corresponds to industry 35 (hardware: computer), and the telecommunications equipment sector corresponds to industry 37 (chips: electronic equipment).

1 implies that an increase in shareholder overlap by one standard deviation (or 0.063) increases patent success in terms of a firm's log patent citation ($ln[1+CITES]$) by 11.3% of its standard deviation of 2.065, suggesting that shareholder overlap has an economically large attenuation effect on patent success. The SOL estimate remains highly significant with the inclusion of firm fixed effects in Column 2. The control variables generally have the expected signs: Firm size correlates positively with the overall number of citations a firm receives, suggesting that large firms may generally be in a better position to assure the long-run success of their patents or may simply launch more successful patents. A higher stock of cumulative R&D spending and a higher capital intensity ratio also correlate positively with future patent success. As expected, *Private Patent Share* has the same sign as SOL because it proxies for the possible underestimation of shareholder overlap due to the unobserved overlap originating from private investors.

Columns 3–4 repeat these regressions for the top three R&D-intensive sectors. As expected, we find a statistically and economically stronger SOL effect in these sectors than in others. The point estimates for SOL increase by about 25% in Columns 3–4, compared with those in Columns 1–2. Not surprisingly, shareholder overlap matters most for patent success in those industries that are most patent-intensive.

4.2 Intensive versus Extensive Margins

Shareholder overlap may affect intensive and extensive margins differently. The intensive margin of patent success is captured by the average number of citations per patent, $\overline{cites}$. Again, we use the logarithmic transformation $ln[1 + \overline{cites}]$ to obtain a suitable dependent variable for the regression

$$ln[1 + \overline{cites}_{s,t}] = \theta_0 + \theta_1 SOL_{s,t-1} + \theta_2 Controls_{s,t-1} + \epsilon_s + \mu_t + \eta_{s,t}, \quad (6)$$

where $\theta_1 > 0$ implies that patent holdup reduces the average success of a firm's patents. A positive value of θ_1 points to ex-post patent value destruction under patent conflict rather than mere rent redistribution to upstream firms. (Specifically, the model presented in Appendix A implies $\theta_1 = \gamma\delta > 0$. The parameter γ measures the *efficiency loss* of patent holdup, whereas δ measures the *distributional loss* from rent transfers to upstream firms. Rejection of $\theta_1 = 0$ in favor of $\theta_1 > 0$ would imply $\gamma > 0$, suggesting that the holdup problem produces an adverse effect on the average success of the innovating firm's patents, beyond the loss of rent redistribution to upstream firms.) As shown in the model, frictionless ex-post rent redistribution should primarily affect the extensive margin of patent production, but not the intensive margin.

Table 3, Columns 1–2 summarize the effect of shareholder overlap on the intensive margin. The regression in Column 1 yields a point estimate of 0.584, which implies that an increase in shareholder overlap by one standard deviation (or 0.063) corresponds to an increase in the average citation count per patent by a modest 3.2% of its standard deviation. Inclusion of firm fixed effects in Column 2 generates a point estimate for SOL of 0.527, which is only slightly smaller than the estimate obtained in Column 1.

The empirical specification for the extensive margin uses the log number of patents as the dependent variable

$$ln[1 + N_{s,t}] = \psi_0 + \psi_1 SOL_{s,t-1} + \psi_2 Controls_{s,t-1} + \epsilon_s + \mu_t + \eta_{s,t}, \quad (7)$$

where the coefficient ψ_1 captures the effect of holdup mitigation through shareholder overlap on the number of successful patent filings. The point estimate of 2.923 for ψ_1, reported in Column 3 of Table 3, suggests a strong economic significance for the shareholder overlap measure; a one-standard-deviation increase in SOL is associated with 18% increase in the number of patents. The coefficient retains its statistical significance again in Column 4 when firm fixed effects are included.

Overall, the results suggest that shareholder overlap is associated with both more citations for each granted patent (i.e., the intensive margin of patent success) and more granted patents in total (i.e., the extensive margin of patent production). The relation between holdup mitigation and patent production appears economically much stronger for the extensive margin. Under shareholder overlap, firms tend to file more patents—presumably because of lower patent rent transfers and their internalization by overlapping shareholders.

4.3 Two Dimensions of SOL Heterogeneity

Which type of overlapping shareholders has the strongest incentives to resolve a potential patent holdup and the greatest ability to influence corporate managers in their R&D decisions? First, we conjecture that long-term institutional investors with concentrated portfolio positions are more likely to devote time and effort to resolving patent-related conflicts. Second, concentration of overlapping ownership among relatively few institutional investors limits free-riding and facilitates the coordination of investor influence.

To test the first hypothesis, we categorize institutional investors into (i) dedicated investors, (ii) intermediate investors, and (iii) transient investors based on a combination of portfolio concentration (proxied by the Herfindahl-Hirschman Index, HHI) and portfolio turnover (proxied by the churn ratio defined in Gaspar, Massa, and Matos, 2005).[17] At the end of each year, we sort all institutional investors by the HHI (in descending order) and churn ratio (in ascending order) and then calculate the combined rank as the average of the HHI rank and the churn ratio rank. We label investors in the top tercile of the combined rank (high concentration and low turnover) dedicated investors, and investors in the bottom tercile (low concentration and high turnover) transient investors. The remaining investors, in the middle tercile, are labeled intermediate investors. The labeling of investor types along the two sorting criteria is shown in Figure 3 and permits the decomposition of shareholder overlap according to the three investor types:

$$SOL_{s,t-1} = SOL\_Dedicated_{s,t-1} + SOL\_Intermediate_{s,t-1} + SOL\_Transient_{s,t-1}. \tag{8}$$

The regression result, reported in Column 2 of Table 4, confirms our hypothesis. The coefficient for $SOL\_Dedicated$ is 19.926, more than five times the estimate for SOL in the baseline regression of Table 2 (reproduced in Column 1 of Table 4). Shareholder overlap originating from the other two groups of investors shows a much weaker effect on patent success. This confirms the first hypothesis that investor type matters for the holdup attenuation effect.

The second hypothesis concerns the concentration of shareholder overlap rather than its types. If joint equity ownership is constituted by a few relatively large shareholders, coordinated investor action is easier to organize and incentives to do so are strongest. To test this hypothesis, we consider a downstream patent p filed by firm s in year t and a related upstream patent p_u owned by firm u. Let $i \in I_{(p,p_u),t-1}$ denote an overlapping investor, who at the end of time $t-1$ owns equity shares (relative to total institutional ownership) $w_{i,s}$ and $w_{i,u}$ in firms s and u, respectively. For a patent pair (p, p_u), we can define a Herfindahl-Hirschman Index ($hhi_{(p,p_u),t-1}$) based on the overlapping ownership shares $\varpi_i = \min[w_{i,s}, w_{i,u}]$ of all overlapping shareholders $i \in I_{(p,p_u),t-1}$. We can further aggregate this concentration measure $hhi_{(p,p_u),t-1}$ over all upstream patents p_u related

[17] We provide detailed definitions of these variables in Appendix B.

to patent p and, subsequently, over all downstream patents p filed by firm s in year t to obtain a weighted Herfindahl-Hirschman Index ($WHHI$) of ownership concentration of overlapping shareholders, defined as

$$WHHI_{s,t-1} = \sum_{p=1}^{N_s} \sum_{u=1}^{N_p} w(p)w(p_u)hhi_{(p,p_u),t-1} \quad , \tag{9}$$

where $w(p)$ and $w(p_u)$ denote (as defined in Eq. (4)) the relative importance weights for patents p and p_u, respectively, and ownership shares are measured at the end of year $t-1$. $WHHI$ describes the concentration of overlapping ownership stakes at the firm level and thus captures the coordination problem among the overlapping investors.

Table 4, Column 3 includes $WHHI$ as a separate control variable. The estimated coefficient is positive and statistically highly significant—suggesting that concentration of joint ownership shares by overlapping shareholders positively correlates with patent success beyond the shareholder overlap SOL itself. The coefficient estimate of 2.206 for $WHHI$ implies that an increase in the ownership concentration of shareholder overlap by one standard deviation (or 0.071) generates the same effect on patent success as raising SOL by 57.9% relative to its mean (= [0.071 × 2.206] / [4.361 × 0.062]). We thus infer that the coordination problem among dispersed overlapping institutional investors represents an important impediment to the exercise of effective shareholder power.

4.4 Patent-Level Regressions

In this section, we first present patent-level regressions that use separate firm and year fixed effects and also interacted firm-year fixed effects. The latter specification identifies the holdup attenuation effect on patent success by relying entirely on the comparison of different patents filed by the same firm in the same year. In the second step, we propose to use the average size of the patent-specific upstream firms as an instrument for patent-level shareholder overlap. The two-stage least squares (2SLS) procedure promises consistent regression estimates because the average size of the patent-specific upstream firms correlates positively with patent-level shareholder overlap and is otherwise irrelevant for the success of the downstream patent, satisfying both the relevance and exogeneity conditions required of an instrument.

Different patent filings by the same firm may cite different upstream patents, resulting in patent-specific holdup and shareholder overlap even within the same firm-year. Patent-specific holdup attenuation is captured by patent-level shareholder overlap $sol_{p,t-1}$ in the regression specification

$$ln[1 + cites_{p,t}] = \lambda_0 + \lambda_1 sol_{p,t-1} + \lambda_2 Controls_{p,t-1} + \delta_{s,t} + \pi_{f,t} + \tau_{p,t}, \tag{10}$$

where $cites_{p,t}$ denotes the future citation count of patent p filed in year t. Similar to the firm-level regressions, the shareholder overlap variable lags the dependent variable by one year. The variable $\delta_{s,t}$ denotes the interacted firm and year fixed effects, $\pi_{f,t}$ represents the interacted technology field and year fixed effects, and $\tau_{p,t}$ is the residual term.[18]

Any omitted variable problem should be less severe for the patent-level regressions because the interacted year-firm fixed effects control for all unobservable (time-variant and time-invariant) influences at the level of the downstream firm. This sharper identification comes at a price: It excludes all effects of shareholder overlap on the extensive margin and focuses exclusively on the intensive margin of patent success. Any remaining omitted variable effect in Eq. (10) needs to influence the patent-level success of the downstream firm (i.e., $ln[1 + cites_{p,t}]$) and simultaneously correlate with the ownership structure of the patent-specific upstream firms (and hence correlate with $sol_{p,t-1}$). To address this remaining endogeneity concern with respect to the ownership of the upstream firms, we conduct a $2SLS$ regression by instrumenting sol with the average market capitalization of the patent-specific upstream firms. The average size of the patent-specific upstream firms correlates positively with shareholder overlap ($sol_{p,t-1}$) but is unlikely to matter for the patent success of the downstream firm except via its correlation with $sol_{p,t-1}$. (That is, it is not correlated with the residual term $\tau_{p,t}$).

Because we control for firm-year fixed effects in the patent-level regressions, we discard all firm-years that feature only one patent application. Such cases account for about 25% of the overall sample. The patent-level data thus feature a strong selection of firms with many patents—51% of all patent filings are from the 1% most patent-intensive firms (as measured by the total number of patent filings over the sample period) and the other 49% are from the remaining 99% of the firms. Note also that the patent-level citation success $cites_{p,t}$ can capture only the intensive margin, not the extensive margin, of patent success.

Columns 1–2 in Table 5 show positive and statistically significant point estimates for patent-

[18] Hall, Jaffe, and Trajtenberg (2001) categorize technology classes into 36 technology fields.

level shareholder overlap. Columns 3–4 report the first- and second-stage results of the 2SLS regression, respectively. The exclusion restriction in the first stage for the instrument has a t-statistic of 99.6, which indicates a strong instrument. The second-stage point estimate for *sol* in Column 4 is 0.283, almost identical to the OLS estimate (0.272) in Column 2. Overall, these results confirm the hypothesis that even within a firm the success of any specific patent is conditioned by the holdup attenuation provided by overlapping shareholders.

4.5 Quasi-Natural Experiment of Financial Institution Mergers

In this section, we conduct a quasi-natural experiment of financial institution mergers to examine the impact of an exogenous increase in shareholder overlap on patent success. The literature (e.g., Holthausen, Leftwich, and Mayers, 1990; Keim and Madhavan, 1996; He and Huang, 2017) suggests that financial institutions often merge for reasons unrelated to the prospects of their portfolio holdings and that the acquirer typically keeps the target's portfolio holdings for an extended period of time without liquidating them for transaction cost concerns. Therefore, if a downstream innovating firm and its upstream firm holding complementary patents are separately held by the two merging financial institutions before the merger, their shareholder overlap should increase right after the merger. Such merger events therefore create plausibly exogenous variation in shareholder overlap between two firms.

We form our merger sample following a similar methodology to that in He and Huang (2017). Specifically, we collect all merger deals between any two 13F financial institutions (with SIC Codes 6000–6999) announced during the period 1992—2006 from the SDC database. We require that a merger is completed within one year of its announcement and that the target stops its 13F filings within one year following the merger completion date. We use a 2.5% cut-off of institutional ownership as our definition of blockholding to increase our sample size, but using a 5% cut-off as in He and Huang (2017) yields qualitatively similar results. We identify as a treatment patent a downstream patent p that meets two criteria: First, the downstream firm owning patent p is blockheld by one of the merging institutions during the quarter immediately prior to the merger announcement. Second, the other merging institution does not blockhold the downstream firm but does blockhold at least one of patent p's upstream firms during the same quarter before the merger. Note that the choice of a relatively large ownership cut-off at 2.5% should predict a large increase in shareholder overlap for the treatment patents, and such an increase is likely to be persistent after the merger. Furthermore, our selection of treatment patents only uses portfolio holdings information prior to the merger, mitigating the concern that the actual post-event portfolio holdings may be endogenous.

For each treatment patent p, we define as control patents all patents q in the same patent class as p and filed in the same year by the same downstream firm owning patent p, but none of patent q's upstream firms are blockheld by the other merging institution. In total, we identify 47 merger deals featuring $17,707$ treated patents and $70,383$ control patents.

We employ a difference-in-difference approach to compare the success of treatment patents and control patents. For each merger deal, we consider a seven-year event window centered around the year of the merger event. We first verify that institution mergers do indeed lead to an increase in shareholder overlap for the treatment patents, and in the second step we examine the effect of such an exogenous increase in shareholder overlap on patent success. Specifically, we estimate the following two regressions:

$$sol_p = \gamma_0 + \gamma_1 Treat + \gamma_2 Post\text{-}Merger + \gamma_3 Treat \times Post\text{-}Merger + \xi_{e,s,f} + \varsigma_p \quad (11)$$

$$ln[1 + cites_p] = \gamma_4 + \gamma_5 Treat + \gamma_6 Post\text{-}Merger + \gamma_7 Treat \times Post\text{-}Merger + \omega_{e,s.f} + \nu_p, \quad (12)$$

in which *Treat* is a dummy for treatment patents and *Post-Merger* a dummy for the post-merger period. $\xi_{e,s,f}$ and $\omega_{e,s,f}$ denote the interacted fixed effects for the merger event e, downstream firm s, and patent class f. ς_p and ν_p are error terms.

Table 5, Column 5 reports the result for Eq. (11). The estimate of 0.017 for the interacted *Treat* $\times$*Post-Merger* term confirms that post-merger treatment patents (but not control patents) do indeed experience an economically significant increase in shareholder overlap (*sol*), at the magnitude of about 11.8% of its mean. Column 6 reports the result for Eq. (12). The estimate of 0.135 for the interacted *Treat*$\times$*Post-Merger* term suggests that post-merger treatment patents experience a 13.5% higher citation count than control patents—a difference that amounts to about 10% of the standard deviation of log patent citations ($ln[1+\text{cites}_p]$). Overall, the institution merger evidence points to an economically significant causal effect of shareholder overlap *sol* on patent success.

4.6 Two Placebo Tests

We propose two different placebo tests to examine whether any unobservable factors could facilitate the formation of joint equity ownership between a downstream innovating firm and its upstream firms and, simultaneously, enhance the patent success of the downstream firm. In each placebo test, we replace the *true* shareholder overlap (SOL) with a *placebo* shareholder overlap ($SOL\_Placebo1$ or $SOL\_Placebo2$). For $SOL\_Placebo1$, we replace each cited upstream firm with a *similar* firm that is *not cited* by the downstream firm in the given patent application year. A placebo firm is chosen based on the criteria that it must have the same four-digit SIC codes as the true upstream firm and have the shortest Euclidean distance to the true upstream firm in terms of (log) firm asset size and (log) number of patents filed in the past five years. $SOL\_Placebo2$ is constructed similarly but the placebo firms are matched to the true upstream firms based on their technological proximity (i.e., the closeness in the distribution of their patents across various technology fields) as defined by Bloom, Schankerman, and Van Reenen (2013). The two placebo measures, $SOL\_Placebo1$ and $SOL\_Placebo2$, have slightly lower means (0.050 and 0.048, respectively) than the true measure of shareholder overlap SOL (mean = 0.062).

Columns 4–5 of Table 4 show that the two placebo measures of shareholder overlap do not feature any statistically significant correlation with patent success. If the positive holdup attenuation effects of shareholder overlap documented in the previous sections are driven by unobservable factors *unrelated to patent citation links*, such omitted variables should similarly lead to a positive relation between placebo shareholder overlap and patent success. Yet, we do not find such evidence for the two placebo measures of shareholder overlap, a result inconsistent with the omitted variable hypothesis.

4.7 R&D Investment and Non-Overlapping Institutional Ownership

In this section, we examine whether shareholder overlap reduces R&D underinvestment, using the following linear panel regression

$$R\&D\ Exp_{s,t}/Assets_{s,t} = \kappa_0 + \kappa_1 SOL_{s,t-1} + \kappa_2 Controls_{s,t-1} + \epsilon_s + \mu_t + \eta_{s,t}, \quad (13)$$

Table 6, Column 1 reports the regression results. The effect of shareholder overlap (SOL) is statistically and economically significant in the specification. The point estimate of 0.24 for SOL

suggests that an increase in shareholder overlap by one standard deviation (or 0.063) increases the R&D expenditure to asset ratio by about 10% of its standard deviation. The holdup attenuation effect of shareholder overlap on R&D investment is therefore economically important.

Previous research has argued that institutional ownership can *ceteris paribus* provide better long-term managerial incentives conducive to the pursuit of R&D (see, e.g., Aghion, Van Reenen, and Zingales, 2013). We therefore control for institutional ownership in Column 2 but find that the shareholder overlap variable (SOL) retains its economic and statistical significance, whereas the institutional ownership variable (IO) is statistically insignificant. To probe this issue further, we decompose institutional ownership itself into ownership by overlapping institutional shareholders (IO^{SOL}) that contributes to shareholder overlap (i.e., the aggregate ownership of all shareholders i with $\min[w_{i,O(p)}, w_{i,O(p_u)}] > 0$ for at least one downstream-upstream patent pair (p, p^u)), and into residual non-overlapping institutional ownership (IO^{NOL}). Formally, for each downstream firm s in year t we have

$$IO_{s,t} = IO_{s,t}^{SOL} + IO_{s,t}^{NOL}. \tag{14}$$

By construction, IO^{SOL} strongly correlates with the shareholder overlap measure SOL, with a correlation of 0.53 during our sample period. If institutional ownership *per se* exerts a positive influence on R&D investment, we expect the same positive coefficient for both $IO_{s,t-1}^{SOL}$ and $IO_{s.t-1}^{NOL}$ in our regressions. By contrast, the property rights view predicts an agency conflict between overlapping institutional shareholders, who advocate for more investment due to patent rent internalization, and non-overlapping shareholders, who view the resulting R&D investment level from patent rent internalization as overinvestment. Hence, a larger share of overlapping (non-overlapping) institutional shareholders should correlate with higher (lower) R&D investment.

Column 3 of Table 6 modifies the specification in Eq. (13) to include both overlapping institutional ownership $IO_{s,t-1}^{SOL}$ and non-overlapping institutional ownership $IO_{s.t-1}^{NOL}$ and reveals opposite signs for the respective coefficients at high levels of statistical significance. The result suggests that different types of institutional owners with different property rights exercise opposing influences on the R&D investment process. Overall, our evidence shows that the property rights perspective is empirically pertinent.

4.8 Shareholder Innovation Focus

Some investors may specialize in acquiring stakes in innovative firms that have a disproportionate share of patents. These technology savvy shareholders may bring particular knowledge to the innovation process, allowing for better governance of the innovating firm. Such a *shareholder innovation focus* is directly measurable based on ownership data in a simple three-step procedure. In the first step, we define for each listed firm s' the *firm innovation focus* (FIF) as the ratio of the future citation count of all patents filed by firm s' in year t to the industry average citation count during the same period. In the second step, we account for all institutional investors i in firm s and calculate their respective *investor innovation focus* (IIF) as the value-weighted average *firm innovation focus* for all stocks s' in their respective investment portfolios except for stock s itself. In the third step, the *shareholder innovation focus* (SIF) for firm s is defined as the value-weighted average of investor innovation focus for all shareholders i in firm s,

$$SIF_{s,t} = \sum_{i} w_{i,s,t} IIF_{i,s,t} \quad , \tag{15}$$

where $w_{i,s,t}$ represents the equity shares held by institutional investor i relative to the aggregate holdings of all institutional investors in firm s at the end of year t. A firm mostly owned by investors with a high IIF should feature a high SIF value. Shareholders' governance competence (proxied by $SIF_{s,t}$) with respect to the innovating firm s should have a positive effect on the patent success of the firm.

Table 7, Column 3 includes shareholder innovation focus $SIF_{s,t}$ as an additional explanatory variable for patent success. As expected, we find that the general innovation focus of a firm's shareholders fosters patent success of the respective firm, but the SOL effect remains strong even after accounting for this factor.

4.9 Reverse Causality?

If investors anticipate a positive effect of shareholder overlap on future patent success and strategically acquire overlapping ownership shares prior to the public disclosure of potentially more valuable patent filings to benefit from such patent rents, then future patent success (at time $t+1$) can cause shareholder overlap (at time t), resulting in a reverse causality problem in our regression setup.

To examine this reverse causality issue, for each yearly cohort of patents filed between 1991 and 2007, we measure the evolution of the average firm-level shareholder overlap relative to the year of the patent filing. For a cohort of downstream patents filed in year t, let $SOL(t,k)$ represent the average shareholder overlap measured based on the ownership data at the end of year $t+k$, where $k=-5,-4,...,0,1$. For example, $SOL(t,-3)$ denotes average shareholder overlap between downstream and upstream firms measured based on ownership at the end of year $t-3$ for all patents filed in year t. The average ownership overlap (measured at lag k) across all patent filing years is denoted by $\overline{SOL}(k)$ and is plotted in Figure 4.[19] As a benchmark, we also plot the evolution of the two placebo shareholder overlap measures, $\overline{SOL}\_Placebo1(k)$ and $\overline{SOL}\_Placebo2(k)$, defined analogous to $\overline{SOL}(k)$. The vertical line segment marks two standard deviations around the mean value for each measure. The figure shows that in the neighborhood of the patent filing year ($k=0$), the average shareholder overlap $\overline{SOL}(k)$ evolves similarly to the two placebo benchmarks, which are by construction devoid of future patent rents and thus not subject to any reverse causality concern.

Overall, we find no evidence that the shareholder overlap $\overline{SOL}(k)$ reacts endogenously in anticipation of patent rents from future patent filing. During the five-year run-up to the patent filing year, $\overline{SOL}(k)$ does not actually change much, with an aggregate change of only -0.00024, which is less than 0.5% of a standard deviation of SOL. This finding may not be surprising for at least two reasons: First, patent developments are generally kept secret so that public information is extremely scarce. Second, legal restrictions on insider trading limit the scope for stock trading on private information.

4.10 Robustness Issues

In this section, we examine two additional alternative hypotheses that may account for the evidence reported in the preceding sections as well as several robustness checks concerning the measurement of shareholder overlap and patent citations.

First, Aghion, Van Reenen, and Zingales (2013) argue that R&D investments have a long-term horizon, and a high share of institutional investors allows management to focus on the long-term return on investment. We therefore include *institutional ownership* ($IO_{s,t-1}$) as an

[19] We note that the full set of $SOL(t,k)$ cannot be calculated for all years. For example, for patents filed in 1992, we can only calculate $SOL(t,k)$ for $k >= -1$. Similarly, for patents filed in 2007, $SOL(t,k)$ can only be calculated for $k <= 0$.

additional explanatory variable of patent success in Column 2 of Table 7. We find that *shareholder overlap* ($SOL_{s,t-1}$) retains its high positive level of statistical significance even with the inclusion of *institutional ownership* in the regression.

Second, Bloom, Schakerman, and Vanreenen (2013) show two countervailing *R&D spillover* effects on a firm's innovation success: A positive effect due to technology spillover (from other firms that operate in similar technology fields) and a negative effect due to product market rivalry (from other firms that operate in similar product markets). Column 4 shows that even after accounting for these two factors, measured by $ln(SpillTech)$ and $ln(SpillSIC)$, the shareholder overlap effect remains quantitatively unchanged. Columns 5 and 6 include all these explanatory variables simultaneously. The former is estimated by OLS with $ln[1+CITES_{s,t}]$ as the dependent variable (as before) and the latter is estimated using a negative binominal model with $CITES_{s,t}$ as the dependent variable. The SOL effect remains strong in both models.

The robustness tests concerning the measurement of shareholder overlap and patent citations are as follows: First, our baseline measure of shareholder overlap ($SOL_{s,t-1}$) is based on ownership stake at the end of year $t-1$. Column 7 of Table 7 replaces $SOL_{s,t-1}$ with $SOL_{s,t-2}$, which is measured based on ownership stake at the end of year $t-2$. The estimate of $SOL_{s,t-2}$ remains highly significant. Using equity stakes measured at years $t-3$ and $t-4$ still produces statistically and economically highly significant point estimates for SOL, albeit at a smaller magnitude.

Second, as an alternative measure, we replace the log citations count $ln[1+cites_s(p)]$ in Eq. (4) with a rank measure of future citations $rank(p)$ to obtain a new shareholder overlap measure SOL_rank. We also create an equal-weight measure that simply aggregates all combinations of downstream and upstream patents under equal weights. The resulted shareholder overlap variable is SOL_equal. Notwithstanding these variable modifications, we still find qualitatively similar results, reported in Columns 8–9, for the holdup attenuation effect of shareholder overlap.

Third, our baseline measure of $CITES$ follows Hall, Jaffe, and Trajtenberg (2001) in adjusting citation count based on the shape of the citation-lag distribution. We reproduce our results using an alternative aggregation proposed by Lerner, Sørensen, and Strömberg (2011), in which we count only the citations received during the calendar year of the patent grant and the three subsequent years. This alternative citation count is denoted by $CITES^{3y}$. The modified shareholder overlap variable is denoted by SOL^{3y}. Our result, reported in Column 10, is again robust to this alternative measure of citation count.

Fourth, we repeat the benchmark regression of Column 1 but use $ln(CITES)$ as an alternative dependent variable. The economic significance of SOL, reported in Column 11, remains high in this smaller sample.

Fifth, as patent citation count is often perceived as a value signal (Harhoff et al., 1999; Hall, Jaffe, and Trajtenberg, 2005), overlapping institutional shareholders may promote cross-citations among firms in which they also have a joint equity stake. To eliminate such spurious effects from our regression, we exclude all citations that come from the upstream firms cited in the patent filings of the downstream firm. Column 12 repeats the baseline regression but uses this modified patent citation $ln(1 + CITES^F)$ as the dependent variable. The estimate of 3.252 for SOL is quantitatively similar to that of 3.45 reported in the baseline regression, suggesting that any potential bias arising from such citation manipulation is small.

5 Conclusion

This paper provides a property rights perspective on the success of corporate innovation processes. We argue that the commercial success of a patent often depends on access to complementary patents not under direct control of its innovator. From a property rights perspective, the "extended boundary" of a downstream innovating firm would include such complementary patents if the downstream innovator and its upstream firm owning those complementary patents are linked together by common shareholders holding joint equity stake in the two firms.

Our identification strategy is based on patent documents that directly list related precursory patents, which may have rival patent claims to new products. We define *shareholder overlap* (SOL) as the (importance-weighted) aggregate minimum ownership share that investors own jointly in both the downstream innovating firm and the firms controlling the complementary assets—an innovating firm with a large SOL value can be interpreted as having an extended firm boundary.

We document the role of *shareholder overlap* for patent success at both the firm and patent levels; it correlates positively with both the intensive and extensive margins of patent production in an economically significant manner. This finding is robust to a variety of control variables and the inclusion of time and firm (or industry) fixed effects. Using interacted firm and time fixed effects, we show that two patents from the same yearly cohort filed by the same firm perform differently depending on their respective patent-level shareholder overlap. In addition, we instrument the patent-level shareholder overlap with the average size of the patent-specific upstream firms holding complementary patents and find a qualitatively similar result to the estimated relation. We also apply two placebo tests to show that the citation link to the upstream patent is crucial for the holdup attenuation effect of shareholder overlap and that the relationship between patent success and shareholder overlap does not appear to be driven by reverse causality.

We highlight two further dimensions of ownership structure. First, shareholder overlap coming from more dedicated investors tends to contribute more to the holdup attenuation—suggesting that the "extended boundary" of the innovating firm also depends on the types of institutional shareholders. Second, the ownership concentration of shareholder overlap matters independently of the overlap level. This could be explained by the existence of coordination and free-rider problems among a large and dispersed group of overlapping shareholders.

References

[1] Aghion, P., J. Van Reenen, and L. Zingales, 2013, Innovation and institutional ownership, *American Economic Review* 103(1), 277–304.

[2] Alcácer, J., M. Gittelman, and B. Sampat, 2009, Applicant and examiner citations in U.S. patents: An overview and analysis, *Research Policy* 38(2), 415–427.

[3] Asker, J., J. Farre-Mensa, and A. Ljungqvist, 2015, Corporate investment and stock market listing: a puzzle? *Review of Financial Studies* 28, 342–390.

[4] Azar, J., M. C. Schmalz, and, I. Tecu, 2017, Anti-competitive effects of common ownership, *Journal of Finance*, forthcoming.

[5] Bena, J., M. A. Ferreira, P. P. Matos, and P. Pires, 2017, Are foreign investors locusts? The long-term effects of foreign institutional ownership, *Journal of Financial Economics* 126 (1), 122–146.

[6] Bessen, James, 2004, Holdup and licensing of cumulative innovations with private information, *Economic Letters* 82(3), 321–326.

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

[8] Blundell, R., R. Griffith, and J. Van Reenen, 1999, Market share, market value and iInnovation in a panel of British manufacturing firms, *Review of Economic Studies* 66 (3), 529–54.

[9] Brown, J.R. , G. Martinsson, B.C. Petersen, 2013, Law, stock markets and innovation, *Journal of Finance* 68, 1517–1549.

[10] Brown, J.R. , G. Martinsson, B.C. Petersen, 2017, Stock markets, credit markets, and technology-led growth, *Journal of Financial Intermediation* 32, 45–59.

[11] Chemmanur T., Y. Shen, and J. Xie, 2017, Innovation beyond firm boundaries: Common blockholders, strategic alliances, and corporate innovation, Working paper, Boston College.

[12] Coase, R. H., 1937, The nature of the firm, *Economica* 4, 386–405.

[13] Creighton, S. A. , and S. A. Sher, 2009, Resolving patent disputes through merger: A comparison of three potential approaches, *Antitrust Law Journal* 75(3), 657–690.

[14] Engelberg, J., P. Gao, and C. Parsons, 2013, The price of a CEO's rolodex, *Review of Financial Studies* 26(1), 79–114.

[15] Galasso, A., and M. Schankerman, 2010, Patent thickets, courts and the market for innovation, *RAND Journal of Economics* 41(3), 472–503.

[16] Gaspar, J., M. Massa, and P. Matos, 2005, Shareholder investment horizons and the market for corporate control, *Journal of Financial Economics* 76(1), 135–165.

[17] Griliches, Z., A. Pakes, and B. H. Hall, 1988, The value of patents as indicators of inventive activity, Working paper. Available at: http://www.nber.org/papers/w2083.

[18] Grossman, S. J., and O. Hart, 1986, The costs and benefits of ownership: A theory of vertical and lateral integration, *Journal of Political Economy* 94(4), 691–719.

[19] Hall, B. H., A. B. Jaffe, and M. Trajtenberg, 2001, The NBER patent citation data file: lessons, insights and methodological tools, Working paper. Available at: http://www.nber.org/papers/w8498.

[20] Hall, B. H., A. B. Jaffe, and M. Trajtenberg, 2005, Market value and patent citations, *RAND Journal of Economics* 36(1), 16–38.

[21] Hall, B. H., and R. H. Ziedonis, 2007, The patent paradox revisited: An empirical study of patenting in the U.S. semiconductor industry, 1979-1995, *Rand Journal of Economics* 32(1), 101–128.

[22] Hansen, R. G., J. R. Lott, 1996, Externalities and corporate objectives in a world with diversified shareholder/consumers, *Journal of Financial and Quantitative Analysis* 31(1), 43–68.

[23] Harford, J., A. Kecskés, S. Mansi, 2017, Do long-term investors improve corporate decision making? *Journal of Corporate Finance*, forthcoming.

[24] Harhoff, D., F. Narin, F. M. Scherer, and K. Vopel, 1999, Citation frequency and the value of patented inventions, *Review of Economics and Statistics* 81(3), 511–515.

[25] Hart, O., 1995, *Firms, contracts, and financial structure*, Clarendon. Oxford University Press, New York.

[26] Hart, O., and J. Moore, 1990, Property rights and the nature of the firm, *Journal of Political Economy* 98(6), 1119–1158.

[27] He, J., and J. Huang, 2017, Product market competition in a world of cross ownership: Evidence from institutional blockholdings, *Review of Economics and Statistics*, forthcoming.

[28] Holthausen, R. W., R. W. Leftwich, and D. Mayers, 1990, Large-block transactions, the speed of response, and temporary and permanent stock-price effects, *Journal of Financial Economics* 26, 71–95.

[29] Hsu, P., X. Tian, and Y. Xu, 2014, Financial development and innovation: Cross-country evidence, *Journal of Financial Economics* 112(1), 116–135.

[30] Keim, D. B., and A. Madhavan, 1996, The upstairs market for large-block transactions: analysis and measurement of price effects, *Review of Financial Studies* 9, 1–36.

[31] Klein, B., R. Crawford, and A. Alchian, 1978, Vertical integration, appropriable rents, and the competitive contracting process, *Journal of Law and Economics* 21, 297–326.

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

[33] Lerner, J., M. Sørensen, and P. Strömberg, 2011, Private equity and long-run investment: The case of innovation, *Journal of Finance* 66(2), 445–477.

[34] McCahery, J. A., Z. Sautner, and L. T. Starks, 2016, Behind the scenes: The corporate governance preferences of institutional investors, *Journal of Finance* 71, 2905–2932.

[35] Noel, M., and M. Schankerman, 2013, Strategic patenting and software innovation, *Journal of Industrial Economics* 61(3), 481–520.

[36] Schmidt, J. J., 2012, Perceived auditor independence and audit litigation: The role of nonaudit services fees, *Accounting Review* 87(3), 1033–1065.

[37] Shapiro, C., 2001, Navigating the patent thicket: Cross-licenses, patent pools, and standard-setting., A. Jaffe, J. Lerner, S. Stern, eds. *Innovation Policy and the Economy.* NBER, Cambridge, MA.

[38] Solow, R. M., 1957, Technical change and the aggregate production function, *Review of Economics and Statistics* 39, 312–320.

[39] Van Nieuwerburgh, S., and L. Veldkamp, 2010, Information acquisition and portfolio under-diversification, *Review of Economic Studies* 77, 779–805.

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

[41] Williamson, O. E., 1975, *Markets and Hierarchies: Analysis and Antitrust Implications* (Free Press, New York).

[42] Williamson, O. E. , 1985, *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting* (Free Press, New York).

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

Table 1: Summary Statistics

Reported are the summary statistics of the regression variables. Firm-level dependent variables are (i) $CITES_{s,t}$, the number of future citations received by the cohort of patents filed by firm s in year t; (ii) $N_{s,t}$, the number of patents filed by firm s in year t; (iii) $\overline{cites}_{s,t}$, the average future citation count per patent for the cohort of patents filed by firm s in year t; (iv) $R\&D\ Exp_{s,t}/Asset_{s,t}$, the R&D expenditure to the total firm assets ratio for firm s in year t, (v) $CITES^{F}_{s,t}$, a filtered citation measure, which removes all citations coming from those upstream firms that firm s has cited in its patent filings in year t, and (vi) $CITES^{3yr}_{s,t}$, the total citations received during the calendar year of the patent grant and the three subsequent years for the cohort of patents filed by firm s in year t. At the patent level we denote by $cites_{p,t}$ the total number of future citations received by patent p, filed in year t. $SOL_{s,t-1}$ and $sol_{p,t-1}$, refer to the shareholder overlap for, respectively, firm s and patent p measured at the end of year $t-1$. We decompose $SOL_{s,t-1}$ into three components, representing the shareholder overlap originating from dedicated investors ($SOL\_Dedicated_{s,t-1}$), intermediate investors ($SOL\_Intermediate_{s,t-1}$), and transient investors ($SOL\_Transient_{s,t-1}$). $SOL\_Placebo1_{s,t-1}$ and $SOL\_Placebo2_{s,t-1}$ are the two placebo measures of shareholder overlap. We also consider three alternative shareholder overlap measures: $SOL^{3yr}_{s,t-1}$ (a cites-weighted measure based on citations received during the calendar year of the patent grant and the three subsequent years), $SOL\_Rank_{s,t-1}$ (a non-parametric rank-based measure), and $SOL\_Equal_{s,t-1}$ (an equal-weight measure). $MktCap^{u}_{p,t-1}$ denotes the average market capitalization value at the end of year $t-1$ of firms owning patent p's upstream patents p^u. $IO_{s,t-1}$, $IO^{SOL}_{s,t-1}$, and $IO^{NOL}_{s,t-1}$ represent the institutional ownership of, respectively, all shareholders, overlapping shareholders, and non-overlapping shareholders in firm s at the end of year $t-1$. $SIF_{s,t-1}$ and $WHHI_{s,t-1}$ are, respectively, the shareholder innovation focus and the weighted Herfindahl-Hirschman Index of the ownership concentration of overlapping shareholders in firm s at the end of year $t-1$. $ln(SpillTECH_{s,t-1})$ and $ln(SpillSIC_{s,t-1})$ measures the extent of technology spillover and product market rivalry effect of $R\&D$, respectively, for firm s in year $t-1$. The control variables include the market capitalization ($MktCap_{s,t-1}$), cumulative R&D investment ($R\&D\ Stock_{s,t-1}$), capital to labor ratio ($K/L_{s,t-1}$), sales ($Sales_{s,t-1}$), and the average proportion of privately owned upstream patents ($Private\ Patent\ Share_{s,t-1}$) for firm s in year $t-1$. Detailed definitions of the variables are given in Appendix B.

	Obs.	Mean	Median	STD	Skewness	Min.	P10	P90	Max.
Dependent Variables (measured in year t)									
$ln(1+CITES)$	19,020	3.948	3.912	2.065	0.116	0.000	1.317	6.619	11.640
$ln(1+N)$	19,020	1.964	1.609	1.340	1.351	0.693	0.693	3.912	8.395
$ln(1+\overline{cites})$	19,020	2.388	2.459	1.140	−0.181	0.000	0.886	3.772	6.643
$R\&D\ Exp/Assets$	19,020	0.082	0.029	0.153	5.777	0.000	0.000	0.210	3.704
$ln(1+cites)$	581,240	1.899	1.962	1.357	0.121	0.000	0.000	3.660	7.129
$ln(1+CITES^{F})$	19,020	3.904	3.869	2.048	0.121	0.000	1.287	6.549	11.565
$ln(1+CITES^{3yr})$	19,020	2.672	2.485	1.921	0.540	0.000	0.000	5.276	10.701
$ln(CITES)$	17,609	4.214	4.091	1.865	0.390	0.180	1.877	6.707	11.640
Independent Variables (measured in year $t-1$)									
SOL	19,020	0.062	0.044	0.063	1.487	0.000	0.000	0.150	0.541
$SOL\_Dedicated$	19,020	0.002	0.000	0.004	8.873	0.000	0.000	0.004	0.174
$SOL\_Intermediate$	19,020	0.027	0.019	0.027	1.610	0.000	0.000	0.064	0.248
$SOL\_Transient$	19,020	0.031	0.020	0.034	1.697	0.000	0.000	0.079	0.289
$SOL\_Placebo1$	19,020	0.050	0.038	0.048	1.513	0.000	0.000	0.114	0.483
$SOL\_Placebo2$	19,020	0.048	0.036	0.047	1.627	0.000	0.000	0.112	0.564
SOL^{3yr}	19,020	0.062	0.043	0.064	1.577	0.000	0.000	0.150	0.750
SOL_Rank	19,020	0.063	0.045	0.063	1.468	0.000	0.000	0.150	0.522
SOL_Equal	19,020	0.172	0.164	0.120	0.436	0.000	0.000	0.336	0.727
sol	581,240	0.144	0.111	0.142	1.169	0.000	0.000	0.342	0.850
$ln(MktCap^{u})$	581,240	8.023	7.574	4.359	0.430	0.042	2.555	14.361	20.216
SIF	19,020	0.281	0.280	0.075	2.294	0.000	0.197	0.367	2.644
$WHHI$	19,020	0.057	0.035	0.071	3.300	0.000	0.000	0.133	1.000
IO	19,020	0.479	0.497	0.266	−0.067	0.000	0.100	0.823	1.000
IO^{SOL}	19,020	0.379	0.367	0.278	0.182	0.000	0.000	0.760	1.000
IO^{NOL}	19,020	0.100	0.037	0.158	2.652	0.000	0.000	0.282	1.000
$ln(SpillTECH)$	19,020	10.607	10.743	1.064	−1.027	1.887	9.244	11.830	12.747
$ln(SpillSIC)$	18,945	8.659	9.061	2.275	−1.157	−8.085	5.650	11.229	12.599
Controls (measured in year $t-1$)									
$ln(MktCap)$	19,020	13.034	12.873	2.100	0.315	6.736	10.462	15.894	20.216
$ln(1+R\&D\ Stock)$	19,020	3.764	3.903	2.232	0.051	0.000	0.000	6.563	10.714
$ln(K/L)$	19,020	4.406	4.319	0.906	0.625	−1.410	3.372	5.533	10.296
$ln(Sales)$	19,020	5.428	5.464	2.564	−0.321	−6.215	2.239	8.685	12.722
$Private\ Patent\ Share$	19,020	0.735	0.769	0.200	−0.891	0.000	0.466	1.000	1.000

Table 2: Baseline Regressions

Reported are the firm-level OLS regressions of patent success ($ln(1+CITES_{s,t})$) on lagged *shareholder overlap* ($SOL_{s,t-1}$) for the sample period 1992–2007. $CITES_{s,t}$ denotes the number of future citations received by the cohort of patents filed by firm s in year t. $SOL_{s,t-1}$ measures the average shareholder ownership overlap at the end of year $t-1$ between the innovating firm s and other firms owning the upstream complementary patents. Columns 1–2 report the full sample results, whereas Columns 3–4 report the subsample results for the top three R&D-intensive sectors (pharmaceuticals, computer hardware, and telecommunications equipment). The control variables include the market capitalization ($MktCap_{s,t-1}$), cumulative R&D investment ($R\&D\ Stock_{s,t-1}$), capital to labor ratio ($K/L_{s,t-1}$), sales ($Sales_{s,t-1}$), and the average proportion of privately owned upstream patents ($Private\ Patent\ Share_{s,t-1}$) for firm s in year $t-1$. All regressions control for a full set of year dummies and industry dummies based on four-digit SIC codes. Firm fixed effects are based on Blundell, Griffith, and Van Reenen (1999). Robust standard errors clustered at the firm level are reported in parentheses. Also reported are the total number of observations (*Obs.*) and adjusted R-squared (*Adj.* R^2). ** and * denote the 1% and 5% significance level, respectively.

Dependent Variable:	$ln(1+CITES)$			
	Full Sample		Top 3 R&D-Intensive Industries	
	(1)	(2)	(3)	(4)
SOL	3.692**	3.450**	4.685**	4.328**
	(0.495)	(0.484)	(0.826)	(0.805)
Controls:				
$ln(MktCap)$	0.317**	0.306**	0.381**	0.367**
	(0.018)	(0.018)	(0.032)	(0.031)
$ln(1+R\&D\ Stock)$	0.317**	0.251**	0.276**	0.213**
	(0.016)	(0.016)	(0.037)	(0.036)
$ln(K/L)$	0.029	−0.009	0.107*	0.040
	(0.032)	(0.031)	(0.059)	(0.056)
$ln(Sales)$	−0.024	−0.076**	−0.008	−0.055*
	(0.015)	(0.015)	(0.024)	(0.024)
Private Patent Share	0.422**	0.321**	0.624**	0.475*
	(0.120)	(0.117)	(0.195)	(0.189)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Firm FE (BGV)	No	Yes	No	Yes
Obs.	19,020	19,020	5,898	5,898
Adj. R^2	0.526	0.542	0.564	0.577

Table 3: Intensive versus Extensive Margin

Reported are OLS regressions of the intensive margin ($ln(1+\overline{cites}_{s,t})$) and the extensive margin ($ln(1+N_{s,t})$) of patent production on the lagged *shareholder overlap* ($SOL_{s,t-1}$) for the sample period 1992–2007. $N_{s,t}$ denotes the number of patents filed by firm s in year t, and $\overline{cites}_{s,t}$ denotes the average future citation count per patent for the cohort of patents filed by firm s in year t. $SOL_{s,t-1}$ measures the average shareholder ownership overlap at the end of year $t-1$ between the innovating firm s and other firms owning the upstream complementary patents. Columns 1–2 and 3–4 report the results for, respectively, the intensive margin and extensive margin of patent production. The control variables include the market capitalization ($MktCap_{s,t-1}$), cumulative R&D investment ($R\&D\ Stock_{s,t-1}$), capital to labor ratio ($K/L_{s,t-1}$), sales ($Sales_{s,t-1}$), and the average proportion of privately owned upstream patents ($Private\ Patent\ Share_{s,t-1}$) for firm s in year $t-1$. All regressions control for a full set of year dummies and industry dummies based on four-digit SIC codes. Firm fixed effects are based on Blundell, Griffith, and Van Reenen (1999). Robust standard errors clustered at the firm level are reported in parentheses. Also reported are the total number of observations (*Obs.*) and adjusted R-squared (*Adj.* R^2). ** and * denote the 1% and 5% significance level, respectively.

Dependent Variables:	$ln(1+\overline{cites})$		$ln(1+N)$	
	(1)	(2)	(3)	(4)
SOL	0.584*	0.527*	2.923**	2.936**
	(0.248)	(0.247)	(0.375)	(0.375)
Controls:				
$ln(MktCap)$	0.081**	0.081**	0.199**	0.184**
	(0.010)	(0.010)	(0.012)	(0.011)
$ln(1+R\&D\ Stock)$	0.023**	0.017*	0.262**	0.183**
	(0.007)	(0.008)	(0.013)	(0.011)
$ln(K/L)$	−0.039*	−0.049**	0.060**	0.051*
	(0.017)	(0.017)	(0.022)	(0.020)
$ln(Sales)$	−0.058**	−0.067**	0.036**	−0.007
	(0.009)	(0.009)	(0.009)	(0.009)
Private Patent Share	0.073	0.056	0.350**	0.282**
	(0.070)	(0.070)	(0.079)	(0.078)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Firm FE (BGV)	No	Yes	No	Yes
Obs.	19,020	19,020	19,020	19,020
Adj. R^2	0.427	0.429	0.614	0.648

Table 4: Structure of Shareholder Overlap and Placebo Measures

Column 1 of the table reproduces the baseline regression in Table 2, Column 1. In Column 2, we decompose $SOL_{s,t-1}$ into three components, representing the shareholder overlap originating from dedicated investors ($SOL\_Dedicated_{s,t-1}$), intermediate investors ($SOL\_Intermediate_{s,t-1}$), and transient investors ($SOL\_Transient_{s,t-1}$). At the end of each year, we rank all institutional investors along two dimensions: Their portfolio concentration (in descending order) and portfolio turnover (in ascending order). We label dedicated, intermediate, and transient investors, respectively, as those in the top, middle, and bottom tercile of the combined rank of the two dimensions of shareholder activism. Column 3 expands the baseline regression by including the Weighted Herfindahl-Hirschman Index of the ownership concentration of overlapping shareholders, $WHHI_{s,t-1}$. Columns 4-5 report the regression results of the two placebo tests, in which we replace $SOL_{s,t-1}$ in the baseline regression with a placebo shareholder overlap measure ($SOL\_Placebo1_{s,t-1}$ or $SOL\_Placebo2_{s,t-1}$). The control variables include the market capitalization ($MktCap_{s,t-1}$), cumulative R&D investment ($R\&D\ Stock_{s,t-1}$), capital to labor ratio ($K/L_{s,t-1}$), sales ($Sales_{s,t-1}$), and the average proportion of privately owned upstream patents ($Private\ Patent\ Share_{s,t-1}$) for firm s in year $t-1$. The sample period is 1992–2007. We report in the last row the p-value for the null hypothesis of equal coefficients in Column 2. All regressions control for a full set of year dummies and industry dummies based on four-digit SIC codes. Firm fixed effects are based on Blundell, Griffith, and Van Reenen (1999). Robust standard errors clustered at the firm level are reported in parentheses. Also reported are the total number of observations (*Obs.*) and adjusted R-squared (*Adj.* R^2). ** and * denote the 1% and 5% significance level, respectively.

Dependent Variable.:	$ln(1+CITES)$				
	(1)	(2)	(3)	(4)	(5)
SOL	3.450**		4.361**		
	(0.484)		(0.502)		
$SOL\_Dedicated$		19.926**			
		(5.090)			
$SOL\_Intermediate$		0.572			
		(0.975)			
$SOL\_Transient$		5.230**			
		(0.869)			
$WHHI$			2.206**		
			(0.264)		
$SOL\_Placebo1$				−0.082	
				(0.556)	
$SOL\_Placebo2$					0.518
					(0.524)
Controls:					
$ln(MktCap)$	0.306**	0.293**	0.324**	0.345**	0.341**
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
$ln(1+R\&D\ Stock)$	0.251**	0.250**	0.252**	0.258**	0.258**
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
$ln(K/L)$	−0.009	−0.008	−0.003	−0.004	−0.005
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)
$ln(Sales)$	−0.076**	−0.075**	−0.072**	−0.070**	−0.070**
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Private Patent Share	0.321**	0.280*	0.958**	−0.305*	−0.215
	(0.117)	(0.117)	(0.147)	(0.121)	(0.116)
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Firm FE (BGV)	Yes	Yes	Yes	Yes	Yes
Obs.	19,020	19,020	19,020	19,020	19,020
Adj. R^2	0.542	0.543	0.545	0.539	0.539
$H_0: SOL\_Dedicated = SOL\_Transient$		0.000			

Table 5: Patent-Level Regressions

Reported are the patent-level OLS and 2SLS regressions of patent success ($ln(1+cites_{p,t})$) on lagged *shareholder overlap* ($sol_{p,t-1}$) for the sample period 1992–2007. Because we control for firm-year fixed effects in the regressions, we discard all firm years that feature only one patent application. $cites_{p,t}$ denotes the total number of future citations received by patent p, filed in year t. The variable $sol_{p,t-1}$ measures the average shareholder ownership overlap at the end of year $t-1$ between the firm owning patent p and other firms owning the upstream complementary patents. Columns 1 and 2 report the patent-level OLS regression results, controlling for year, firm, and technology field fixed effects or the interacted year-firm and year-technology field fixed effects. Columns 3 and 4 report, respectively, the first and second stage result of the 2SLS regression, with $ln(MktCap^u_{p,t-1})$ as an instrumental variable for $sol_{p,t-1}$. $MktCap^u_{p,t-1}$ denotes the average market capitalization value, at the end of year $t-1$, of firms owning patent p's upstream patents p^u. Columns 5 and 6 report the regression results based on financial institution mergers. For each merger deal, we consider a seven-year event window centered around the year of the merger event. *Post-Merger* is a dummy of 1 for observations during the post-merger period. *Treat* is a dummy of 1 for treatment patents. The two regressions control for the interacted merger event-downstream firm-patent class fixed effects. Robust standard errors are reported in parentheses, which are clustered at the firm-year level for Columns 1–4 and at merger event-downstream firm level for Columns 5–6. Also reported are the total number of observations (*Obs.*) and adjusted R-squared (*Adj.* R^2). ** and * denote the 1% and 5% significance level, respectively.

	Full Sample OLS		Full Sample 2SLS		Institution Mergers OLS	
			1^{st} Stage	2^{nd} Stage		
Dependent Variable:	$ln(1+cites)$	$ln(1+cites)$	sol	$ln(1+cites)$	sol	$ln(1+cites)$
	(1)	(2)	(3)	(4)	(5)	(6)
sol	0.192^{**}	0.272^{**}		0.283^{**}		
	(0.019)	(0.018)		(0.013)		
$ln(MktCap^u)$			0.024^{**}			
			(0.000)			
Treat × *Post-Merger*					0.017^{**}	0.135^{**}
					(0.006)	(0.046)
Post-Merger					0.001	-0.566^{**}
					(0.004)	(0.035)
Treat					0.071^{**}	0.066^{**}
					(0.008)	(0.025)
Year FE	Yes	No	No	No	No	No
Tech. FE	Yes	No	No	No	No	No
Firm FE	Yes	No	No	No	No	No
Year × Firm FE	No	Yes	Yes	Yes	No	No
Year × Tech. FE	No	Yes	Yes	Yes	No	No
Event × Firm × Tech. FE	-	-	-	-	Yes	Yes
Obs.	581,240	581,240	581,240	581,240	88,090	88,090
Adj. R^2	0.312	0.339	0.851		0.251	0.266

Table 6: R&D Expenditure and Shareholder Overlap

Reported are OLS regressions of the R&D expenditure (relative to assets) for the sample period 1992–2007. $R\&D\ Exp_{s,t}/Assets_{s,t}$ denotes the R&D expenditure to the total firm assets ratio for firm s in year t. $SOL_{s,t-1}$ measures the average shareholder ownership overlap at the end of year $t-1$ between the innovating firm s and other firms owning the upstream complementary patents. $IO_{s,t-1}$, $IO^{SOL}_{s,t-1}$, and $IO^{NOL}_{s,t-1}$ represent the institutional ownership of, respectively, all shareholders, overlapping shareholders, and non-overlapping shareholders in firm s at the end of year $t-1$. The control variables include the capital to labor ratio ($K/L_{s,t-1}$), sales ($Sales_{s,t-1}$), and the average proportion of privately owned upstream patents ($Private\ Patent\ Share_{s,t-1}$) for firm s in year $t-1$. All regressions control for a full set of year dummies and industry dummies based on four-digit SIC codes. Robust standard errors clustered at the firm level are reported in parentheses. Also reported are the total number of observations (*Obs.*) and adjusted R-squared (*Adj.* R^2).

Dependent Variable:	*R&D Exp/Assets*		
	(1)	(2)	(3)
SOL	0.240**	0.237**	
	(0.043)	(0.042)	
IO		0.004	
		(0.007)	
IO^{SOL}			0.025**
			(0.008)
IO^{NOL}			−0.047**
			(0.009)
Controls:			
$ln(K/L)$	0.008**	0.008**	0.009**
	(0.002)	(0.002)	(0.002)
$ln(Sales)$	−0.019**	−0.019**	−0.018**
	(0.002)	(0.002)	(0.002)
Private Patent Share	0.005	0.005	−0.021**
	(0.012)	(0.012)	(0.008)
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	No
Obs.	19,020	19,020	19,020
Adj. R^2	0.278	0.278	0.279

Table 7: Robustness

Column 1 of the table reproduces the baseline regression in Table 2, Column 1. Additional explanatory variables, including *institutional ownership* ($IO_{s,t-1}$), *shareholder innovation focus* ($SIF_{s,t-1}$), *knowledge spillover* ($ln(SpillTech_{s,t-1})$), and *product market rivalry effect of R&D* ($ln(SpillSIC_{s,t-1})$) are added to Columns 2–5. The dependent variable in Columns 1–5 is $ln(1+CITES_{s,t})$. All specifications are estimated using OLS regressions except for Column 6, which estimates a negative binomial model with $CITES_{s,t}$ as the dependent variable. Column 7 replaces $SOL_{s,t-1}$ in Column 1 with $SOL_{s,t-2}$, for which the ownership stake is measured at the end of year $t-2$. Columns 8–10 replace $SOL_{s,t-1}$ in Column 1 with, respectively, a non-parametric rank-based measure ($SOL\_Rank_{s,t-1}$), an equal-weight measure ($SOL\_Equal_{s,t-1}$), and an alternative cites-weighted measure ($SOL^{3yr}_{s,t-1}$) of shareholder overlap. Columns 10 and 11 measures the patent success by $ln(1+CITES^{3yr}_{s,t})$ and $ln(CITES_{s,t})$, respectively. Column 12 uses a filtered citation measure, $ln(1+CITES^{F}_{s,t})$, as the dependent variable, which removes all citations coming from those upstream firms that firm s has cited in its patent filings in year t. All regressions control for the firm market capitalization ($MktCap_{s,t-1}$), cumulative R&D investment ($R\&D\ Stock_{s,t-1}$), capital to labor ratio ($K/L_{s,t-1}$), sales ($Sales_{s,t-1}$), and the average proportion of privately owned upstream patents ($Private\ Patent\ Share_{s,t-1}$) for firm s in year $t-1$. All regressions control for a full set of year dummies and industry dummies based on four-digit SIC codes. Firm fixed effects are based on Blundell, Griffith, and Van Reenen (1999). Robust standard errors clustered at the firm level are reported in parentheses. Also reported are the total number of observations (*Obs.*) and adjusted R-squared (*Adj.* R^2). ** and * denote the 1% and 5% significance level, respectively. The variable definitions are the same as those in Table 1.

Dependent Var.:	$ln(1+CITES)$					Neg. Bino. $CITES$
	(1)	(2)	(3)	(4)	(5)	(6)
SOL	3.450**	3.429**	3.337**	3.476**	3.354**	2.633**
	(0.484)	(0.475)	(0.485)	(0.486)	(0.478)	(0.602)
IO		−0.567**			−0.560**	−0.702**
		(0.097)			(0.097)	(0.088)
SIF			0.722**		0.661**	0.755**
			(0.223)		(0.219)	(0.242)
$ln(SpillTECH)$				0.099**	0.093**	0.095**
				(0.032)	(0.032)	(0.032)
$ln(SpillSIC)$				−0.028	−0.028	−0.008
				(0.018)	(0.018)	(0.019)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year & Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE (BGV)	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	19,020	19,020	19,020	18,945	18,945	18,945
Adj. R^2	0.542	0.545	0.543	0.543	0.546	

Dependent Var.:	$ln(1+CITES)$			$ln(1+CITES^{3yr})$	$ln(CITES)$	$ln(1+CITES^{F})$
	(7)	(8)	(9)	(10)	(11)	(12)
SOL					3.224**	3.252**
					(0.481)	(0.484)
$SOL(t-2)$	2.385**					
	(0.432)					
SOL_Rank		3.621**				
		(0.482)				
SOL_Equal			2.518**			
			(0.200)			
SOL^{3yr}				2.515**		
				(0.373)		
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year & Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE (BGV)	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	18,109	19,020	19,020	19,020	17,609	19,020
Adj. R^2	0.539	0.543	0.547	0.605	0.525	0.538

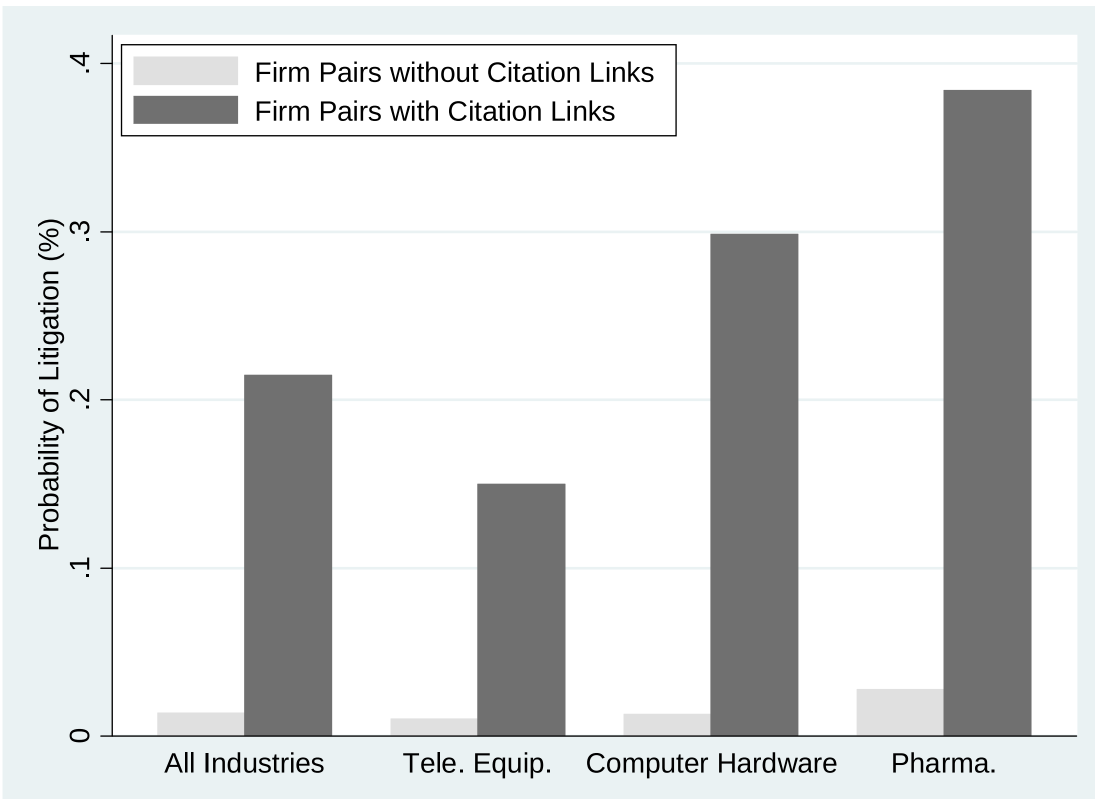

Figure 1: Probability of patent litigation for intra-industry firm pairs. This figure compares the likelihood of patent litigation for listed firm pairs with citation links and those without any citation link in the preceding three years, based on the litigation cases reported in the Audit Analytics Litigation database. Each year from 2000 to 2007, we form intra-industry firm pairs (based on the Fama-French 49 industry classification scheme) of all U.S. listed firms with at least one patent application in the past three years and sort them into pairs with at least one patent citation link and pairs without any such link. The litigation probability is 0.215% for the former and 0.014% for the latter in the full sample. The corresponding probabilities are 0.150% and 0.010% for the telecommunications equipment sector, 0.299% and 0.013% for the computer hardware sector, and 0.384% and 0.028% for the pharmaceuticals sector.

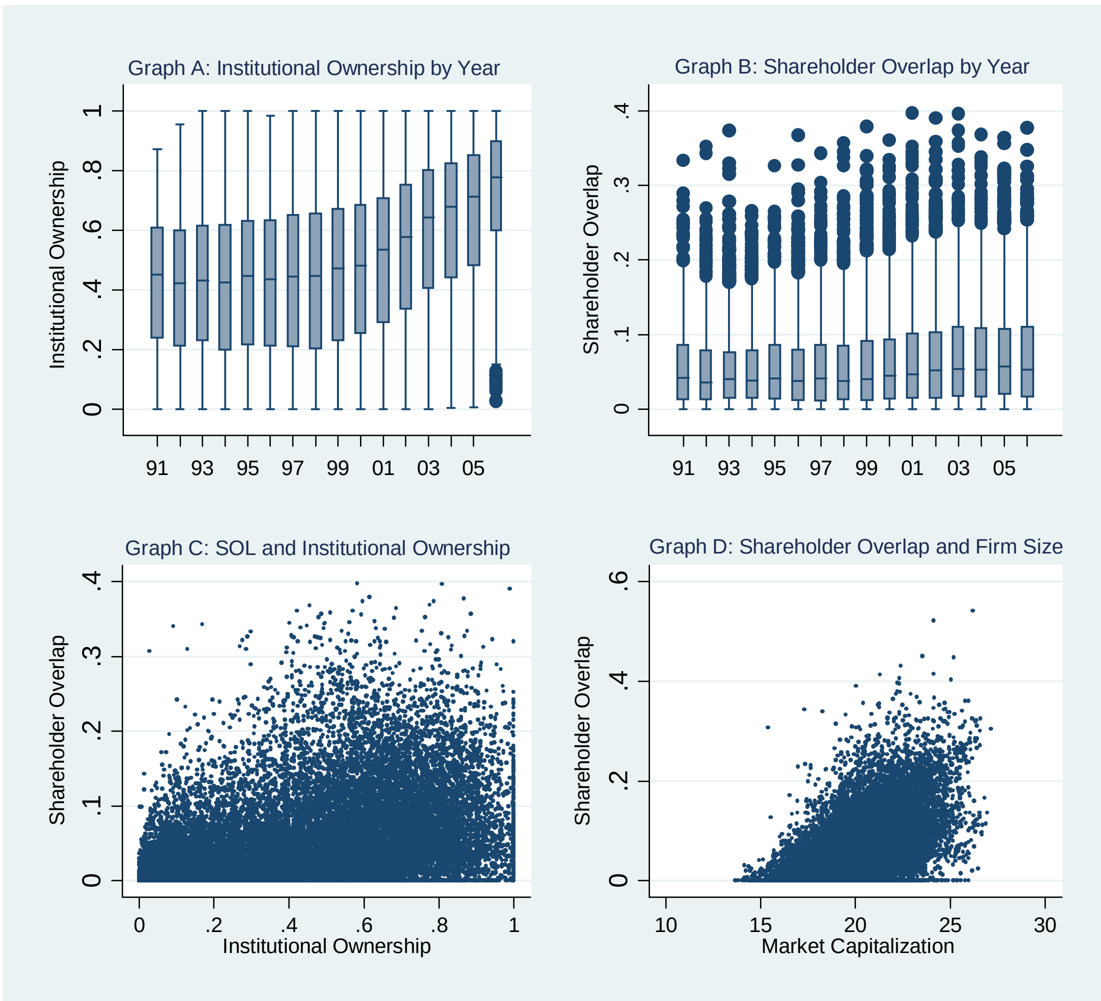

Figure 2: Institutional ownership and shareholder overlap. Graphs A and B are the box plots for the distribution of institutional ownership ($IO_{s,t}$) and shareholder overlap ($SOL_{s,t}$), respectively, by year from 1991 to 2006. The top, middle, and bottom values of each box represent the 75th, 50th, and 25th percentile of the distribution in the given year; the maximum and minimum of each vertical bar represent the upper and lower adjacent values, and the dots denote the observations outside the adjacent values. Graph C plots our sample along the dimension of shareholder overlap $SOL_{s,t}$ and institutional ownership $IO_{s,t}$, whereas Graph D plots along the dimension of shareholder overlap $SOL_{s,t}$ and firm size $MktCap_{s,t}$ for all firm-years.

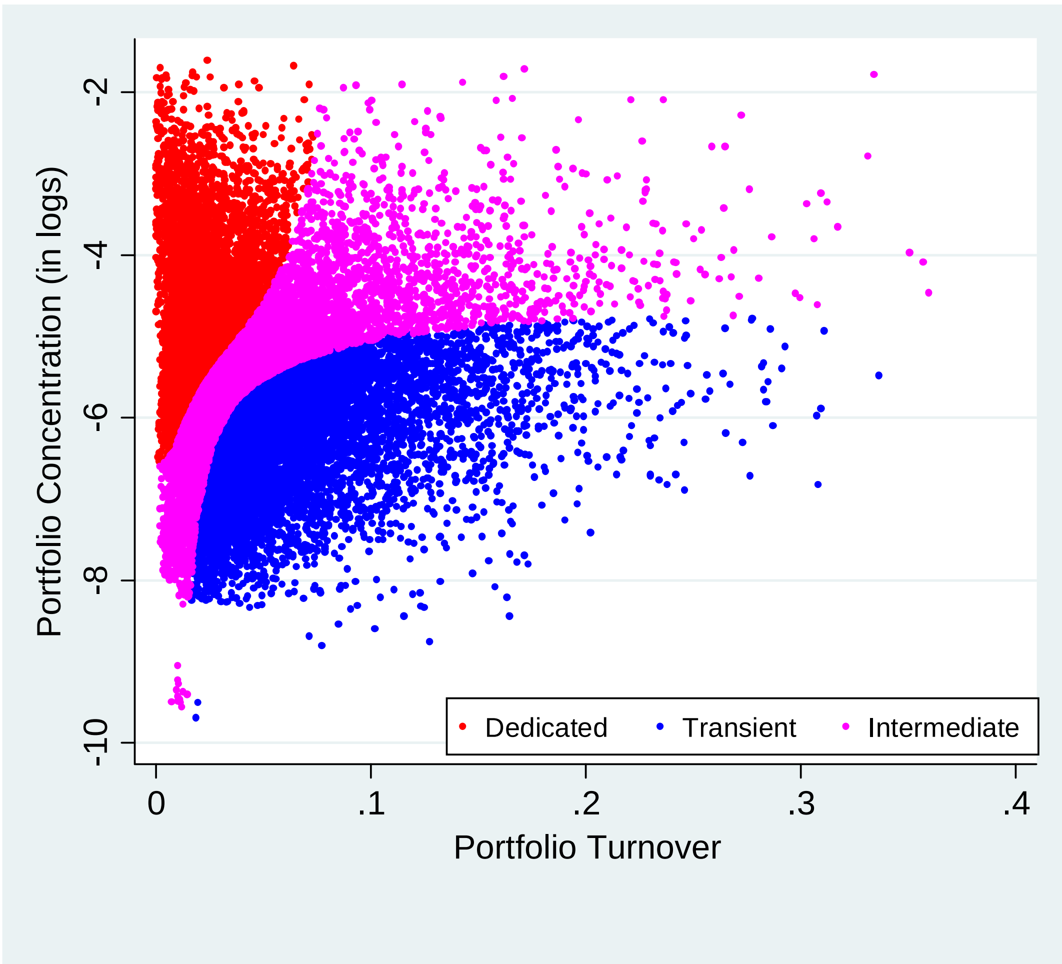

Figure 3: Shareholder overlap by investor types. Plotted are the (log) portfolio concentration and portfolio turnover of institutional investors (dedicated, intermediate, or transient investors) in our sample over the period 1991–2006. Specifically, at the end of each year, we rank all institutional investors along two dimensions: Their portfolio concentration (i.e., the Herfindahl-Hirschman Index of their equity portfolio holdings) in descending order and their portfolio turnover ratios in ascending order. We label the dedicated, intermediate, and transient investors as those in the top, middle, and bottom tercile of the combined rank, respectively.

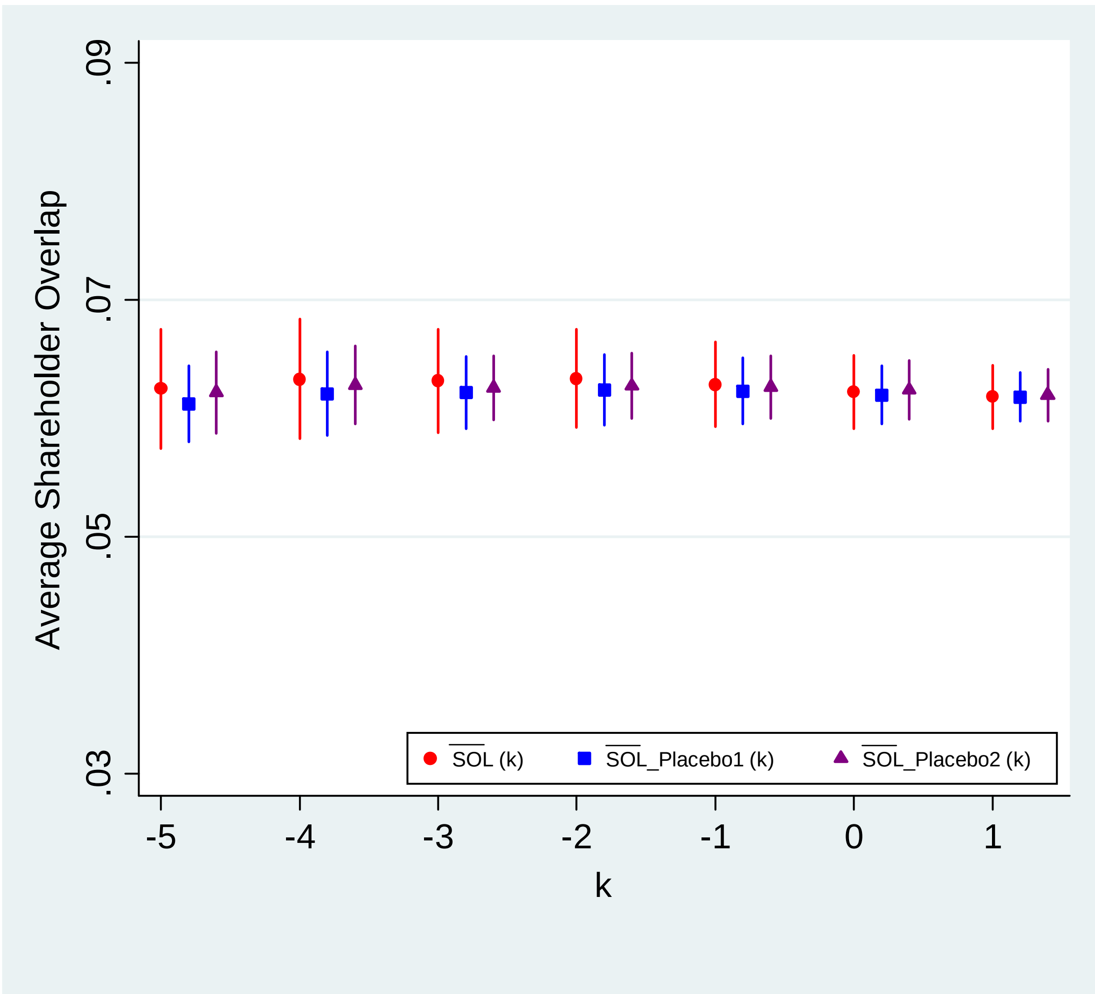

Figure 4: Evolution of shareholder overlap. The evolution of the average shareholder overlap $\overline{SOL}(k)$ between the innovating firm and other firms owning complementary upstream patents is plotted for the period from five years prior to the patent filing year to one year after the filing (i.e., $k = -5$ to 1), with the patent filing year denoted by $k = 0$. Each dot in the figure denotes the mean value of shareholder overlap for the given year k relative to the patent filing year, and the vertical segment above and below the dot denotes the standard deviation of the distribution of shareholder overlap for the given year. The evolution of the two placebo measures of shareholder overlap is also plotted. For ease of comparison, we adjust the value of $\overline{SOL}\_Placebo1(k)$ and $\overline{SOL}\_Placebo2(k)$ upward by 0.012 and 0.014, respectively, so that they would have the same mean value as $\overline{SOL}(k)$ for $k = 0$.

Internet Appendix

(Not for Journal Publication)

Patent Success, Patent Holdup and the Structure of Property Rights

Heng Geng*

Victoria University of Wellington

Harald Hau**

University of Geneva and Swiss Finance Institute

Sandy Lai***

National Taiwan University

Appendix A. A Model of Patent Investment

A.1 A Simple Benchmark (with No Holdup Effect)

A risk-neutral firm s can invest in a continuum of patent projects. Each project is represented by the index number p on the interval $[0, \infty)$, with a higher index number corresponding to higher patent development costs. For simplicity, we assume a continuous increasing convex cost function $C(p)$, with $C'(p) > 0$ and $C''(p) > 0$, for the patent development costs. The present value from commercialization of the patent project, $V_s(p)$, is proportional to the success of the patent, proxied by the future citation count $cites_s(p)$.[20] Hence,

$$V_s(p) = \alpha \times cites_s(p), \tag{A.1}$$

where $cites_s(p)$ is a random variable with the expected value $E[cites_s(p)] = \mu_s$, and $\alpha > 0$ is a constant. The total expected firm value Π_s follows as

$$\Pi_s = \max_{\overline{p}} \int_0^{\overline{p}} [\alpha \mu_s - C(p)] \, dp, \tag{A.2}$$

where the interval $[0, \overline{p}]$ denotes the range of patent projects the firm pursues. Value maximization implies the first-order condition $\alpha \mu_s = C(\overline{p})$. For a convex cost function $C(p) = cp^b$ $(b > 1)$, we find that

$$\overline{p} = \left(\frac{\alpha \mu_s}{c} \right)^{\frac{1}{b}} \tag{A.3}$$

characterizes the optimal range of patent production. We summarize the model implications as follows:

Proposition 1: Patent Production without Patent Holdup

A value maximizing firm optimally invests in the production of patents on the line interval $[0, \overline{p}]$. Given a patent-level expected citation count $E[cites_s(p)] = \mu_s$ that is proportional to each patent's expected value and a convex cost function $C(p) = cp^b$, we find for

(i) the (log) extensive margin of patent production

$$ln[\overline{p}] = \frac{1}{b} ln \frac{\alpha}{c} + \frac{1}{b} ln(\mu_s) \tag{A.4}$$

(ii) the (log) intensive margin of patent production

$$ln[E(cites_s)] = ln(\mu_s) \tag{A.5}$$

[20] See, for example, Harhoff et al. (1999) and Kogan et al. (2017) for a positive relation between future citation count and the economic value of a patent, and Hall et al. (2005) for a positive link between patent citation count and firm value.

(iii) the firm-level (log) citation count

$$ln[E(CITES_s)] = ln\int_0^{\overline{p}} E[cites_s(p)]dp = \frac{1}{b}ln\frac{\alpha}{c} + \frac{b+1}{b}ln(\mu_s), \qquad \text{(A.6)}$$

(iv) the (log) R&D expenditure

$$ln[R\&D\ Exp] = ln\int_0^{\overline{p}} cp^b dp = ln\frac{c}{1+b} + \frac{b+1}{b}ln\frac{\alpha\mu_s}{c}. \qquad \text{(A.7)}$$

The firm-level (log) citation count in Eq. (A.6) is equal to the sum of the (log) extensive margin in Eq. (A.4) and the (log) intensive margin in Eq. (A.5). Empirically, we can approximate the intensive margin by the average citation count $\overline{cites}_s$ of a firm's patents.

A.2 The Case of Efficient Ex-Ante Contracting

Next, we enrich the model and assume that commercialization of each new downstream patent p requires licensing of complementary upstream patents p^u, $u = 1, 2, ...N_p$, and that efficient ex-ante contracting prior to patent investment is possible. We assume that such contracting transfers a share $L_s(p, p_u)$ of patent p's net profit to the respective upstream firm owning patent p^u. The value of $L_s(p, p_u)$, with $0 < L_s(p, p_u) < 1$, captures the relative strength of the negotiating position of the upstream and downstream firms. For simplicity, we assume that the ex-ante expected share of value loss is identical for all patents p produced by the same firm s, denoted by $\overline{L}_s$. The profit function of the downstream innovating firm becomes

$$\Pi_s = \max_{\overline{p}}(1 - \overline{L}_s)\int_0^{\overline{p}} [\alpha\mu_s - C(p)]\, dp. \qquad \text{(A.8)}$$

It is straightforward to see that the term $1 - \overline{L}_s$ does not enter the first-order conditions for the investment choice of the downstream firm.

Proposition 2: Efficient Ex-Ante Contracting with Upstream Patent Owners

If commercialization of each downstream patent p requires access to complementary upstream patents p^u, efficient ex-ante contracting on the sharing of the net profit, $\alpha\mu_s - C(p)$, of each downstream patent would yield the same patent investment outcome for the downstream innovating firm as that in Proposition 1.

A.3 Patent Holdup (in the Absence of Efficient Ex-Ante Contracting)

Next, we assume that efficient ex-ante contracting is not possible, and consequently ex-post negotiation always take place between a downstream innovating firm and its upstream firms holding complementary patents. Suppose that commercialization of a patent p requires consent from the owners of upstream patents (p_u, $u = 1, 2, ...N_p$). These upstream patents allow their owners to extract part of patent p's value (through, e.g., license fees) so that the innovating firm's expected patent value decreases. We denote the share of the patent value lost to the owner of an upstream patent p_u by $L_s(p, p_u)$ and the aggregate value loss to all of the owners of the upstream patents by

$$L_s(p) = \sum_{u=1}^{N_p} L_s(p, p_u). \tag{A.9}$$

The share $L_s(p) \in [0, 1]$ and its component $L_s(p, p_u)$ depend on the "toughness" of ex-post bargaining by the owner of the upstream patent p_u. In the ideal case in which the shareholders of firm s coincide with those of the firms owning patents p_u, $u = 1, 2, ...N_p$, no rent extraction should take place and therefore $L_s(p) = L_s(p, p_u) = 0$. By contrast, maximal rent extraction occurs when there is no overlap in institutional ownership between the downstream innovating firm and the upstream firms. Again, for simplicity we assume that the ex-ante expected share of value loss is identical for all patents p produced by the same firm, with $E[L_s(p)] = \overline{L}_s$.

Besides the direct value loss due to rent extraction, the holdup situation might also reduce the total value prospect of each individual patent itself. For example, patent litigation may retard the commercial adoption of a patent and jeopardize its long-run success. We assume that the expected number of citations diminishes according to

$$E\left[cites_s(p)\right] = \mu_s[1 - \overline{L}_s]^{\gamma}, \tag{A.10}$$

where γ denotes the elasticity of expected patent success (measured by future citation count) to the retained value share, $1 - \overline{L}_s$, with $\gamma \geq 0$. In the special case $\gamma = 0$, patent holdup does not compromise the overall long-term success of each patent and instead amounts to only a simple redistribution of future patent rents. The expected commercial value from patent p follows as

$$E[V_s(p)] = \alpha[1 - \overline{L}_s]\ E\left[cites_s(p)\right] = \alpha\mu_s[1 - \overline{L}_s]^{1+\gamma}. \tag{A.11}$$

The optimal investment policy in the holdup case requires maximization of the expected present value function

$$\max_{\overline{p}_L} \Pi_s = \int_0^{\overline{p}_L} \left[\alpha\mu_s[1 - \overline{L}_s]^{1+\gamma} - C(p)\right] dp, \tag{A.12}$$

where the optimal patent range $[0, \overline{p}_L]$ has the upper limit

$$\overline{p}_L = \left(\frac{\alpha\mu_s}{c}[1-\overline{L}_s]^{1+\gamma}\right)^{\frac{1}{b}}. \tag{A.13}$$

In the derivation above, we assume, for simplicity, the patent development costs $C(p)$ have been sunk for the downstream innovating firm at the time of ex-post negotiations with upstream firms holding complementary patents.

Proposition 3: Patent Production in the Patent Holdup Case

A firm accounting for an expected value loss $\overline{L}_s$ per patent optimally invests in the production of patents on the line interval $[0, \overline{p}_L]$. Given a patent-level ex-ante expected citation count $E[cites_s(p)] = \mu_s[1-\overline{L}_s]^\gamma$, a convex cost function $C(p) = cp^b$, and an ex-ante expected value loss $\overline{L}_s = E[L_s(p)]$ for each patent due to patent holdup, we find for

(i) the (log) extensive margin of patent production

$$ln[\overline{p}_L] = \frac{1}{b}ln\frac{\alpha}{c} + \frac{1}{b}ln(\mu_s) + \frac{1+\gamma}{b}ln[1-\overline{L}_s] \tag{A.14}$$

(ii) the (log) extensive margin of patent production

$$ln[E(cites_s)] = ln(\mu_s) + \gamma ln[1-\overline{L}_s] \tag{A.15}$$

(iii) the firm-level (log) citation count

$$ln[E(CITES_s)] = \frac{1}{b}ln\frac{\alpha}{c} + \frac{b+1}{b}ln(\mu_s) + \frac{1+\gamma+b\gamma}{b}ln[1-\overline{L}_s], \tag{A.16}$$

(iv) the (log) R&D expenditure

$$ln[R\&D\ Exp] = ln\frac{c}{1+b} + \frac{b+1}{b}ln\frac{\alpha\mu_s}{c} + (1+\gamma)\frac{b+1}{b}ln[1-\overline{L}_s]. \tag{A.17}$$

Eqs. (A.14)–(A.17) are exactly the same as Eqs. (A.4)–(A.7) except for the last term. The last term in Eqs. (A.14)–(A.17) features the same (log) loss term $ln[1-\overline{L}_s] < 0$ and captures how the holdup problem reduces, respectively, the extensive margin, intensive margin, overall patent success, and R&D expenditure.

A.4 Patent Holdup and Shareholder Overlap

The model estimation has to define empirical proxies for the patent-specific holdup loss $L_s(p)$ and its unconditional expected value $E[L_s(p)] = \overline{L}_s$. We assume that shareholder overlap influences $\overline{L}_s$ through two channels: First, a *transfer internalization channel* implies that the management of the downstream firm will account for the portion of the transfer payments received by the overlapping shareholders (but not the portion paid to the upstream firms' other shareholders) in its value maximization. Second, a *transfer reduction channel* suggests that if the rent extraction by upstream firms involves frictions that generate costs for overlapping shareholders without a commensurate benefit, overlapping investors would exercise their influence over the upstream firms in favor of swift conflict resolution and therefore reduce the overall patent transfer payments by the downstream firm. Both channels imply that $\overline{L}_s$ should decrease with *shareholder overlap*.

We can formalize the role of shareholder overlap as follows: Let $O(p)$ be an ownership function that assigns a patent p to a (single) firm owner at time t. The *pairwise (institutional) shareholder overlap* between the downstream patent p and an upstream patent p_u can be defined as

$$PSOL(p, p_u) = \sum_i \min[w_{i,O(p)}, w_{i,O(p_u)}], \tag{A.18}$$

where $w_{i,O(p)}$ and $w_{i,O(p_u)}$ are the ownership share (relative to the total institutional ownership of the respective firm) of institutional investor i in, respectively, firms $O(p)$ and $O(p_u)$ at time t. Without loss of clarity, we denote firm $O(p)$ by subscript s in all subsequent discussions. We assume the following reduced form for the distributive value loss function $L_s(p, p_u)$, with the share of patent p's value loss to its upstream patent p_u decreasing in their *pairwise shareholder overlap*:

$$L_s(p, p_u) = \delta w(p_u) \left[1 - PSOL(p, p_u)\right], \tag{A.19}$$

where the weight function $w(p_u)$ measures the importance of the upstream patent p_u relative to all other upstream cited patents of the follow-up patent p. Presumably, the more important the upstream patent p_u is, the more bargaining power its owner has in terms of rent extraction. The parameter $\delta \in [0, 1]$ denotes the degree to which separate asset ownership translates into patent revenue sharing; a larger value for δ implies more rent redistribution due to ownership separation. The total redistributed rents to the N_p upstream patent holders aggregate to a redistributive loss for patent p, given by

$$\begin{aligned} L_s(p) &= \sum_{u=1}^{N_p} \delta w(p_u) \left[1 - PSOL(p, p_u)\right] \\ &= \delta \left[1 - \sum_{u=1}^{N_p} w(p_u) PSOL(p, p_u)\right]. \end{aligned} \tag{A.20}$$

We can define *patent-level shareholder overlap* as the weighted average *pairwise shareholder overlap* over all N_p upstream patents of patent p.

$$sol_p = \sum_{u=1}^{N_p} w(p_u) PSOL(p, p_u). \tag{A.21}$$

For the N_s patents filed by firm s at year t, we can approximate the average holdup loss as

$$\begin{aligned}\overline{L}_s &= \sum_{p=1}^{N_s} w(p) L_s(p) \\ &= \delta \left[1 - \sum_{p=1}^{N_s} \sum_{u=1}^{N_p} w(p) w(p_u) PSOL(p, p_u) \right],\end{aligned}$$

where the weight $w(p)$ denotes the relative importance of patent p. Presumably, any percentage value loss from a more important patent should translate into a higher absolute value loss for the firm. The *firm-level shareholder overlap* can be defined as

$$SOL_s = \sum_{p=1}^{N_s} \sum_{u=1}^{N_p} w(p) w(p_u) PSOL(p, p_u), \tag{A.22}$$

which captures shareholder commonality between firm s and all other firms owning the upstream complementary patents. The holdup loss term in Proposition 3 can be approximated by

$$ln(1 - \overline{L}_s) \simeq -\overline{L}_s = -\delta[1 - SOL_s], \tag{A.23}$$

and substitution of Eq. (A.23) into Eqs. (A.14)–(A.17) makes the model directly testable. The expression δSOL_s captures the holdup attenuation through firm-level shareholder overlap relative to a total (non-attenuated) holdup effect embodied by $-\delta$.

Appendix B. Variable Definitions

Variable	Description
$N_{s,t}$	Number of patents filed by firm s in year t. Only those patents that are ultimately granted are included in our sample. [Source: Kogan et al., 2017]
$cites_{p,t}$	Total future citation count for patent p, which is filed in year t and subsequently granted by the United States Patent and Trademark Office (USPTO). All self-citations are excluded. Because we only observe citations up to the end of 2010, we correct for this truncation bias using the estimated citation-lag distribution suggested by Hall, Jaffe, and Trajtenberg (2001). [Source: Kogan et al., 2017; Hall et al., 2001]
$CITES_{s,t}$	Total future citation count for the cohort of patents filed by firm s in year t. Only those patents that are subsequently granted by USPTO are included in our sample. [Source: Kogan et al., 2017; Hall et al., 2001]
$\overline{cites}_{s,t}$	Average future citation count per patent for the cohort of patents filed by firm s in year t. [Source: Kogan et al., 2017; Hall et al., 2001]
$CITES^{3yr}_{s,t}$	Three-year citations received by the cohort of patents filed by firm s in year t. For each patent, we count citations received during the calendar year of patent grant and the three subsequent years. [Source: Kogan et al., 2017; Hall et al., 2001]
$CITES^{F}_{s,t}$	Total filtered future citation count for the cohort of patents filed by firm s in year t. It removes from $CITES_{s,t}$ citations from the upstream firms cited in the patent filings of the downstream firm s in year t. [Source: Kogan et al., 2017]
$R\&D$ $Exp_{s,t}/Assets_{s,t}$	The ratio of total $R\&D$ expenditure (in \$U.S. million) to total firm assets (in \$U.S. million) for firm s in year t. The Compustat Mnemonic is XRD for the former and AT for the latter. They are measured based on the latest fiscal year-end value as of the end of calendar year t. [Source: Compustat-CRSP merged database]
$sol_{p,t}$	Shareholder overlap for patent p, filed in year t. It is the weighted average of pairwise shareholder overlap $PSOL(p, p_u)$ across all upstream patents (p_u, $u = 1, 2, .., N_p$) cited by patent p, where $PSOL(p, p_u)$ is measured according to Eq. (1). The weight for an upstream patent p_u is the ratio of its future citations to the aggregate future citations of all cited upstream patents. In cases where multiple upstream patents are owned by the same firm, we aggregate the citation count of these patents and treat them as one single patent. [Source: Kogan et al., 2017; Thomson Reuters 13F database]
$SOL_{s,t}$	Shareholder overlap for firm s in year t. It is the weighted average of $sol_{p,t}$ across all patents p filed by firm s in year t. The weight for a patent p is the ratio of its future citations to the aggregate future citations of all patents filed by the firm in the year. [Source: Kogan et al., 2017; Thomson Reuters 13F database]

Variable	Description
$SOL\_Dedicated_{s,t}$	Shareholder overlap of dedicated investors for firm s in year t. It is exactly the same as $SOL_{s,t}$ except that only the overlapping shares of dedicated investors are counted. Dedicated shareholders are the one-third of investors with the most concentrated portfolio and least portfolio turnover. Specifically, at the end of each year, we rank all institutional investors by the Herfindahl-Hirschman Index (HHI) of their equity portfolio holdings (in descending order) and the turnover ratio (in ascending order). We label dedicated investors as those in the top tercile of the combined rank of the two ranks. The HHI is calculated as the sum of squares of each individual stock's weight in the investor's overall equity portfolio. The turnover ratio for investor i in year t is calculated based on Gaspar, Massa, and Matos (2005) as $\frac{\sum_{j\in\Omega}\lvert Q_{j,i,t}P_{j,t}-Q_{j,i,t-1}P_{j,t-1}-Q_{j,i,t-1}\Delta P_{j,t}\rvert}{\frac{1}{2}\sum_{j\in\Omega}(Q_{j,i,t}P_{j,t}+Q_{j,i,t-1}P_{j,t-1})}$ where $Q_{j,i,t}$ is the number of shares of stock j held by investor i at the end of year t, $P_{j,t}$ is the price of stock j at the end of year t, and Ω is the pool of all stocks held by the investor in the year. [Source: Kogan et al., 2017; CRSP and Thomson Reuters 13F databases]
$SOL\_Intermediate_{s,t}$	Shareholder overlap of intermediate investors for firm s in year t. The overlapping shares are counted only for intermediate investors, who are the middle one-third of shareholders based on the combined rank of the HHI of their equity portfolio holdings (in descending order) and the turnover ratio (in ascending order). [Source: Kogan et al., 2017; CRSP and Thomson Reuters 13F databases]
$SOL\_Transient_{s,t}$	Shareholder overlap of transient investors for firm s in year t. The overlapping shares are counted only for transient investors, who are the bottom one-third of shareholders based on the combined rank of the HHI of their equity portfolio holdings (in descending order) and the turnover ratio (in ascending order). [Source: Kogan et al., 2017; CRSP and Thomson Reuters 13F databases]
$SOL\_Rank_{s,t}$	(Non-parametric) Rank-measure-based shareholder overlap for firm s in year t. We define $rank(p)$ as patent p's rank in future citation count among all patents filed in the given year under the same technology class as defined by USPTO. We then replace the log citation counts $ln[1+cites(p)]$ and $ln[1+cites(p_u)]$ in Eq. (4) with $rank(p)$ and $rank(p_u)$, respectively, to obtain a new shareholder overlap measure SOL_rank. [Source: Kogan et al., 2017]
$SOL^{3yr}_{s,t}$	Three-year cites-weighted shareholder overlap for firm s in year t. It is defined in the same way as $SOL_{s,t}$ except that the log citation count $ln[1+cites(p)]$ in Eq. (4) is replaced with $ln[1+cites^{3yr}(p)]$, in which $cites^{3yr}(p)$ is the future citations received by patent p during the calendar year of patent grant and the three subsequent years. [Source: Kogan et al., 2017]

Variable	Description
$SOL\_Equal_{s,t}$	Equally-weighed shareholder overlap for firm s in year t. It is the same as $SOL_{s,t}$ except that we use equal weights for each patent in the calculation of shareholder overlap. [Source: Kogan et al., 2017; Thomson Reuters 13F database]
$SOL\_Placebo1_{s,t}$	First placebo shareholder overlap measure for firm s in year t. It is constructed in the same way as $SOL_{s,t}$ except that we replace every cited upstream firm with a *similar* firm that is *not* cited by the downstream firm s in the given patent application year t. A placebo firm is chosen based on the criteria that it must have the same four-digit SIC code as the true upstream firm and that it has the shortest Euclidean distance from the upstream firm in terms of (log) firm asset size and (log) number of patents filed in the past five years. Specifically, the Euclidean distance between a true upstream firm u and a placebo firm x is $\sqrt{(\frac{\ln(Asset_x)}{\ln(Asset)\_mean} - \frac{\ln(Asset_u)}{\ln(Asset)\_mean})^2 + (\frac{\ln(1+M_x)}{\ln(1+M)\_mean} - \frac{\ln(1+M_u)}{\ln(1+M)\_mean})^2}$, where $Asset$ and M denote the total firm assets and the number of patents a firm files in the past three years (from $t-2$ to t), respectively. The suffix $\_mean$ refers to the industry average based on four-digit SIC codes. [Source: Kogan et al., 2017; Compustat-CRSP merged database]
$SOL\_Placebo2_{s,t}$	Second placebo shareholder overlap measure for firm s in year t. It is constructed in the same way as $SOL\_Placebo1_{s,t}$ except that the placebo firms are matched to the true upstream firms based on their technological proximity. Following Bloom, Schankerman, and Van Reenen (2013), we measure technological proximity between a true upstream firm u and a placebo firm x by $\frac{T_u T_x'}{\sqrt{T_u T_u'}\sqrt{T_x T_x'}}$, where $T_u = (T_{u,1}, ..., T_{u,K})$ and $T_x = (T_{x,1}, ..., T_{x,K})$. $T_{u,k}$ denotes the ratio of the number of patents filed by firm u in technological field $k \in [1, K]$ in the past three years to the total number of patents it filed during the same period. $T_{x,k}$ is defined analogously. The chosen placebo firm features the greatest value in the technological proximity measure among all firms not cited by the downstream firm in the given year. [Source: Kogan et al., 2017]
$IO_{s,t}$	Aggregate institutional ownership of firm s in year t. It is the ratio of the number of shares held by institutional investors to the total number of shares outstanding for firm s at the end of year t. [Source: Thomson Reuters 13F and Compustat-CRSP merged databases]
$IO_{s,t}^{SOL}$	Overlapping institutional ownership of firm s in year t. For each patent application year t, we identify all *overlapping shareholders* that hold joint equity stakes in firm s and its upstream patent-owning firms. $IO_{s,t}^{SOL}$ measures the ratio of the total number of shares held by overlapping institutional shareholders to the total number of shares outstanding for firm s at the end of year t. [Source: CRSP and Thomson Reuters 13F]

Variable	Description
$IO^{SOL}_{s,t}$	Overlapping institutional ownership of firm s in year t. For each patent application year t, we identify all *overlapping shareholders* that hold joint equity stakes in firm s and its upstream patent-owning firms. $IO^{SOL}_{s,t}$ measures the ratio of the total number of shares held by overlapping institutional shareholders to the total number of shares outstanding for firm s at the end of year t. [Source: CRSP and Thomson Reuters 13F]
$IO^{NOL}_{s,t}$	Non-overlapping institutional ownership of firm s in year t. For each patent application year t, we identify all *overlapping shareholders* that hold joint equity stakes in firm s and its upstream patent-owning firms. The remaining shareholders of firm s are identified as *non-overlapping shareholders*. $IO^{NOL}_{s,t}$ measures the ratio of the total number of shares held by non-overlapping institutional shareholders to the total number of shares outstanding for firm s at the end of year t. [Source: Thomson Reuters 13F and Compustat-CRSP merged databases]
$MktCap_{s,t}$	Market capitalization value (in \$U.S. K) of firm s at the end of year t. [Source: Compustat-CRSP merged database]
$MktCap^{u}_{p,t}$	Average market capitalization value (in \$U.S. K) of firms owning patent p's upstream patents u at the end of year t. [Source: Kogan et al., 2017; CRSP database]
$R\&D\ Stock_{s,t}$	Cumulative R&D investment (in \$U.S. million) of firm s at the end of year t. Following Hall, Jaffe, and Trajtenberg (2005), we measure $R\&D\ Stock_{s,t}$ as $R\&D\ Expenditure_{s,t} + (1-\delta) \times R\&DStock_{s,t-1}$, where δ represents the private depreciation rate of knowledge and is set to be 0.15. [Source: Compustat-CRSP merged database]
$SIF_{s,t}$	Shareholder innovation focus for firm s in year t. In the first step, we define for each listed firm s' the *firm innovation focus* (FIF) as the ratio of the future citation count of all patents filed by firm s' in year t to the industry average during the same period. In the second step, we account for all institutional investors i in firm s and calculate their respective *investor innovation focus* (IIF) as the value-weighted average *firm innovation focus* for all stocks s' in their respective investment portfolios except for stock s itself at the end of year t. In the third step, the *shareholder innovation focus* (SIF) for firm s is defined as the value-weighted average of investor innovation focus for all shareholders i in firm s at the end of year t, with each investor i being weighted based on their relative investment value in the firm. [Source: Kogan et al., 2017; Compustat-CRSP merged database]
$Sales_{s,t}$	Total amount of sales (in \$U.S. million) for firm s in year t. The variable (Compustat Mnemonic: $SALE$) is based on the latest fiscal year-end value prior to the end of calendar year t. [Source: Compustat-CRSP merged database]

Variable	Description
$K/L_{s,t}$	Capital (in \$U.S. million) to labor (in thousands) ratio for firm s in year t. K and L denote capital (Compustat Mnemonic: $PPEGT$) and labor (Compustat Mnemonic: EMP), respectively. Both are based on the latest fiscal year-end values prior to the end of calendar year t. [Source: Compustat-CRSP merged database]
$WHHI_{s,t}$	Weighted Herfindahl-Hirschman Index of shareholder overlap concentration for firm s in year t. For each patent p filed by firm s in year t, we identify all the overlapping shareholders $i \in I_{p,pu}$ who have a joint equity stake in firm s and the firm owning the upstream patent p_u. We then calculate $hhi_{p,p_u,t}$ as the Herfindahl-Hirschman Index based on the overlapping ownership share of each overlapping shareholder $i \in I_{p,pu}$, with the ownership measured at the end of year t. $WHHI_{s,t}$ is the weighted average of $hhi_{p,p_u,t}$ across all patents p owned by firm s and their respective upstream patents p_u, where the weight for each patent is given by Eq. (4) [Source: Kogan et al., 2017; Thomson Reuters 13F database]
$Private\ Patent\ Share_{s,t}$	Average proportion of private upstream patents for firm s in year t. For each patent p filed by firm s in year t, we calculate the share of privately owned upstream patents. We then average this private patent share across all patents filed by firm s in year t, with the weight of each patent p given by w(p) in Eq. (4). [Source: Kogan et al., 2017]
$SpillTech_{s,t}$	Technology (or knowledge) spillover from other firms for firm s in year t. It is the technological proximity-weighted sum of *R&D Stock* (in \$U.S. million) of all firms in year t except firm s. Technological proximity between firms m and s is defined by $\frac{T_m T_s'}{\sqrt{T_m T_m'}\sqrt{T_s T_s'}}$, where $T_m = (T_{m,1}, ..., T_{m,K})$ and $T_s = (T_{s,1}, ..., T_{s,K})$. $T_{m,k}$ denotes the ratio of the number of patents filed by firm m in technological class $k \in [1, K]$ over the whole sample period to the total number of patents it filed during the same period. $T_{s,k}$ is defined analogously. [Source: Kogan et al., 2017; Compustat-CRSP merged database]
$SpillSIC_{s,t}$	Product market rivalry effect of *R&D* for firm s in year t. It is the product market proximity-weighted sum of *R&D Stock* (in \$U.S. million) of all firms in year t except firm s. Product market proximity between firms m and s is defined by $\frac{X_m X_s'}{\sqrt{X_m X_m'}\sqrt{X_s X_s'}}$, where $X_m = (X_{m,1}, ..., X_{m,Q})$ and $X_s = (X_{s,1}, ..., X_{s,Q})$. $X_{m,q}$ denotes the share of firm m's sales in industry $q \in [1, Q]$ relative to its total sales during the year, averaged over the whole sample period. Industries are defined by four-digit SIC codes. $X_{s,q}$ is defined analogously. [Source: Kogan et al., 2017; Compustat-CRSP merged database]