# Technological Progress and Ownership Structure

Heng Geng*

Victoria University of Wellington

Harald Hau**

University of Geneva and Swiss Finance Institute

Sandy Lai***

University of Hong Kong

November 6, 2016

**Abstract**

Innovation processes under patent protection generate holdup problems if complementary patents are owned by different firms. We show that in line with Hart and Moore (1990), shareholder ownership overlap across firms with patent complementarities helps mitigate such holdup problems and correlates significantly with higher patent investment and more patent success as measured by future citations. The positive innovation effect is strongest for concentrated overlapping ownership and for the cases in which the overlapping shareholders are dedicated investors.

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.

***Faculty of Business and Economics, University of Hong Kong, K.K. Leung Building, Pokfulam Road, Hong Kong. Tel.: (+852) 3917 4180. E-mail: sandy_lai@hku.hk.

# 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, 2016). 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. 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.[1]

The property rights literature (Williamson, 1975, 1985; Grossman and Hart, 1986; Hart and Moore, 1990) argues that joint ownership of complementary assets increases (ex-ante) investment incentives. Applying this insight to the patent process, we conjecture that joint equity ownership (i.e., shareholder overlap) between an innovating (downstream) firm and other (upstream) firms controlling complementary patents can similarly attenuate the holdup problem and contribute to the patent success of the innovating firm. 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 avoid underinvestment in downstream patents.[2] 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

[1] Recent economic research (e.g., Heller and Eisenberg, 1998; Bessen and Maskin, 2009) has documented a negative impact of recent patent proliferation on R&D investment and follow-on innovation. Heller and Eisenberg (1998) in particular point out that such proliferation of upstream patents in biomedical research forces some downstream firms to divert resources to less promising projects with fewer licensing obstacles.

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

investors can contribute to a swift conflict resolution, thereby reducing the overall patent transfer payments and increasing ex-ante investment incentives by the downstream firm.

Anecdotal evidence on legal conflict resolution provides support for the holdup attenuating role of overlapping shareholders. Hansen and Lott (1996) cite Albert J. Wilson, Vice President and Secretary of TIAA-CREF (which was an overlapping shareholder in the firms involved in the conflicts), as stating that his large pension fund was actively involved in applying pressure to ensure that the Pennzoil v. Texaco and Apple v. Microsoft conflicts were resolved, and he claimed that this pressure resulted in Pennzoil and Texaco settling their lawsuit much sooner than they would have done otherwise.

To subject this property-right 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 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 Alcácera, Gittelmanb, and Sampatc (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 a limited scope in manipulating the reference list by the patent filing firms.

Prior research (Ziedonis, 2004; Galasso and Schankerman, 2010; Noel and Schankerman, 2013) suggests that the owners of upstream cited patents are reasonable proxies for the potential licensors of downstream citing patents. Ziedonis (2004), in particular, documents that some patent strategy consulting firms reported screening the list of companies that cited their clients' patents in an effort to identify potential licensees.[3] In fact, in 2005 two U.S. inventors, Stephen K. Boyer and

[3] Ziedonis (2004) specifically names three consulting firms: Mogee Associates, InteCap, and Delphion. Sep-

Alex Miller, were granted a patent (US6879990) for proposing a systematic approach to identify potential licensees from patent citation references for a group of target patents.[4] Following this line of the literature and industry practice, our analysis assumes that the (potential) need for acquiring licenses for complementary patents listed in the patent filing creates potential patent holdup for the downstream firm. Figure 1 shows that firms with citation links are 35 times as likely to engage in patent-related lawsuits against each other as those without any citation links.[5] The evidence lends support to the previous literature that relies on citation links as a proxy for patent complementarity and potential patent holdup. Notwithstanding the imperfect nature of the proxy, it allows us to identify asset complementarity for a large sample of firms, particularly among firms in the forefront of the innovation process.

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 R&D investment, and contributes to the long-run patent success of the innovating firm. We further explore two refinements of this basic hypothesis. Specifically, we examine whether shareholder activism and the concentration of overlapping equity stakes matter for the holdup attenuation effect.

To test these hypotheses, we first construct a new explanatory variable, *firm-level shareholder overlap* ($SOL$), characterizing the percentage of equity ownership for which investors own an equally large equity stake in both the downstream innovating firm and the upstream firms owning the precursory patents. Consider a patent $p$ owned by the downstream firm $O(p)$ that cites a precursory patent $p_u$ owned by the 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 aggregating over all upstream

arately, Ambercite (an intellectual property services company) in a recent internet posting suggested its clients should look for potential patent licensees from the list of owners of forward citation patents (www.ambercite.com 2014).

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

[5] The figure is based on the Audit Analytics Litigation database collected primarily from corporate disclosures to the Securities and Exchange Commission (SEC). Reported are a total of 604 patent lawsuits over the period 2000-2007 for our firm sample. 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.

patents cited in the patent filing of patent $p$, and the *firm-level shareholder overlap* ($SOL$) is obtained by jointly aggregating 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 $t$.[6]

### Main Findings

Consistent with the *holdup attenuation hypothesis* of shareholder overlap, we find that $SOL$ emerges as a statistically and economically significantly positive covariate of patent success, and the effect is more pronounced in the top three R&D-intensive sectors. Overall, an increase in shareholder overlap $SOL$ by one standard deviation enhances patent success in terms of log firm citation (*ln[1+CITES]*) by 11.3% of its standard deviation and lead to approximately 18% more patents to be filed by the downstream innovator. 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 find a much stronger effect of shareholder overlap on patent success when such overlap or joint ownership 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. Besides, the Herfindahl-Hirschman index of (overlapping) shareholder ownership concentration correlates positively with the firm-level patent success beyond the shareholder overlap itself, suggesting 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. We also note that large overlapping shareholders of innovating firms are more likely to simultaneously serve on the boards of both upstream and 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.[7]

[6] See, for example, Aghion, Van Reenen, and Zingales (2013) for a similar definition of firm-level patent success.
[7] We obtain board data from the BoardEx database. The database has limited coverage prior to 2000, and it

We pursue four different strategies to address the potential omitted variables and reverse causality issues in the empirical relation between shareholder overlap and patent success. 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. Any remaining omitted variable effect thus needs to operate on the patent-level success of the downstream firm and simultaneously correlate with the ownership structure of the patent-specific upstream firms (and therefore correlate with $sol$).

Second, to address the remaining endogeneity concern arising from the ownership structure of upstream firms, we instrument the patent-level shareholder overlap $sol$ with the average market capitalization of patent-specific upstream firms. The average size of these patent-specific upstream firms should correlate positively with $sol$ but have little partial effect on the success of the patent (once all other relevant variables are controlled for), 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 even when the latter is instrumented by the average market capitalization of the upstream firms.

Third, to further probe omitted variables operating across firms, we design two placebo tests. We replace the actual shareholder overlap ($SOL$) with a placebo shareholder overlap. The latter is constructed by replacing 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 features no statistically significant effect on holdup mitigation and patent success.

Fourth, we address the issue of reverse causality using the two placebo measures of shareholder

covers about 66% of CRSP stocks in 2000 and 74% in 2007 (Engelburg, 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.

overlap constructed earlier. If investors anticipate a positive effect of shareholder overlap on potential future patent success and strategically acquire overlapping ownership shares prior to the public disclosure of 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 may not be surprising because patent developments are generally kept secret and trading on insider information is sanctioned by law.

One of the proximate causes of patent success is R&D investment. The property rights literature particularly emphasizes the negative effect of holdup on firm investment. Empirically, we find an economically strong positive relation between shareholder overlap and R&D expenditure, again pointing to the holdup attenuation of shareholder overlap on R&D investment. This result also serves as a robustness check for the earlier results based on patent citations.

We conduct an additional falsification test based on the relation between shareholder overlap and R&D expenditure. If the governance influence of overlapping shareholders is the main cause of higher R&D expenditure, then non-overlapping shareholders should have an opposing interest. From their perspective, any internalization of rent transfers to upstream patent owners implies R&D overinvestment because unlike overlapping shareholders, non-overlapping shareholders are not entitled to any transfer payment made to upstream firms. Accordingly, we can test whether a higher share of non-overlapping institutional investors $IO^{NOL}$ is negatively correlated with R&D expenditure. We find that shareholder overlap ($SOL$) and non-overlapping institutional ownership ($IO^{NOL}$) indeed feature opposite correlations with R&D investment. This finding again is consistent with the holdup attenuation hypothesis of overlapping shareholders. Importantly, it also highlights the fact that institutional investors can have opposing interests with respect to firm policy and that aggregate institutional ownership itself may not be a very meaningful variable in characterizing agency conflicts with respect to patent investment.

To the best of our knowledge, the role of stock market ownership structure in mitigating holdup problems in patent processes has not been subject to any systematic analysis. Ex-ante contracting about access to auxiliary patents is difficult before the feasibility and commercial potential of a new patent are established. Holdup expectations should reduce ex-ante investment incentives unless overlapping shareholders internalize such rent extraction through simultaneous ownership in upstream and downstream firms. Costly patent rent extraction (for which efficiency losses occur) might also be reduced through the power of overlapping shareholders vis-à-vis upstream firms, justifying higher ex-ante R&D investments. The well-established positive link of firm value with R&D efficiency and patent and citation count (Hall, Jaffe, and Trajtenberg, 2005) also points to a positive relation between shareholder overlap and overall firm value.

In the following section, we survey the related literature. In Section 3, we discuss the data, variable construction, and summary statistics. In Appendix A, we present a simple model of patent holdup in the spirit of Hart and Moore (1990) to motivate our empirical analysis. In Section 4, we present the empirical evidence, and we conclude in Section 5.

# 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 requires a level of disaggregation typically not available for investment data. Any firm-level analysis is clouded by the fact that a firm can shift investments to other projects for which holdup problems are less severe. Third, investments may involve intangible resources (such as managerial attention), which pose additional measurement problems. In this study, explicit citation of precursory patents in patent filings provides a unique opportunity to identify the potential holdup problems in a firm's patent success. We infer the (latent) project underinvestment indirectly from diminished patent success.[8] Future patent citations provide a sufficiently precise proxy for patent success at the firm and patent levels to allow for a comprehensive study of holdup in patent process. (See, e.g., Harhoff, Narin, Scherer, and Vopel, 1999, for a direct positive link between the future citation count and the economic value of a patent.)

Can firms avoid patent conflicts? Previous studies on patent holdup (e.g., Shapiro, 2001; Ziedonis, 2004; Hall and Ziedonis, 2007) find that licensing agreements are commonly used in

[8] One of the key assumptions in Hart and Moore (1990) is that investment has to be specific to an asset or product such that the realization of the investment cannot be used for other purposes. In the setting of corporate innovation, the applied research is usually done with a product in mind and is therefore asset-specific in this sense.

practice—yet these might typically concern the ex-post rent allocation. Licensing agreements can involve substantial royalty fees and their negotiation might not be a frictionless process. A firm can also try to invent around a patented technology, but given the cumulative and sequential nature of technological development, such a strategy is not always possible. There is also evidence that firms seek outright ownership integration via mergers to resolve patent disputes. But firm mergers involve high transaction costs and might be challenged in court for anti-competitive reasons (Creighton and Sher, 2009). Our findings in this study suggest that in liquid equity markets, partial ownership integration via ownership overlap might be achieved at lower costs, and such partial integration may already exist if large institutional shareholders happen to hold shares in both firms concerned.

Recent empirical work on the determinants of patent success focuses on the role of institutional shareholders. Aghion, Van Reenen, and Zingales (2013) and Bena, Ferreira, Matos, and Pires (2016) argue that a large share of institutional shareholders is conducive to patent investment as these shareholders tend to pursue a long-run objective. Our evidence shows that it is important to decompose institutional ownership into an overlapping component and a non-overlapping component as the latter correlates negatively with patent investment. Generally, institutional investors may have opposing interests in R&D investment depending on their ownership stakes in those upstream firms that benefit from licensing rents.

Our work is also related to a nascent literature on the coordination role of common shareholders in corporate policies. Hansen and Lott (1996) provide evidence that overlapping shareholders internalize conflicts between firms and pursue a strategy that maximizes their overall portfolio value. He and Huang (2014) find that firms with common shareholders are more likely to coordinate their product market strategies and experience higher market share growth. Azar, Schmalz, and Tecu (2015) show that shareholder overlap induced by increasing institutional ownership softens product market competition. We note that our evidence on the positive relation between patent success and shareholder overlap is robust to the control of product market competition (as reported in the robustness section, 4.8).

Recent empirical work has also highlighted the complementarity between equity market development and the degree of patent innovation (Hsu, Tian, and Xu, 2014). Insofar as equity market development allows for a 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, Papanikolaou, Seru, and Stoffman (2016). The data set provides annual patent and citation information for patents granted over the period 1926–2010.[9] 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 the United States Patent and Trademark Office (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 metrics by summing up the patent-level metrics by patent filing year instead of grant year, as 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 cites 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 adjust carefully for the two truncation problems commonly associated with patent data. First, the patent data set only includes those patents that are eventually granted, so many patent

[9] 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.

applications filed in the last two years of our patent data set (i.e., years 2009 and 2010) were still not granted by the end of 2010 and therefore not included in the data. To mitigate this patent truncation bias, we use only patent applications up to 2007 in our empirical analysis. Second, patents tend to receive citations over a long period of time, but in our data set we observe citations only up to 2010. We correct for the truncation bias in citation count based on the shape of the citation-lag distribution suggested by Hall, Jaffe, and Trajtenberg (2001, 2005). Furthermore, because expired patents should not create any holdup problems, we ignore those upstream cited patents that have expired by the time the shareholder overlap measure is constructed.[10]

## 3.2 Ownership Data

We obtain the ownership data from the Thomson Reuters 13F database. The SEC requires all institutional organizations, companies, universities, and so on that exercise discretionary management of investment portfolios over $100 million in equity assets to report those 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 only use ownership data 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 30 patents per year.

[10] 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 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 $p_u$ (listed in the patent filing) 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 weights reflecting the relative importance of any patents $p$ and $p_u$. In the context of our model (presented in Appendix A), a higher weight assigned to a more important upstream patent reflects the fact that its owner is likely to have a stronger bargaining power in terms of future rent extraction. A higher weight for a more important downstream patent reflects 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 the 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, 4.8, 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 uses equal weights. The results are qualitatively similar.

A limitation of our analysis is that due to data constraint we can measure ownership only for publicly listed firms, but not for 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 from 1991 to 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 cap, 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. We develop a simple model of holdup attenuation through shareholder overlap in the spirit of Hart and Moore's (1990) property rights theory (Appendix A). 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 $ln[1{+}CITES]$ instead of $ln[CITES]$ because some firms register patents that have never been cited throughout the sample period.[11] The baseline regression linking patent success to shareholder overlap is

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

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, $\beta_0$ represents the overall constant for all observations, $\beta_1$ is the coefficient for $SOL$, $\beta_2$ denotes the vector of coefficients for control variables, $\epsilon_s$ and $\mu_t$ 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

[11] As discussed in the robustness section, 4.8, 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 firms in the top three R&D-intensive sectors (pharmaceuticals, computer hardware, and telecommunications equipment).[12] Columns 1 and 3 control for year fixed effects and industry fixed effects based on four-digit SIC codes, whereas Columns 2 and 4 control for year and firm fixed effects. We report in parentheses robust standard errors with two-way clustering at the firm-year level.

The baseline regression shows that *shareholder overlap SOL* represents a statistically and economically significant explanatory variable. The point estimate of 3.692 in Column 1 implies that an increase in shareholder overlap by one standard deviation (or 0.063) increases patent success in terms of log firm 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 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.

The inclusion of firm fixed effects in Column 2 limits the explanatory power of $SOL$ to the intertemporal variation of patent success within a firm. The point estimate for $SOL$ drops to 1.586 but still remains highly significant at the 1% level. This reduced economic significance level is explained by the double inclusion of the firm fixed effects and the five firm-level controls, which together absorb much of the cross-sectional variation in patent success. As expected, *Private Patent Share* has the same sign as $SOL$ because it proxies for the possible underestimation

[12]We identify the three R&D-intensive sectors following the approach suggested by Bloom, Schankerman, and Van Reenan (2013). Specifically, they include firms in the following sectors: Pharmaceuticals (SIC codes 2834 and 2835), computer hardware (SIC codes 3570, 3571, 3572, 3575, 3576, and 3577), and telecommunications equipment (SIC codes 3661, 3663, and 3669).

of shareholder overlap due to the unobserved overlap originating from private investors. It is statistically significant in Column 1 but insignificant in Column 2, suggesting that firm fixed effects capture much of the firm-specific underestimation of $SOL$.

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 estimate for $SOL$ in Column 4 nearly doubles that in Column 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

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

where $\theta_1 > 0$ implies that patent holdup reduces the average success of a firm's patents. A positive value of $\theta_1$ 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 $\gamma$ measures the efficiency loss of patent holdup, whereas $\delta$ 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. Column 1 excludes firm fixed effects, so both cross- and within-firm variation in shareholder overlap are reflected in the point estimate of 0.584, implying an increase in shareholder overlap by one standard deviation (or 0.063) corresponds to an increase in the average citation count per patent by about 3.2% of its standard deviation. Inclusion of firm fixed effects in Column 2 restricts the identification of the shareholder overlap effect to intertemporal firm variation. The insignificant coefficient for $SOL$ suggests that much of the attenuation effect for the intensive margin coming from the cross-sectional variation is now absorbed by the combination of firm-level controls and firm fixed effects.

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 $\psi_1$ captures the effect of holdup mitigation through shareholder overlap on the number of successful patent filings. The point estimate of 2.923 for $\psi_1$, 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. Moreover, the coefficient retains its statistical significance in the specification with firm fixed effects, reported in Column 4.

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 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

An important condition for finding any holdup attenuation under shareholder overlap is that shareholders seek to influence the corporate decision process. We identify *potential influence* of shareholders on the patent holdup problem along two dimensions: We hypothesize that shareholder overlap is inconsequential if (i) the overlapping shareholders are non-dedicated investors, who do not seek to influence the corporate decision process, or (ii) the overlapping ownership shares are so fragmented that coordinated actions are difficult to organize. We subject these two aspects of shareholder influence to further testing.

We first categorize institutional investors into (i) dedicated investors, (ii) intermediate investors, and (iii) transient investors based on a combination of portfolio diversification (proxied by the Herfindahl-Hirschman Index, HHI) and portfolio turnover (proxied by the churn ratio defined in Gaspar, Massa, and Matos, 2005).[13] 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 churn ratio rank in percentile. 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 rest of the investors, in the middle tercile, are labeled intermediate investors. The distribution of investor types along the two sorting criteria is shown in Figure 3.

Then, we decompose the shareholder overlap of each firm-year according to the three investor types:

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

We hypothesize that shareholder overlap from dedicated investors attenuates holdup problems more effectively than shareholder overlap from the other investor groups. The regression result, reported in Column 2 of Table 4, confirms this prediction. The coefficient for $SOL\_Dedicated$ is 25.151, more than six 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.

The second aspect of shareholder influence concerns the potential coordination problem among overlapping shareholders. If the downstream innovating firm and the upstream cited firm are jointly owned by a few relatively large shareholders, coordinated action might be easier to organize, and shareholders could have stronger incentives to resolve a potential holdup. To test this hypothesis, let's consider a downstream patent $p$ filed by firm $s$ in year $t$ and a related upstream patent $p_u$ owned by firm $u$. Let $i \in I_{p,p_u}$ 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. We can define a Herfindahl-Hirschman Index ($hhi$) of shareholder overlap based on the overlapping ownership shares $\varpi_i = \min[w_{i,s}, w_{i,u}]$ of all overlapping shareholders $i \in I_{p,p_u}$. We can further aggregate this shareholder overlap concentration index over all downstream patents $p$ filed by firm $s$ in year $t$ and over their respective upstream patents $p_u$ to obtain a weighted Herfindahl-Hirschman Index ($WHHI$) of ownership concentration of overlapping shareholders,

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

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 before) the relative importance weights for patents $p$ and $p_u$, respectively, and the ownership shares are measured at 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 statistically significantly positive, 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 2.449 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 59.7% relative to its mean (= [0.071 × 2.449] / [4.698 × 0.062]). This suggests 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 present the patent-level regression specification by first including the separate firm and year fixed effects and then the 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. 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. The patent-specific holdup attenuation is captured by patent-level shareholder overlap $sol_{p,t-1}$ in the regression specification

$$ln[1 + cites_{p,t}] = \beta_0 + \beta_1 sol_{p,t-1} + \beta_2 Controls_{p,t-1} + \epsilon_{s,t} + \theta_{f,t} + \eta_{p,t}, \tag{10}$$

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 $\epsilon_{s,t}$ denotes the interacted firm and year fixed effects, $\theta_{f,t}$ represents the interacted technology field and year fixed effects, and $\eta_{p,t}$ is the residual term.[14]

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) influences at the level of the downstream firm. Therefore, any omitted variable effect in Eq. (10) needs to operate on the patent-level success of the downstream firm (i.e., $ln[1+\text{cites}_{p,t}]$) and also simultaneously correlate with the ownership structure of the patent-specific upstream firms (and hence correlate with $sol_{p,t-1}$). To address the endogeneity concern with respect to the ownership of the upstream firms, we conduct a two-stage least square ($2SLS$) regression by instrumenting $sol$ with the average market capitalization of the patent-specific upstream firms. The average size of these 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 (i.e., has no correlation with the residual term $\eta_{p,t}$).

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 bias toward those 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. It is also noted that the patent-level citation success $cites_{p,t}$ can capture only the intensive margin, but not the extensive margin of patent success.

Even with the above constraint, we still find statistically significant point estimates of $sol$ in Columns 1–2 of Table 5, providing strong evidence that shareholder overlap features a holdup attenuation effect even at the patent level. Columns 3 and 4 report the first- and second-stage results of the $2SLS$ regression. The $2SLS$ specification yields a very similar point estimate for $sol$ (0.283) to that (0.272) obtained from $OLS$. Overall, the result is consistent with the argument that patent success within a firm is also correlated with the patent-specific shareholder overlap $sol$, which differentiates between patents within the same firm-year.

## 4.5 Two Placebo Tests

We propose two different placebo tests to examine the extent to which the relation between *firm-level* patent success and shareholder overlap is driven by unobservable factors. In each placebo test,

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

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 the shortest Euclidean distance to the 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).[15] The two placebo measures, $SOL\_Placebo1$ and $SOL\_Placebo2$, have a slightly lower mean (0.050 and 0.047, respectively) than the true measure of shareholder overlap $SOL$ (mean = 0.062).

By design, placebo firms have never been cited by the downstream firm in the given year. Therefore, any *placebo* shareholder overlap effect must be driven by factors unrelated to the citation link between the upstream and downstream firms. Columns 4–5 of Table 4 show that the two placebo measures of shareholder overlap ($SOL\_Placebo1$ and $SOL\_Placebo2$) do not feature any statistically significant correlation with patent success. Therefore, the positive correlation between shareholder overlap and firm-level patent success (shown in Column 1 of Table 4) is contingent on picking exactly those upstream firms that are cited in the patent filings of the downstream firm, for the construction of $SOL$.

In conclusion, our findings on the holdup attenuation of true $SOL$ (reported in Column 1 of Table 4) cannot be driven by any unobservable factors influencing both firm-level patent success and shareholder overlap. Any such omitted variables should similarly lead to a positive relation between placebo shareholder overlap and patent success, inconsistent with the evidence from the two placebo tests (reported in Columns 4–5).

### 4.6 R&D Expenditure and Non-Overlapping Institutional Ownership

The proximate cause of patent success is patent investment. We therefore examine the direct effect of shareholder overlap on the investment behavior of a firm using the linear panel regression

$$ln[1 + R\&D\ Exp_{s,t}] = \kappa_0 + \kappa_1 SOL_{s,t-1} + \kappa_2 Controls_{s,t-1} + \epsilon_s + \mu_t + \eta_{s,t}, \quad (11)$$

[15] We provide detailed definitions of the variables in Appendix B.

where we include the same control variables as in the previous regressions with the exception of $ln(1+R\&D\ Stock)$, which is excluded because it summarizes past R&D expenditure. (The model in Appendix A implies $\kappa_1 = (1+\gamma)(1+1/b)\delta > 0$.) Because the data on R&D expenditure are sourced from the Compustat database compiled independently from the patent data, using it as an alternative dependent variable also alleviates concerns about any possible measurement errors or biases in the patent data.

Table 6 reports the regression results. The effect of shareholder overlap is statistically and economically significant in the specifications both without firm fixed effects (Column 1) and with firm fixed effects (Column 2). The estimates in Column 1 suggest that an increase in shareholder overlap by one standard deviation (or 0.063) increases R&D expenditure by 11.9% of its standard deviation. The holdup attenuation effect of shareholder overlap on R&D investment is therefore economically important.

If the higher R&D expenditure for firms with shareholder overlap is indeed driven by the governance influence of overlapping shareholders, the ensuing agency conflict implies that non-overlapping institutional shareholders in the downstream firm might oppose patent rent internalization by the overlapping shareholders because from their perspective such rent internalization leads to R&D overinvestment that needs to be curtailed.[16]

Columns 3 and 4 of Table 6 extend the specifications in Eq. (11) to include the non-overlapping institutional ownership $IO^{NOL}$. Their ownership share is obtained by subtracting the holdings of overlapping shareholders from the aggregate institutional ownership. In accordance with our hypothesis, the coefficient for $IO^{NOL}$ has the predicted negative sign and is statistically highly significant. The evidence suggests that non-overlapping institutional shareholders constrain the holdup internalization efforts of overlapping shareholders.

## 4.7 Reverse Causality?

Asset ownership structure could dynamically adjust to patent holdup and evolve toward an efficient combination of complementary assets. Under private information about future patent holdup, investors have an incentive to achieve this joint ownership of complementary patents through shareholder overlap (between upstream and downstream firms)—thus internalizing the

[16] This can be easily inferred from a firm's value maximization problem in Eq. (23) by comparing the case with patent rent internalization by overlapping shareholders and the case without such rent internalization.

patent holdup problem—and subsequently benefiting from the value gains from the success of the patent projects. In this case, (potential) future patent success (proxied by $CITES_{s,t}$) could cause shareholder overlap $SOL_{s,t-1}$ to increase, resulting in a reverse causality problem in our baseline regression specification.

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 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 $k$) for all patent filing years follows

$$\overline{SOL}(k)=\begin{cases}\frac{1}{17-|k|}\sum_{t=1991+|k|}^{2007}SOL(t,k), & if \ \ -5\leq k\leq -1\\ \frac{1}{17-|k|}\sum_{t=1991}^{2007-|k|}SOL(t,k), & if \ \ 0\leq k\leq 1\end{cases} \tag{12}$$

and is plotted in Figure 4.[17] 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 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. The vertical line segment marks two standard deviations around the mean value for each measure. We find no evidence that the shareholder overlap $\overline{SOL}(k)$ endogenously reacts in anticipation of patent rents from future patent filing. During the five-year run-up to the patent filing year, $\overline{SOL}(k)$ actually does not 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 should be extremely scarce. Second, legal restrictions on insider trading limit the scope for stock trading on private information.

[17] 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 $k<=0$.

## 4.8 Robustness Issues

We examine several alternative hypotheses that may account for the evidence reported in the preceding sections. 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 the management to focus on the long-term return on investment. We therefore include *institutional ownership* ($IO_{s,t-1}$) as an 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, 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. We therefore 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. As expected, Column 3 shows that the general innovation focus of a firm's shareholders fosters the patent success of the firm, but the $SOL$ effect still remains strong.

Third, 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 the aforementioned 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.

We also conduct a variety of robustness tests concerning the measurement of shareholder overlap and patent citations. 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 $SOL_{s,t-2}$ is only slightly smaller than that of $SOL_{s,t-1}$ reported in Column 1. 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, our baseline measure of $CITES$ follows Hall, Jaffe, and Trajtenberg (2001) in scaling the raw future citation count of each patent by a specific factor (see Table 5 of Hall, Jaffe, and Trajtenberg, 2001) that increases in the time span until the terminal year of our sample. We reproduce our results using an alternative aggregation proposed by Lerner, Sorensen, and Stromberg (2011). The modified shareholder overlap variable is denoted by $SOL\_rel$. As another 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 8–10, for the holdup attenuation effect of shareholder overlap.

Third, 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.

Fourth, as patent citation count is often perceived as a value signal (Harhoff, Narin, Scherer, and Vopel, 1999; Hall, Jaffe, and Trajtenberg, 2005), overlapping institutional shareholders may promote cross-citations among firms in which they also have 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 3.498 for $SOL$ is quantitatively similar to that of 3.692 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 success of patents often depends on access to complementary patents not under the direct control of the innovating firm. From a property rights perspective, the "extended boundary" of the innovating firm includes such complementary patents if both the downstream innovator and the upstream firms owning these complementary patents are linked by joint shareholder ownership.

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 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 upstream firms (holding the precursory patents to the downstream patent) without a qualitative change 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.

# 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 $C''(p) > 0$. 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 $cites_s(p)$.[18] Hence,

$$V_s(p) = \alpha \times cites_s(p), \tag{13}$$

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 $\Pi_s$ follows as

$$\Pi_s = \max_{\overline{p}} \int_0^{\overline{p}} [\alpha\mu_s - C(p)]\, dp, \tag{14}$$

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

$$\overline{p} = \left(\frac{\alpha\mu_s}{c}\right)^{\frac{1}{b}} \tag{15}$$

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

$$ln[\overline{p}] = \frac{1}{b} ln\frac{\alpha}{c} + \frac{1}{b} ln(\mu_s) \tag{16}$$

[18] Using survey data, Harhoff, Narin, Scherer, and Vopel (1999) provide direct evidence that future citation count is positively related to the economic value of a patent.

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

$$ln[E(cites_s)] = ln(\mu_s) \tag{17}$$

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

$$ln[E(CITES_s)] = ln\int_0^{\overline{p}} E[cites_s(p)]dp = \frac{1}{b}ln\frac{\alpha}{c} + \frac{b+1}{b}ln(\mu_s), \tag{18}$$

(iv) the (log) R&D expenditure

$$ln[R\&D\ Exp] = ln\int_0^{\overline{p}} cp^b dp = ln\frac{c}{1+b} + \frac{b+1}{b}ln\frac{\alpha\mu_s}{c}. \tag{19}$$

The firm-level (log) citation count in Eq. (18) is equal to the sum of the (log) extensive margin in Eq. (16) and the (log) intensive margin in Eq. (17). Empirically, we can approximate the intensive margin by the average citation count $\overline{cites_s}$ of a firm's patents.

## A.2 The Patent Holdup Effect

Next, we enrich the model setting to account for holdup problems with respect to the patent value $V_s(p)$. Suppose that commercialization of a patent $p$ requires consent from the owners of upstream patents ($p_u$, $u = 1, 2, ...N_p$).[19] 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

$$L_s(p) = \sum_{u=1}^{N_p} L_s(p, p_u). \tag{20}$$

The share $L_s(p) \in [0, 1]$ and its component $L_s(p, p_u)$ depend on the "toughness" of 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 $p_u$ with $u = 1, 2, ...N_p$, 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

[19] Note that $p_u$ does not include any expired patents because they do not pose any threat to the commercialization of the citing patent.

no overlap in institutional ownership between the downstream innovating firm and the upstream 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, 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

$$E\left[cites_s(p)\right] = \mu_s[1 - \overline{L}_s]^\gamma, \tag{21}$$

where $\gamma$ denotes the elasticity of the 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

$$E[V_s(p)] = \alpha[1 - \overline{L}_s]\ E\left[cites_s(p)\right] = \alpha\mu_s[1 - \overline{L}_s]^{1+\gamma}. \tag{22}$$

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

$$\max_{\overline{p}_L} \Pi_s = \int_0^{\overline{p}_L} \left[\alpha\mu_s[1 - \overline{L}_s]^{1+\gamma} - C(p)\right] dp, \tag{23}$$

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

$$\overline{p}_L = \left(\frac{\alpha\mu_s}{c}[1 - \overline{L}_s]^{1+\gamma}\right)^{\frac{1}{b}}. \tag{24}$$

**Proposition 2: 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 $E[cites_s(p)] = \mu_s[1 - \overline{L}_s]^\gamma$, which is proportional to the expected patent value, 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

$$ln[\overline{p}_L] = \frac{1}{b}ln\frac{\alpha}{c} + \frac{1}{b}ln(\mu_s) + \frac{1+\gamma}{b}ln[1-\overline{L}_s] \tag{25}$$

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

$$ln[E(cites_s)] = ln(\mu_s) + \gamma ln[1-\overline{L}_s] \tag{26}$$

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

$$ln[E(CITES_s)] = \frac{1}{b}ln\frac{\alpha}{c} + \frac{b+1}{b}ln(\mu_s) + \frac{1+\gamma+b\gamma}{b}ln[1-\overline{L}_s], \tag{27}$$

(iv) the (log) R&D expenditure

$$ln[R\&D\ Exp] = ln\frac{c}{1+b} + \frac{b+1}{b}ln\frac{\alpha\mu_s}{c} + (1+\gamma)\frac{b+1}{b}ln[1-\overline{L}_s]. \tag{28}$$

Eqs. (25)–(28) are exactly the same as Eqs. (16)–(19) except for the last term. The last term in Eqs. (25)–(28) 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.3 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 in *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

$$PSOL(p, p_u) = \sum_i \min[w_{i,O(p)}, w_{i,O(p_u)}], \tag{29}$$

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*:

$$L_s(p, p_u) = \delta w(p_u) \left[1 - PSOL(p, p_u)\right], \tag{30}$$

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 $\delta$ 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

$$\begin{aligned} L_s(p) &= \sum_{u=1}^{N_p} \delta w(p_u) \left[1 - PSOL(p, p_u)\right] \\ &= \delta \left[1 - \sum_{u=1}^{N_p} w(p_u) PSOL(p, p_u)\right]. \end{aligned} \tag{31}$$

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

$$sol_p = \sum_{u=1}^{N_p} w(p_u) PSOL(p, p_u). \tag{32}$$

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

$$SOL_s = \sum_{p=1}^{N_s}\sum_{u=1}^{N_p} w(p)w(p_u)PSOL(p,p_u), \quad (33)$$

which captures shareholder commonality between firm $s$ and all other firms owning the upstream complementary patents. The holdup loss term in Proposition 2 can be approximated by

$$ln(1 - \overline{L}_s) \simeq -\overline{L}_s = -\delta[1 - SOL_s], \quad (34)$$

and substitution of Eq. (34) into Eqs. (25)–(28) makes the model directly testable. The expression $\delta SOL_s$ 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., 2016] |
| $cites_{p,t}$ | 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., 2016; 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., 2016; 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., 2016; 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 $CITES_{s,t}$ those citations that come from the upstream firms cited in the patent filings of the downstream firm $s$ in year $t$. [Source: Kogan et al., 2016] |
| $R\&D\ Exp_{s,t}$ | Total $R\&D$ expenditure (in million U.S. dollars) for firm $s$ in year $t$. The Compustat Mnemonic is $XRD$. It is 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 the cases in which 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., 2016; Thomson Reuters 13F database] |

| Variable | Description |
| --- | --- |
| $SOL_{s,t}$ | 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., 2016; Thomson Reuters 13F database] |
| $SOL\_$ $Dedicated_{s,t}$ | 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 $\Omega$ is the pool of all stocks held by the investor in the year. [Source: Kogan et al., 2016; 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., 2016; 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., 2016; CRSP and Thomson Reuters 13F databases] |

| Variable | Description |
|---|---|
| $SOL\_Placebo1_{s,t}$ | 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 five years (from $t-4$ to $t$), respectively. The suffix $\_mean$ refers to the industry average based on four-digit SIC codes. [Source: Kogan et al., 2016; 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 five 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., 2016] |
| $SOL\_Rank_{s,t}$ | (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., 2016] |

| Variable | Description |
|---|---|
| $SOL\_Rel_{s,t}$ | Relative cites-weighted shareholder overlap for firm $s$ in year $t$. It is the same as $SOL_{s,t}$ except that the weight for each patent is calculated based on the approach proposed by Lerner, Sorensen and Stromberg (2011). Specifically, we define the relative citation count of a patent $p$ filed in year $t$ as $cites\_rel_t(p) = \frac{cites^{3y}_{p,t}}{\frac{1}{N_k}\sum_{p\in k} cites^{3y}_{p,t}}$, where patent success is captured by the citation count over a three-year period after the patent is granted relative to the aggregate citation count of all $N_k$ patents in the same technology class $k$. The weights $w(p)$ and $w(p_u)$ in the calculation of shareholder overlap in Eq. (4) are based on $cites\_rel(p)$ and $cites\_rel(p_u)$, respectively. [Source: Kogan et al., 2016] |
| $SOL\_Equal_{s,t}$ | 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., 2016; Thomson Reuters 13F database] |
| $IO_{s,t}$ | 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] |
| $IO^{NOL}_{s,t}$ | 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 stake 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 thousand U.S. dollars) of firm $s$ at the end of year $t$. [Source: Compustat-CRSP merged database] |
| $R\&D\ Stock_{s,t}$ | Cumulative R&D investment (in million U.S. dollars) 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 $\delta$ represents the private depreciation rate of knowledge and is set to be 0.15. [Source: Compustat-CRSP merged database] |

| Variable | Description |
|---|---|
| $K/L_{s,t}$ | Capital (in million U.S. dollars) 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] |
| $Sales_{s,t}$ | Total amount of sales (in million U.S. dollars) 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] |
| $WHHI_{s,t}$ | 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., 2016; Thomson Reuters 13F database] |
| $SIF_{s,t}$ | Shareholder innovation focus for firm $s$ in year $t$. In the first step, we define for each listed firm 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., 2016; Compustat-CRSP merged database] |
| $MktCap^u_{p,t}$ | Average market capitalization value (in thousand U.S. dollars) of firms owning patent $p$'s upstream patents $u$ at the end of year $t$. [Source: Kogan et al., 2016; CRSP database] |

| Variable | Description |
|---|---|
| *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 than 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., 2016] |
| $SpillTech_{s,t}$ | 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 million U.S. dollars) 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., 2016; 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 million U.S. dollars) 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., 2016; Compustat-CRSP merged database] |

# 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] Azar, J., M. C. Schmalz, and, I. Tecu, 2015, Anti-competitive effects of common ownership, Working paper. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2427345.

[4] Bena, J., M. A. Ferreira, P. P. Matos, and P. Pires, 2016, Are foreign investors locusts? The long-term effects of foreign institutional ownership, Working paper, University of Virginia.

[5] Bessen, J., and E. Maskin, 2009, Sequential innovation, patents, and imitation, *RAND Journal of Economics* 40(4), 611–635.

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

[7] 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.

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

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

[10] 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.

[11] 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.

[12] 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.

[13] 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.

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

[15] 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.

[16] 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.

[17] 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.

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

[19] He, J., and J. Huang, 2014, Product market competition in a world of cross ownership: Evidence from institutional blockholdings, Working paper, Available at: http://ssrn.com/abstract=2380426.

[20] Heller M. A., and R. S. Eisenberg, 1998, Can patents deter innovation? The anticommons in biomedical research, *Science* 280(5364), 698–701.

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

[22] Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman, 2016, Technological innovation, resource allocation and growth, *Quarterly Journal of Economics*, forthcoming.

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

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

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

[26] 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.

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

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

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

[30] 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) $R\&D\ Exp_{s,t}$, the R&D expenditure for firm $s$ in year $t$, and (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 $t$. 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: $SOL\_Rel_{s,t-1}$ (an alternative cites-weighted measure), $SOL\_Rank_{s,t-1}$ (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 $u$. $IO_{s,t-1}$ and $IO^{NOL}_{s,t-1}$ represent the institutional ownership of, respectively, all 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 |
| $ln(1+R\&D\ Exp)$ | 19,020 | 2.267 | 2.195 | 2.075 | 0.575 | 0.000 | 0.000 | 5.087 | 9.408 |
| $ln(1+cites)$ | 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 |
| $ln(CITES)$ | 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 | 0.047 | 0.035 | 0.046 | 1.608 | 0.000 | 0.000 | 0.110 | 0.546 |
| $SOL\_Rel$ | 19,020 | 0.061 | 0.042 | 0.064 | 1.588 | 0.000 | 0.000 | 0.149 | 0.751 |
| $SOL\_Rank$ | 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 |
| $WHHI$ | 19,020 | 0.057 | 0.035 | 0.071 | 3.300 | 0.000 | 0.000 | 0.133 | 1.000 |
| $IO^{NOL}$ | 19,020 | 0.100 | 0.037 | 0.158 | 2.652 | 0.000 | 0.000 | 0.282 | 1.000 |
| $IO$ | 19,020 | 0.479 | 0.497 | 0.266 | −0.067 | 0.000 | 0.100 | 0.823 | 1.000 |
| $SIF$ | 19,020 | 0.281 | 0.280 | 0.075 | 2.294 | 0.000 | 0.197 | 0.367 | 2.644 |
| $ln(SpillTECH)$ | 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 $t-1$. The regressions also control for year and industry (based on four-digit SIC codes) or firm fixed effects. All regressions report robust standard errors clustered at the firm-year level 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** | 1.586** | 4.685** | 2.914** |
| | (0.374) | (0.371) | (0.684) | (0.677) |
| Controls: | | | | |
| $ln(MktCap)$ | 0.317** | 0.145** | 0.381** | 0.158** |
| | (0.011) | (0.015) | (0.021) | (0.026) |
| $ln(1+R\&D\ Stock)$ | 0.317** | 0.154** | 0.276** | 0.258** |
| | (0.009) | (0.025) | (0.020) | (0.050) |
| $ln(K/L)$ | 0.029 | −0.067* | 0.107** | −0.034 |
| | (0.019) | (0.031) | (0.034) | (0.054) |
| $ln(Sales)$ | −0.024** | −0.010 | −0.008 | −0.012 |
| | (0.009) | (0.017) | (0.015) | (0.024) |
| $Private\ Patent\ Share$ | 0.422** | 0.124 | 0.624** | 0.186 |
| | (0.094) | (0.094) | (0.160) | (0.161) |
| Year FE | YES | YES | YES | YES |
| Industry FE | YES | NO | YES | NO |
| Firm FE | NO | YES | NO | YES |
| *Obs.* | 19,020 | 19,020 | 5,898 | 5,898 |
| *Adj.* $R^2$ | 0.526 | 0.727 | 0.564 | 0.747 |

**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 $t-1$. The regressions also control for year and industry (based on four-digit SIC codes) or firm fixed effects. All regressions report robust standard errors clustered at the firm-year level 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** | 0.220 | 2.923** | 1.207** |
| | (0.207) | (0.236) | (0.238) | (0.203) |
| Controls: | | | | |
| *ln(MktCap)* | 0.081** | −0.010 | 0.199** | 0.136** |
| | (0.007) | (0.010) | (0.007) | (0.007) |
| *ln(1 + R&D Stock)* | 0.023** | −0.020 | 0.262** | 0.163** |
| | (0.005) | (0.015) | (0.005) | (0.015) |
| *ln(K/L)* | −0.039** | −0.092** | 0.060** | 0.025 |
| | (0.012) | (0.021) | (0.010) | (0.015) |
| *ln(Sales)* | −0.058** | −0.038** | 0.036** | 0.023** |
| | (0.006) | (0.012) | (0.005) | (0.008) |
| *Private Patent Share* | 0.073 | 0.047 | 0.350** | 0.058 |
| | (0.060) | (0.065) | (0.052) | (0.044) |
| Year FE | YES | YES | YES | YES |
| Industry FE | YES | NO | YES | NO |
| Firm FE | NO | YES | NO | YES |
| *Obs.* | 19,020 | 19,020 | 19,020 | 19,020 |
| *Adj.* $R^2$ | 0.427 | 0.593 | 0.614 | 0.835 |

**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 along the two dimensions. 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. $p$-values for the two null hypotheses described in the last two rows of the table for Column 2 are also reported. All regressions report robust standard errors clustered at the firm-year level 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) |
| $SOL$ | 3.692** | | 4.698** | | |
| | (0.374) | | (0.393) | | |
| $SOL\_Dedicated$ | | 25.151** | | | |
| | | (5.073) | | | |
| $SOL\_Intermediate$ | | 2.277** | | | |
| | | (0.820) | | | |
| $SOL\_Transient$ | | 4.041** | | | |
| | | (0.699) | | | |
| $WHHI$ | | | 2.449** | | |
| | | | (0.227) | | |
| $SOL\_Placebo1$ | | | | −0.418 | |
| | | | | (0.429) | |
| $SOL\_Placebo2$ | | | | | 0.653 |
| | | | | | (0.452) |
| Controls: | | | | | |
| $ln(MktCap)$ | 0.317** | 0.309** | 0.337** | 0.361** | 0.355** |
| | (0.011) | (0.012) | (0.012) | (0.011) | (0.011) |
| $ln(1+R\&D\ Stock)$ | 0.317** | 0.315** | 0.317** | 0.326** | 0.326** |
| | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) |
| $ln(K/L)$ | 0.029 | 0.029 | 0.035 | 0.036 | 0.034 |
| | (0.019) | (0.019) | (0.019) | (0.019) | (0.019) |
| $ln(Sales)$ | −0.024** | −0.025** | −0.022* | −0.017 | −0.018 |
| | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) |
| $Private\ Patent\ Share$ | 0.422** | 0.385** | 1.126** | −0.298** | −0.140 |
| | (0.094) | (0.095) | (0.119) | (0.098) | (0.098) |
| Year FE | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES |
| *Obs.* | 19,020 | 19,020 | 19,020 | 19,020 | 19,020 |
| *Adj.* $R^2$ | 0.526 | 0.527 | 0.530 | 0.523 | 0.523 |
| $H_0 : SOL\_Dedicated = SOL\_Transient$ | | 0.000 | | | |
| $H_0 : SOL\_Intermediate = SOL\_Transient$ | | 0.168 | | | |

**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 $t$. $sol_{p,t-1}$ 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 $ln(1+MktCap^u_{p,t-1})$ 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 $u$. All regressions report robust standard errors clustered at the firm-year level 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: | $ln(1+cites)$ | | | |
|---|---|---|---|---|
| | OLS | | 2SLS | |
| | | | $1^{st}$ Stage | $2^{nd}$ Stage |
| | (1) | (2) | (3) | (4) |
| *sol* | $0.192^{**}$ | $0.272^{**}$ | | $0.283^{**}$ |
| | (0.019) | (0.019) | | (0.013) |
| $ln(MktCap^u)$ | | | $0.024^{**}$ | |
| | | | (0.000) | |
| Year FE | YES | NO | NO | NO |
| Tech. FE | YES | NO | NO | NO |
| Firm FE | YES | NO | NO | NO |
| Year × Firm FE | NO | YES | YES | YES |
| Year × Tech. FE | NO | YES | YES | YES |
| *Obs.* | $581,240$ | $581,240$ | $581,240$ | $581,240$ |
| *Adj.* $R^2$ | 0.312 | 0.339 | 0.855 | |

**Table 6: R&D Expenditure and Shareholder Overlap**

Reported are OLS regressions of (log) R&D expenditure for the sample period 1992-2007. $R\&D\ Exp_{s,t}$ denotes the R&D expenditure 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. $IO^{NOL}_{s,t-1}$ denotes the institutional ownership of non-overlapping shareholders in firm $s$ at the end of year $t-1$. The control variables include the market capitalization ($MktCap_{s,t-1}$), 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 regressions also control for year and industry (based on four-digit SIC codes) or firm fixed effects. All regressions report robust standard errors clustered at the firm-year level in parentheses. Also reported are the total number of observations (*Obs.*) and adjusted R-squared (*Adj.* $R^2$).

| Dependent Variable: | $ln(1 + R\&D\ Exp)$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $SOL$ | 3.906** | 1.246** | 3.368** | 0.904* |
| | (0.375) | (0.367) | (0.368) | (0.367) |
| $IO^{NOL}$ | | | −0.712** | −0.359** |
| | | | (0.035) | (0.037) |
| Controls: | | | | |
| $ln(MktCap)$ | 0.506** | 0.286** | 0.481** | 0.283** |
| | (0.010) | (0.014) | (0.010) | (0.014) |
| $ln(K/L)$ | 0.184** | −0.021 | 0.172** | −0.020 |
| | (0.017) | (0.027) | (0.017) | (0.027) |
| $ln(Sales)$ | 0.125** | 0.112** | 0.119** | 0.109** |
| | (0.009) | (0.016) | (0.009) | (0.016) |
| $Private\ Patent\ Share$ | 0.210* | 0.107 | 0.538** | 0.261** |
| | (0.084) | (0.082) | (0.084) | (0.083) |
| Year FE | YES | YES | YES | YES |
| Industry FE | YES | NO | YES | NO |
| Firm FE | NO | YES | NO | YES |
| *Obs.* | 19,020 | 19,020 | 19,020 | 19,020 |
| *Adj.* $R^2$ | 0.578 | 0.767 | 0.586 | 0.769 |

**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, an alternative cites-weighted measure ($SOL\_Rel_{s,t-1}$), a non-parametric rank-based measure ($SOL\_Rank_{s,t-1}$), and an equal-weight measure ($SOL\_Equal_{s,t-1}$) of shareholder overlap. Column 11 measures the patent success by $ln(CITES_{s,t})$. 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 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 $t-1$, as well as year and industry (based on four-digit SIC codes) fixed effects. The sample period is 1992-2007. Robust standard errors clustered at the firm-year 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: | $ln(1+CITES)$ | | | | | Neg. Bino. $CITES$ |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $SOL$ | 3.692** | 3.658** | 3.567** | 3.724** | 3.581** | 2.802** |
| | (0.374) | (0.374) | (0.375) | (0.374) | (0.374) | (0.386) |
| $IO$ | | −0.681** | | | −0.666** | −0.798** |
| | | (0.059) | | | (0.059) | (0.073) |
| $SIF$ | | | 0.792** | | 0.710** | 0.791** |
| | | | (0.196) | | (0.192) | (0.246) |
| $ln(SpillTECH)$ | | | | 0.135** | 0.126** | 0.103** |
| | | | | (0.018) | (0.018) | (0.019) |
| $ln(SpillSIC)$ | | | | −0.035** | −0.034** | −0.012 |
| | | | | (0.011) | (0.011) | (0.012) |
| Control variables and Year and Industry FE | YES | YES | YES | YES | YES | YES |
| *Obs.* | 19,020 | 19,020 | 19,020 | 18,945 | 18,945 | 18,945 |
| *Adj.* $R^2$ | 0.526 | 0.530 | 0.527 | 0.527 | 0.531 | |

| Dependent Variables: | $ln(1+CITES)$ | | | | $ln(CITES)$ | $ln(1+CITES^F)$ |
|---|---|---|---|---|---|---|
| | (7) | (8) | (9) | (10) | (11) | (12) |
| $SOL$ | | | | | 3.454** | 3.498** |
| | | | | | (0.364) | (0.373) |
| $SOL(t-2)$ | 3.650** | | | | | |
| | (0.376) | | | | | |
| $SOL\_Rel$ | | 3.232** | | | | |
| | | (0.329) | | | | |
| $SOL\_Rank$ | | | 3.885** | | | |
| | | | (0.374) | | | |
| $SOL\_Equal$ | | | | 2.870** | | |
| | | | | (0.166) | | |
| Control variables and Year and Industry FE | YES | YES | YES | YES | YES | YES |
| *Obs.* | 19,020 | 19,020 | 19,020 | 19,020 | 17,609 | 19,020 |
| *Adj.* $R^2$ | 0.526 | 0.526 | 0.526 | 0.532 | 0.506 | 0.506 |

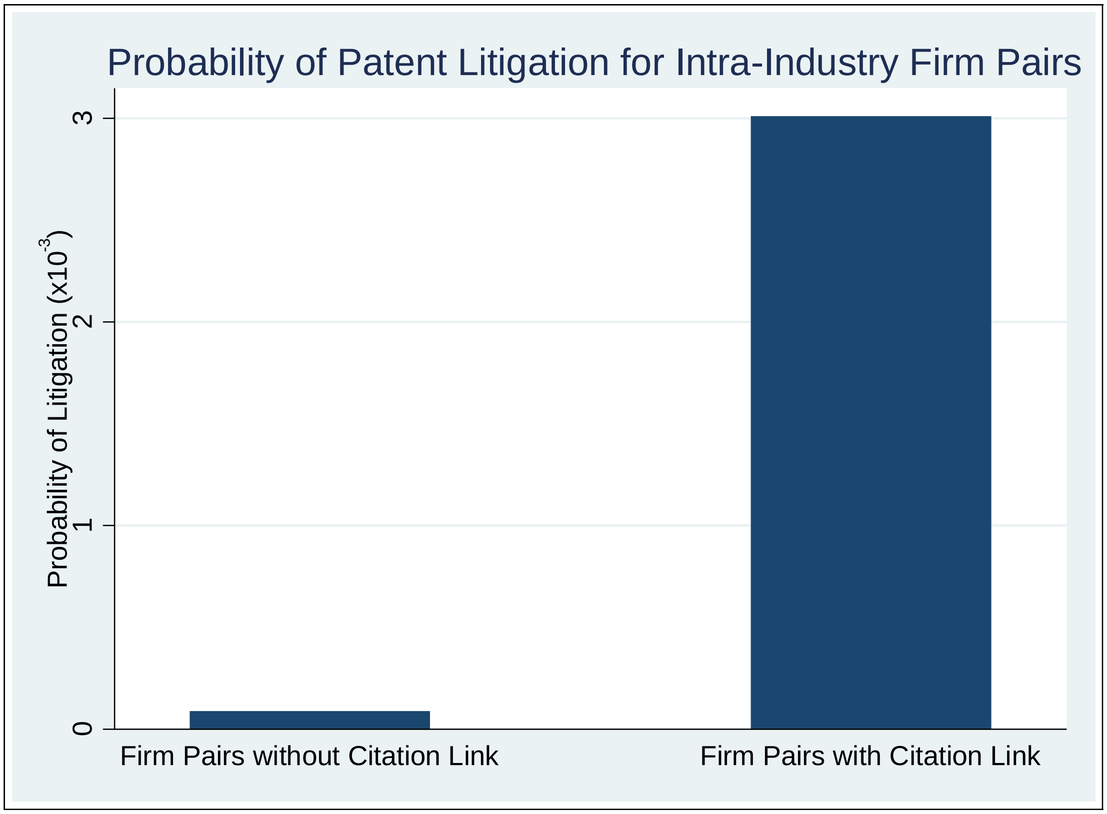


Figure 1: We compare the likelihood of patent litigation for listed firm pairs with citation links and those without any citation link in the past three years. For each year from 2000 to 2007, we form all intra-industry firm pairs (defined based on two-digit SIC codes) of all U.S. listed firms with at least one patent application in the past three years and sort them into pairs with at least one patent citation link and pairs without. The Audit Analytics Litigation database reports a total of 187 patent litigation cases for intra-industry firm pairs with citation links, which translates into a litigation likelihood per firm pair of 0.301% (= 187 cases/62,210 firm pairs), whereas 125 cases concern intra-industry firm pairs without any citation link, which translates into a litigation likelihood of 0.0085% (= 125 cases/1,462,972 firm pairs).

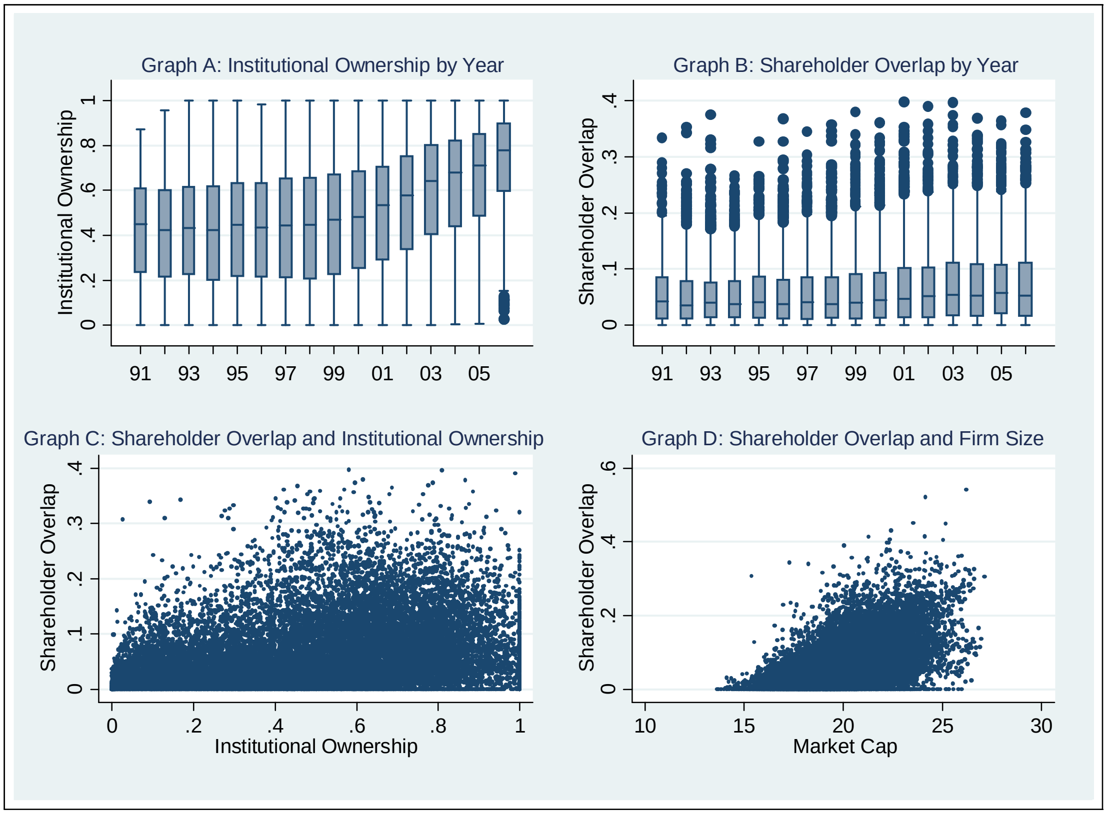


Figure 2: Graphs A and B are the box plots for the distribution of institutional ownership ($IO_{s,t}$) and shareholder overlap ($SOL_{s,t}$), respectively, by year from 1991 to 2006. The top, middle, and bottom value of each box represent the 75th, 50th, and 25th percentile of the distribution in the 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. Detailed definitions of the variables are given in Appendix B.

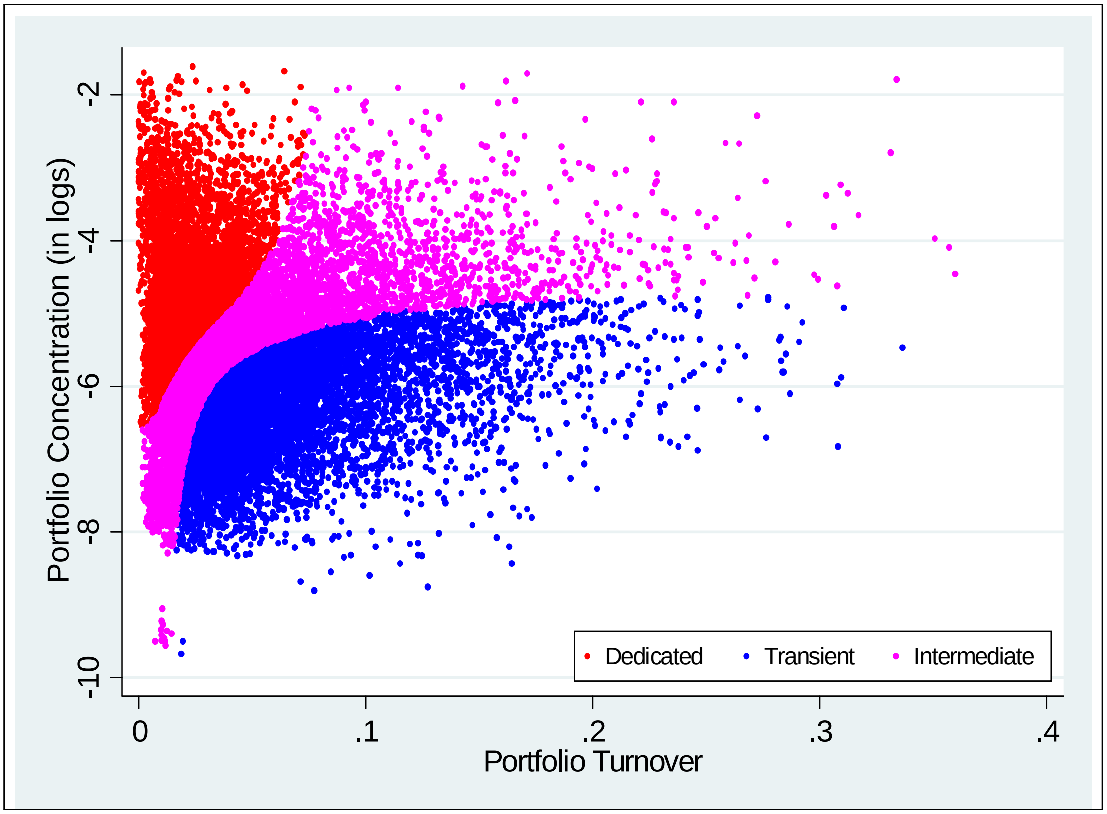


Figure 3: Plotted are the (log) portfolio concentration and (log) portfolio turnover of institutional investors (dedicated, intermediate, or transient investors) in our sample over the period 1991-2006. Specifically, at the end of each year, we rank all institutional investors along two dimensions: Their 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, respectively, the top, middle, and bottom tercile of the combined rank along the two dimensions. Detailed definitions of the variables are given in Appendix B.

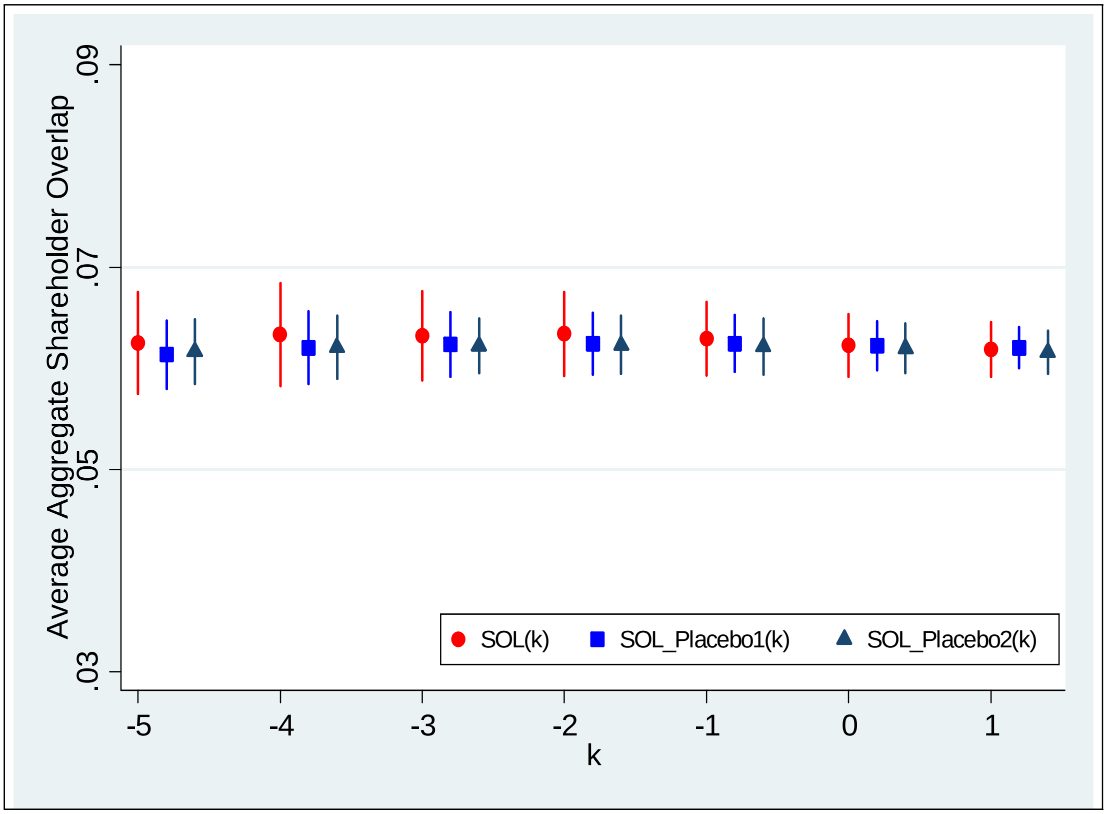


Figure 4: The evolution of the average shareholder overlap $\overline{SOL}(k)$ between the innovating firm and other firms owning the upstream complementary patents is plotted for the period from five years prior to the patent filing year to five years after the filing (i.e., $k = -5$ to 1), with the patent filing year denoted by $k = 0$. $\overline{SOL}(k)$ is calculated according to Eq. (12). 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 are also plotted. For ease of comparison, in the plot we adjust the value of $\overline{SOL\_Placebo1}(k)$ and $\overline{SOL\_Placebo2}(k)$ upward by 0.012 and 0.015, respectively, so that they would have the same mean value as $\overline{SOL}(k)$.