Technological Progress and Ownership Structure

Heng Geng∗
University of Hong Kong

Harald Hau∗∗
University of Geneva and Swiss Finance Institute

Sandy Lai∗∗∗
University of Hong Kong

May 25, 2015

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 intensive and extensive margins of patent production. The positive innovation effect of shareholder overlap is strongest for active investors (such as hedge funds) and for concentrated overlapping ownership. We also show that shareholder overlap is related to higher R&D expenditure and less patent litigation and increases for active investors before the patent filing of the downstream firm becomes public information.

JEL Classification: L22, G31, G32
Keywords: Patents, holdup Problems, Innovation, Institutional Ownership

∗Faculty of Business and Economics, University of Hong Kong, K.K. Leung Building, Pokfulam Road, Hong Kong. E-mail: gengheng1989@hku.hk.

∗∗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

While technological progress is recognized as the main source of long-run economic growth, its relation with corporate ownership structure and property rights in patents is less understood. This paper provides a new empirical perspective on the role of the equity ownership structure in attenuating holdup problems induced by patent protection in the corporate innovation process.

Patents protection provides the inventor with exclusive rights to the commercial use of his discoveries. But, such discoveries are often part of a larger technological process of interdependent innovations, in which 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 holdup problems whenever such complementary patents are owned by different firms and ex-ante contracting is incomplete.¹

Building on the property rights theory of Hart and Moore (1990), we argue that joint equity ownership in the innovating firm and firms controlling complementary patents can attenuate the holdup problem and contribute to the patent success of the innovating firm. From the perspective of an equity investors who owns shares in both the upstream and downstream firm, any hold-up in pursuit of patent rents is costly and his interest consists in a swift conflict resolution without patent litigation. We develop a simple model of optimal patent investment to show how such joint equity ownership can increase the intensive and extensive margins of the patent production. An important implication from our model is that a liquid equity market with flexible equity ownership across various innovating firms can promote innovation and technological progress, and such an institution is particularly important under the current patent environment featuring an increasingly extensive patent protection and a proliferation of patent rights.

To subject this property right perspective of patent success to an 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 1991–2010. In particular, we track stock ownership not only for the innovating firms, but also for firms owning the complementary patents. The complementarities are identified directly from the patent filings that explicitly list important precursory patents owned by other firms. By law, each newly filed

¹The recent economic research has documented a negative impact of recent patent proliferation on R&D investment and follow-on innovation (Heller and Eisenberg, 1998; Bessen and Maskin, 2009; Galasso and Schankerman, 2015) and highlighted abusive patent enforcement by so-called “non-practising entities” (Cohen, Gurun, and Kominers, 2014).
patent must list prior precursory (upstream) patents that are technologically related and material to the patentability of the new application.\textsuperscript{2} These upstream patents might have to be licensed to the (downstream) innovator for him to realize the full value of the new patent (Ziedonis, 2004; Noel and Schankerman, 2013; Galasso and Schankerman, 2010).\textsuperscript{3} Our analysis identifies potential patent holdup based on this list of precursory patents and assumes that the list is exogenously determined by the technology to be patented. Patent examiners frequently add precursory patent to the reference list, suggesting a limited scope by the patent filing firms in manipulating the reference list of precursory patents (Alcácer, Gittelman, and Sampat, 2009).\textsuperscript{4}

Our main hypothesis states that joint equity ownership between the downstream innovator and the upstream firms controlling complementary patents attenuates holdup problems, increases R&D investment, and contributes to the long-run patent success of the innovating firm. Following the existing literature, we measure \textit{patent success} by the cumulative citation count of each patent that was filed and granted; this measure can be viewed as the intensive margin of patent production.\textsuperscript{5} The extensive margin of patent production is measured by the number of successful patents (i.e., patent applications that are eventually approved by USPTO) a firm files in a given year. Our baseline analysis relates \textit{firm-level patent success} ($\ln(1+CITES)$) to \textit{shareholder overlap} ($SOL$). The former is calculated as the (log of) total number of future citations for all patents a firm filed in a given year, whereas the latter measures the average aggregate (percentage) shares held by the group of shareholders that invest in both the patent filing firm and the firms owning upstream complementary patents. Consistent with the \textit{holdup attenuation hypothesis} of joint equity ownership, we find that $SOL$ emerges as the statistically and economically most significant determinant of patent success, and it is positively related to both the intensive and extensive margin of patent production. This result holds regardless of whether $SOL$ is measured based on equity ownership overlap in the year just before the patent application or two to four years prior

\textsuperscript{2}The U.S. patent law requires an invention to be useful, novel, and non-obvious to be patented.

\textsuperscript{3}Ziedonis (2004) argues that owners of the (upstream) cited patents are reasonable proxies for the potential licensors of the citing patent. Noel and Schankerman (2013) and Galasso and Schankerman (2010) also suggest that a greater number of upstream assignees can signal a greater number of negotiations and disputes required for the commercialization of the downstream patent.

\textsuperscript{4}Patent examiners in USPTO are officially responsible for constructing the list of prior art references. However, inventors also have a “duty of candor” to disclose all material prior art, and failure to do so can result in an “inequitable conduct” and the court may render the patent unenforceable. Using data from USPTO for all patents granted over the period 2001–2003, Alcácer, Gittelmanb, and Sampate (2009) document that examiners insert at least one citation in 92% of patent applications. Overall, examiner citations account for 63% of all citations made by an average patent.

\textsuperscript{5}See for example Aghion, Van Reenen and Zingales (2013) for a similar definition of firm level patent success.
to the application date.

We further hypothesize that shareholder overlap may represent an even more powerful mechanism for holdup resolution if the respective shareholders are so-called “activist” shareholders. In particular, more active investors such as hedge funds should contribute more to the resolution of rent extraction problems than other less active investors. Consistent with this intuition, we find a much stronger effect of shareholder overlap on patent success when such overlap or joint ownership originates in hedge fund holdings.

Another related hypothesis concerns the ownership concentration of the group of overlapping investors. If the downstream innovating firm and upstream firms are jointly owned by only a few relatively large shareholders, coordinated action might be easier to organize, and shareholders may have stronger incentives to resolve a potential holdup. In accordance with this prediction, we find that the Herfindahl-Hirschmann index of (overlapping) shareholder ownership concentration correlates positively with firm-level patent success beyond the shareholder overlap itself. Thus, coordination and contracting problems not only are an issue with respect to ex-ante bargaining between different patent owners, but also constrain the overlapping shareholders in their effort to overcome the ex-ante bargaining failures.

Our analysis controls for a variety of firm characteristics and also applies firm and time fixed effects. Yet, unobservable time-varying and firm-specific factors may still pose an inference problem if they influence both shareholder overlap and patent success. Therefore, we reproduce our regressions at the patent level while controlling for interacted firm and time fixed effects. These specifications directly compare the success of any two patents (in term of their future citations) filed by the same firm in the same year. We show that even within the same firm-year, patent success is correlated with the varying degree of holdup across a firm’s cohort of patents.

Besides patent success itself, we also analyze three additional variables related to the patent process. First, our model implies the holdup attenuation of joint shareholder ownership should generally mitigate the adverse effect of patent holdup on R&D investment. Empirically, we indeed find that R&D expenditure features a statistically and economically strong positive correlation with shareholder overlap. Second, we show that potential holdup situations are less likely to lead to costly patent litigation if the shareholder overlap between the down-stream innovator and the up-stream firm is larger [section to be added].

Third, if shareholder overlap can mitigate frictions

---

6Both the Innovation Act (H.R.3309) and the Patent Transparency and Improvement Act (S. 1720) propose
in ex-ante bargaining about patent rights, we expect it to efficiently adjust over the patent life cycle. In particular, privately informed investors may internalize hold-up problems by strategically increasing their overlapping ownership position before the filing of the downstream patent becomes public information. We find indeed that such shareholder overlap increases (modestly) for the most activist investors (like hedge funds) by approximately 11% in the 3 years prior to the public disclosure of patent filings. Yet this dynamic market timing effect appears economically small compared to the long-run increase in shareholder overlap due to the increase in institutional investment.

Our analysis also examines two alternative hypotheses. First, having tech-savvy shareholders, who invest mostly in innovative firms, might lead to an increase in shareholder overlap and simultaneously constitute a governance advantage for a firm engaged in patent competition. We control for this factor by constructing a firm-level proxy of shareholder innovation focus SIF. Second, a larger share of institutional ownership IO might procure a general governance advantage for the firm. Aghion, Van Reenen, and Zingales (2013) highlight institutional investors’ monitoring role and their willingness to support risky R&D investments that typically pay off only in the long run. We find shareholder innovation focus and institutional ownership only feature an economically weak relation with a firm’s long-run patent success once we control for shareholder overlap. We also control for a number of other variables, which include the previous R&D investment ln(1 + R&D Stock), a measure of relative capital intensity ln(1 + Capital/Labor), a firm size measure ln(1 + Sales), and year and industry (or firm) fixed effects. Yet, the relation between the holdup attenuation factor—shareholder overlap—and future patent success remains both statistically and economically significant even with these controls.

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. Consequently, ex-post holdup should reduce ex-ante investment incentives unless such rent extraction can be mitigated by overlapping equity ownership (especially from activist investors). We also note that for such a positive effect on ex-ante investment incentives to occur, the shareholder overlap needs to be acquired before the share prices of the upstream firms
fully reflect the holdup rents.\textsuperscript{7}

In the following section, we survey the related literature. Section 3 develops a simple model of patent holdup in the spirit of Hart and Moore (1990); it develops the main hypotheses and motivates the regression specifications. Section 4 discusses the data; Section 5 presents the main evidence for the role of shareholder overlap for patent success. Section 6 features related evidence on the role of shareholder overlap for R&D expenditure and patent litigation as well as the endogenous adjustment of shareholder overlap over the patent cycle. Various robustness checks are undertaken in Section 7, followed by conclusions in Section 8.

\section{Related Literature}

Existing studies on patent holdup problems (e.g., Shapiro, 2001; Ziedonis, 2004; and Hall and Ziedonis, 2007) find that licensing agreements are commonly used in practice—yet these might typically concern the ex-post rent allocation. Licensing agreement may involve substantial royalty fees and their negotiation may not represent a frictionless process. Alternatively, a firm may invent around the patented technology to avoid being held up, but this is not always possible given the cumulative and sequential nature of technological development. There is also evidence that firms seek outright ownership integration via mergers to resolve patent disputes. However, such merger cases are often challenged by court and eventually fail for anti-competitive reasons (Creighton and Sher, 2009). Our study suggests that in liquid equity markets, partial ownership integration via ownership overlap may be achieved at lower costs or may already exist if large institutional shareholders happen to hold shares in both concerned firms.

Notwithstanding its prominence in economic theory, the property rights view of the boundaries of the firm has seen very few empirical applications because of a variety of obstacles. First, non-contractible holdup problems might often be difficult to identify in a complicated business environment. Explicit citation of precursory patents in the patent documents provides a unique identification opportunity. Second, any underinvestment at the project level also tends to be difficult to measure because a firm can shift investments to other projects for which holdup problems are less severe. Such an analysis, therefore, requires a level of disaggregation typically not

\textsuperscript{7}If ownership integration can only be achieved ex-post at a share price for the upstream firm that fully reflects the holdup rents, then such ex-post integration amounts to transferring the holdup rents to the upstream firm’s shareholders. Therefore, no mitigation of the ex-ante underinvestment problem is obtained.
available for investment data. Third, investments may also involve intangibles resources (such as managerial attention) that pose additional measurement problems. For these reasons, we infer the (latent) underinvestment indirectly from diminished project (or patent) success. Future patent citations provide a sufficiently precise proxy for patent success at the firm (and patent) level to allow for a comprehensive study of holdup in the patent process.

Patent reform has become a hotly debated policy issue. President Obama in his 2014 State of the Union Address singled out the patent system as a priority for economic reform. The Administration has pushed USPTO to examine patent requests more rigorously and defined their patentable component more narrowly ex ante in order to reduce the reliance on the courts to make those determinations ex post. Galetovic, Haber, and Levine (2014), however, argue that there is no evidence that more patent litigation is associated with patent holders stymieing the commercialization of complex technologies or hindering innovation. The evidence in our paper suggests that shareholder overlap represents an important palliative to holdup problems with respect to patent investment. Moreover, we show that the liquidity of the U.S. stock market allows for the dynamic rearrangement of shareholder overlap and thus represents a complementary institution helping to maintain high levels of R&D expenditure under extensive patent protection.

Other empirical work on the determinants of patent success focuses on the role of institutional shareholders. Aghion, Van Reenen, and Zingales (2013) argue that a large share of institutional shareholders is conducive to patent investment as these shareholders tend to pursue a long-run objective. Our more extensive data, however, shows that the statistical and economic significance of the institutional ownership variable diminishes once we control for the role of shareholder overlap as the relevant determinant of patent success. Bena, Ferreira, and Matos (2014) relate patent success to foreign ownership, but it is unclear whether foreign ownership just proxies for more shareholder overlap in complementary patents as identified in this study. Brav, Jiang, and Tian (2014) show that hedge fund activism leads to more efficient use of innovative resources and human capital. Our study complements their finding and identifies activist shareholders as an important mechanism to alleviate holdup problems in corporate innovation processes. Recent empirical work has also highlighted the complementarity between equity market development and the degree of patent innovation both in the cross-section of countries (Hsu, Tian, Xu, 2014) and

---

8A deterioration in the patent environment is sometimes attributed to the growing role of non-practicing entities (NPEs) referred to as “patent trolls”, which specialize in the enforcement of patent rights without having a commercial activity of their own.
for some particular event (Ostinelli, 2014). Insofar as equity market development allows for a better internalization of holdup problems (though enhanced and adjustable shareholder overlap), our study offers a deeper microeconomic interpretation rooted in the theory of the firm for these documented findings.

3 A Model of Patent Investment

3.1 A Simple Benchmark (with No Holdup Effect)

A risk neutral firm $s$ can invest into a continuum of patent projects. Each project is represented by the index number $p$ on the interval $[0, \infty)$, where a higher index number corresponds 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 number of future citation counts $\text{cites}_s(p)$. Hence,

$$ V_s(p) = \alpha \times \text{cites}_s(p), $$

where $\text{cites}_s(p)$ is a random variable with the expected value $E[\text{cites}_s(p)] = \mu_s$, and $\alpha > 0$ is a constant. The total expected firm value $\Pi_s$ follows as

$$ \Pi_s = \max_p \int_0^p [\alpha \mu_s - C(p)] \, dp, $$

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

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

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 $E[\text{cites}_s(p)] = \mu_s$ 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

\[ \ln[L] = \frac{1}{b} \ln \alpha + \frac{1}{b} \ln(\mu_s) \]  

(ii) the firm-level (log) citation counts

\[ \ln[CITES_s] = \ln \int_0^{P_L} E[cites_s(p)] dp = \frac{1}{b} \ln \alpha + \frac{b + 1}{b} \ln(\mu_s), \]  

(iii) the (log) R&D expenditure

\[ \ln[R&D\ Exp] = \ln \int_0^{P_L} cP^b dp = \ln \frac{c}{1 + b} + \frac{b + 1}{b} \ln(\mu_s). \]

The firm-level (log) citation count in Eq.(5) is equal to the (log) extensive margin in Eq. (4) plus the (log) intensive margin \( \ln E[cites_s(p)] = \ln(\mu_s) \). Empirically, we can approximate the intensive margin by the average citation count \( \text{cites}_s \) of a firm’s patents.

3.2 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 each patent \( p \) requires consent from the owners of upstream patent \( p_u \) with \( u = 1, 2, ..., N_p \). \(^9\) These upstream patents allow their owners to extract part of the value (through, e.g., license fees) so that the firm’s expected patent value decreases. We denote the share of patent value lost to each upstream patent by \( \ell_s(p, p_u) \) and the aggregate value loss by

\[ L_s(p) = \sum_{u=1}^{N_p} L_s(p, p_u). \]  

\(^9\)Note that \( p_d \) does not include any expired patents because they do not pose any threat to the commercialization of the citing patent.
place so that \( L_s(p) = L_s(p, p_u) = 0 \). By contrast, the maximal rent extraction occurs if there is no overlap in institutional ownership between the downstream innovating firm and the upstream cited firms. For simplicity, we assume that the ex-ante expectation for value loss is identical for all patents \( p \) produced by the same firm; hence, \( E[L_s(p)] = \bar{T}_s \).

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

\[
E[cites_s(p)] = \mu_s[1 - \bar{T}_s]^{\gamma},
\]  

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

\[
E[V_s(p)] = \alpha[1 - \bar{T}_s] \cdot E[cites_s(p)] = \alpha \mu_s[1 - \bar{T}_s]^{1+\gamma}.
\]  

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

\[
\max_{\bar{P}_L} \Pi_s = \int_0^{\bar{P}_L} \left[ \alpha \mu_s[1 - \bar{T}_s]^{1+\gamma} - C(p) \right] dp,
\]

where the optimal patent range \([0, \bar{P}_L]\) has the upper limit

\[
\bar{P}_L = \left( \frac{\alpha \mu_s}{c}[1 - \bar{T}_s]^{1+\gamma} \right)^{\frac{1}{\beta}}.
\]

**Proposition 2: Patent Production in the Patent Holdup Case**

A value maximizing firm optimally invests in the production of patents on the line interval \([0, \bar{p}]\). Given a patent-level (ex-ante) expected citation count \( E[cites_s(p)] = \mu_s[1 - \bar{T}_s]^{\gamma} \) that is proportional to the expected patent value, a convex cost function \( C(p) = \alpha p^\beta \), and an (ex-ante) expected value loss \( \bar{T}_s = E[L_s(p)] \) for each patent due to patent holdup, we find for
(i) the (log) extensive margin of patent production

\[
\ln[\overline{P}_L] = \frac{1}{b} \ln \frac{\alpha}{c} + \frac{1}{b} \ln(\mu_s) + \frac{1 + \gamma}{b} \ln[1 - \overline{L}_s] \tag{12}
\]

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

\[
\ln[CITES_s] = \frac{1}{b} \ln \frac{\alpha}{c} + \frac{b + 1}{b} \ln(\mu_s) + \frac{1 + \gamma + b\gamma}{b} \ln[1 - \overline{L}_s], \tag{13}
\]

(iii) the (log) R&D expenditure

\[
\ln[R&D \ Exp] = \ln \frac{c}{1 + b} + \frac{b + 1}{b} \ln \frac{\alpha_s}{c} + (1 + \gamma) \frac{b + 1}{b} \ln[1 - \overline{L}_s]. \tag{14}
\]

The first and second terms in Eq. (13) are the same as those in Eq. (5). The third term in Eq. (13) and Eq. (14) captures how the holdup problem reduces, respectively, the overall patent success and R&D expenditure. All three equations feature the same (log) loss term \(\ln[1 - \overline{L}_s] < 0\).

We also note that the holdup problem affects not only the extensive margin (\(\overline{P}_L\)), but also the intensive margin \(E[cites_s(p)]\) of patent production if \(\gamma > 0\).

### 3.3 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 \(E[L_s(p)] = \overline{L}_s\). Our basic assumption is that shareholder overlap reduces holdup. 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 upstream patent \(p_u\) (listed in the patent filings) can be defined as

\[
PSOL(p, p_u) = \sum_i \min[w_{i,O(p)}, w_{i,O(p_u)}], \tag{15}
\]

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 omit the time index \(t\) from all variable expressions in this subsection. We assume the following reduced form for the distributive value loss function associated with the
upstream patent \( p_u \) cited by patent \( p \):

\[
L_s(p, p_u) = \delta w(p_u) [1 - PSOL(p, p_u)],
\]

where 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 \( p \). 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

\[
L_s(p) = \sum_{u=1}^{N_p} \delta w(p_u) [1 - PSOL(p, p_u)]
\]

\[
= \delta \left[ 1 - \sum_{u=1}^{N_p} w(p_u) PSOL(p, p_u) \right].
\]

We can define \textit{patent-level shareholder overlap} as

\[
sol_p = \sum_{u=1}^{N_p} w(p_u) PSOL(p, p_u).
\]

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

\[
\overline{L}_s = \sum_{p=1}^{N_s} w(p)L_s(p) = \delta \sum_{p=1}^{N_s} \sum_{u=1}^{N_p} w(p)w(p_u) [1 - PSOL(p, p_u)]
\]

\[
= \delta \left[ 1 - \sum_{p=1}^{N_s} \sum_{u=1}^{N_p} w(p)w(p_u) PSOL(p, p_u) \right],
\]

where the weight \( w(p) \) denotes the relative importance of patent \( p \). The \textit{firm-level shareholder overlap} can be defined as

\[
SOL_s = \sum_{p=1}^{N_s} \sum_{u=1}^{N_p} w(p)w(p_u) PSOL(p, p_u),
\]
which captures shareholder commonality between firm $s$ and all other firms owning the upstream patents. The holdup loss term in Proposition 2 can be approximated by

$$ln(1 - T_s) \simeq -T_s = \delta[SOL_s - 1],$$ \hspace{1cm} (20)

and substitution makes the model directly testable. The expression $\delta SOL_s$ capture the holdup attenuation through firm-level shareholder overlap relative to a total (non-attenuated) holdup effect embodied by $\delta$.

A final measurement issue concerns the choice of weights reflecting the relative importance of any patents $p$ and $p_u$. Empirically, we measure the relative importance by the relative (log) citation count as follows:

$$w(p) = \frac{ln[1 + cites_s(p)]}{\sum_{p=1}^{N}\ln[1 + cites_s(p)]} \text{ and } w(p_u) = \frac{ln[1 + cites(p_u)]}{\sum_{u=1}^{N_p}\ln[1 + cites(p_u)]}. \hspace{1cm} (21)$$

In the robustness section (Section 7), we show that an alternative weighting scheme using a (non-parametric) rank measure of future citations $\text{rank}(p)$ in Eq. (21) delivers very similar results. The results are also robust to using equal weights, albeit at a slightly weaker level of economic significance.

3.4 Hypotheses

We summarize the main testable hypotheses in this subsection. The hypotheses H1, H2, H4, and H5 follow directly from model, whereas hypotheses H3, H6, and H7 are extensions based on simple economic arguments.

**H1: Firm Patent Success and Holdup Attenuation**

The patent success of firm $s$ (in terms of future citation $CITES_{s,t}$) for its cohort of patents filed in year $t$ should increase in the *firm-level shareholder overlap* $SOL_{s,t-1}$ between the firm itself and all other firms owning cited upstream patents that pose potential holdup problems. In our main empirical analysis, we measure $SOL_{s,t-1}$ based on equity ownership at the end of year $t - 1$. In the robustness analysis (Section 7), we further verify our results using ownership measured in years $t - 2$ to $t - 4$.

**H2: Extensive and Intensive Margins of Patent Production**
The extensive margin of patent production (proxied by the number of patents \( N_{s,t} \) filed by firm \( s \) in year \( t \) and eventually granted) correlates positively with the firm-level shareholder overlap \( SOL_{s,t-1} \). The intensive margin \( cites_{s,t} \) (which measures the average citation success of a firm’s patents) should also correlate positively with \( SOL_{s,t-1} \) if patent holdup involves not only value redistribution but also (inefficient) value destruction (i.e., \( \gamma > 0 \)).

A straightforward extension of the holdup hypothesis distinguishes between shareholder types. We conjecture that active shareholders (such as hedge funds) should be more willing and/or more capable of exercising their ownership power to resolve patent holdup than more passive shareholders (such as most pension funds). Moreover, a more concentrated ownership among overlapping shareholders should overcome the free-rider problem of costly lobbying and contribute to patent holdup resolution.

**H3: Shareholder Type and Concentration of Shareholder Overlap**

Shareholder overlap should feature a more positive correlation with patent success if the respective overlap is contributed by more active shareholders and if the overlapping equity stake is concentrated among few shareholders.

The patent holdup and its patent-specific attenuation through shareholder overlap should operate not only at the firm level, but also across different patents filed by the same firm in the same year. Accordingly, we can formulate the following within firm hypothesis:

**H4: Patent Success within the Firm**

Within a firm’s cohort of patents filed in the same year \( t \), those with the largest patent-level shareholder overlap (denoted by \( sol_{p,t-1} \) and measured with respect to each individual patent \( p \)’s cohort of upstream cited patents) should feature the largest patent-level success (denoted by \( cites_{p,t} \)) in terms of the total future citations for the patent.

This latter hypothesis has the advantage that it can be tested using a rich set of interacted firm-year fixed effects—thereby controlling for time varying unobservable firm heterogeneity.

Based on Eq. (14) we can be summarize the role of shareholder overlap for the ex-ante investment incentives as follows:
**H5: Firm R&D Expenditure**

A firm’s (log) R&D expenditure (denoted by $R&D\ Exp_{s,t}$) increase (linearly) in firm-level shareholder overlap $SOL_{s,t-1}$.

Direct evidence for patent holdup consist in patent litigation between the upstream and downstream patent firms. We conjecture that the likelihood of such costly judiciary conflict is reduced by shareholder overlap because the joint owners of the subject companies have strong incentives to avoid the ‘negative-sum-game’ of litigation. We can directly test the following hypothesis:

**H6: Patent Litigation**

Less holdup under more shareholder overlap is reflected in less patent litigation.

If shareholder overlap is indeed an effective mechanism to alleviate patent holdup and contributes to efficiency through better ex-ante investment incentives, we expect shareholder ownership to dynamically adjust so as to reduce or even minimize holdup inefficiencies. Importantly, shareholder overlap should increase before holdup rents are reflected in the share prices of the upstream firms. Otherwise, the acquisition of ownership overlap amounts to an implicit payment of the holdup rents by the acquirer to the shareholders of the upstream firms, without the positive ex-ante effects on patent investment of the downstream firm. The dynamic extension of the model can be summarized as follows:

**H7: Ownership Overlap over the Patent Life Cycle**

Shareholder overlap between upstream and downstream patent owning firms should increase prior to the public release of the patent filing for the downstream patent. The dynamic adjustment in shareholder overlap should be strongest for active investors (such as hedge funds).

The last hypothesis also relates to the benefits of equity market liquidity in general. The lower the costs of rearranging equity ownership in a liquid stock market, the less any ex-ante R&D expenditure is constrained by future holdup in a world of imperfect ex-ante contracting. In other words: Equity market liquidity represents a complementary institution for any innovation process based on extensive patent protection. The constant dynamic rearrangement of equity ownership can help minimize holdup costs.
Our sample combines institutional ownership data with annual patent and citation data for publicly-listed firms in the United States. The ownership data is from the Thomson Reuters 13F database. SEC requires all institutional organizations, companies, universities, etc. 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) document reporting inconsistencies in the ownership data prior to 1991, so we only use the ownership data from 1991 onwards.

We collect patent and citation information from the latest version of the National Bureau of Economic Research (NBER) Patent Citation database, which includes annual data for patents granted during 1976–2006. We further supplement the NBER data with the data from Kogan, Papanikolaou, Seru, and Stoffman (2014). The combined data set provides annual patent and citation information for patents granted over 1976–2010.10

Our measurement of innovation success follows the existing literature (Griliches, Pakes, and Hall, 1988). To distinguish influential innovations from incremental technological discoveries, we use the total number of a patent’s future citations \( c_{it} \) from the patent filing year \( t \) to 2010 as our proxy for patent success. Typically, a patent starts to receive citations only after it is granted unless the invention is already known by others at the patent application stage.11 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 3 years of application, so the lag between patent filing and the first citation can range from zero to three years or longer.

At the firm level, we aggregate the count statistic \( c_{it} \) to the total number of future patent citations generated by all granted patents filed by firm \( s \) in year \( t \), denoted by \( CITES_{s,t} \). Self-citations are excluded, and 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 successful patent filings (i.e., patent applications that are eventually granted) by firm $s$ in year $t$. The corresponding intensive margin is measured by the average cites per patent $\text{cites}_{s,t}$ (which equals the ratio of $\text{CITES}_{s,t}$ to $N_{s,t}$). Because most of these patent related measures feature highly right skewed distribution, we generally apply a log transformation $\ln(1 + X)$ in order to obtain more normally distributed variables for regression analysis.

We carefully adjust 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 applications filed in 2009 and 2010 are still not included in the data set. 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 the citations only up to 2010. Following Hall, Jaffe, and Trajtenberg (2001, 2005), we correct for the truncation bias in citation counts by estimating the shape of the citation-lag distribution.

The key explanatory variable is the lagged firm-level shareholder overlap ($\text{SOL}_{s,t-1}$) between the innovating firm and all other firms controlling complementary patents. Calculation of $\text{SOL}_{s,t-1}$ follows the definition in Eq. (19) and is based on the ownership data at the end of year $t - 1$ for patents filed in year $t$. We measure ownership overlap at year $t - 1$ to reduce the scope for reverse causality from patent applications (in year $t$) to shareholder ownership changes and therefore variations in $\text{SOL}$.

Because expired patents should not create any holdup problems, we ignore cited patents that are filed 20 years before the application date of the citing patents in constructing $\text{SOL}$. Moreover, we discard firm-year observations when a firm does not have any successful patent application. We track potential holdup situations only for those cases for which both the patent filing firm and the upstream patent owner are publicly listed firms. Upstream patents owned by private (non-listed) companies do not enter into the $\text{SOL}$ measure as we cannot determine any ownership overlap in these cases. Successful patent filings for which none of the upstream patent owners is a publicly listed firm are again discarded from the sample. Generally, downstream firm owners should find it difficult to acquire an overlapping ownership stake in a private firm, thus limiting the scope of

---

12 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 of the greater of the 20 year term discussed earlier or 17 years from the grant date. (See http://www.uspto.gov/inventors/patents.jsp#heading-5.)
the attenuation effect.\textsuperscript{13}

The citation count variable as a proxy for patent success has the important advantage that it can be measured not only at the firm level but also at the patent level. Analogously, we can also measure shareholder overlap at both the firm level (\textit{SOL}) and the patent level (\textit{sol}), based on Eq. (18). The weighted sum of the patent-level shareholder overlap amounts to the firm-level shareholder overlap, as implied by Eq. (19).

We also use a series of control variables, namely the previous R&D investment $ln(1 + R\&D \textit{Stock}_{s,t-1})$, a measure of relative capital intensity $ln(1 + \textit{Capital/Labor}_{s,t-1})$, and a firm size measure $ln(1 + \textit{Sales}_{s,t-1})$. For simplicity, $ln(1 + \textit{Capital/Labor})$ is abbreviated as $ln(1 + K/L)$. To calculate these control variables, we obtain the accounting data from Compustat and the stock price and shares outstanding data from CRSP. Firm-level regressions also control for the pre-sample mean citation count for each firm, which is measured by the aggregate citation count of a firm’s patents filed over the pre-sample period 1976–1991, divided by the number of pre-sample years. For firms without patent filings during this period, the value is set to zero.

Our final sample includes all U.S. publicly listed firms that have at least two successful patent applications 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 3,053 firms. We exclude all firm-year observations with missing values for the explanatory variables. The summary statistics are reported in Table 1. The sample features 19,622 firm-years of patent production involving a total of 713,319 patents. On average, a firm produces 13 patents and obtains an average number of 108 citations per year. The average (median) firm-level shareholder overlap (\textit{SOL}) is 14.1\% (12.7\%) with a large standard deviation of 10.8\%. The patent-level shareholder overlap (\textit{sol}) shows an average (median) value of 24.5\% (22.7\%) with a standard deviation of 17.2\%. The higher mean and standard deviation for the patent-level shareholder overlap is explained by the fact that firms with many patent filings tend to be both larger and feature a higher level of shareholder overlap. The institutional ownership (relative to the total number of shares outstanding of a firm) generally exhibits an upward time trend, from 29\% in 1992 to 61\% in 2007. The shareholder overlap statistics feature a similar upward time trend as well. Fixed time effects are included in all regressions to exclude that the shareholder overlap measure

\textsuperscript{13}The exclusion of privately held patents presumably creates a measurement error for \textit{SOL} unless shareholder overlap with such patent owning firms is zero. However, we conjecture that such shareholder overlap with privately held companies may indeed be generally negligible.
does not capture any parallel time trend in patent success. We provide the detailed definitions of all variables in the appendix.

5 Evidence on Patent Success

5.1 Baseline Specification

Our baseline regression specification follows Eq. (13) and Eq. (20). As some firms in our sample feature patents without any citations, we replace the term ln[CITES] with ln[1+CITES] in our main regression specification as follows:14

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

(22)

where the coefficient of interest is \( \beta_1 = (\frac{1}{t} + \gamma + \delta) \geq 0 \). More shareholder overlap with the cited upstream patents of a firm should boost its patent success as holdup problems are attenuated.

We estimate Eq. (22) over the period 1992–2007. The citation count \( CITES_{s,t} \) for patents filed 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 previous R&D investment \( ln(1+R&D \ Stock_{s,t-1}) \), a measure of relative capital intensity \( ln(1+K/L_{s,t-1}) \), and a firm size measure \( ln(1+Sales_{s,t-1}) \).15

In line with the existing literature, regressions without firm fixed effects control for the pre-sample mean citation count of a firm’s patent stock up to 1991 (See, e.g., Aghion, Van Reenen, and Zingales, 2013).

Table 2 presents the baseline regressions using the dependent variable \( ln(1+CITES) \). Columns 1–3 present the results for all firms and Columns 4–6 for firms in three R&D intensive sectors (pharmaceuticals, computer hardware, and telecommunications equipment).16 Columns 1 and 4 control for year fixed effects as well as industry fixed effects based on four-digit SIC codes, whereas

---

14 We note that all results remain qualitatively similar if we restrict the sample to firms with a strictly positive number of citations and use \( \ln[CITES] \) as the dependent variable.

15 See for example Gompers and Metrick (2001) and Hall et al. (2005).

16 We identify the three R&D intensive sectors following Bloom, Schakerman, and Vanreenen (2013). Specifically, they are 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).
Columns 2–3 and 5–6 control for year and firm fixed effects. We report robust t-statistics allowing for two-way clustering at the firm and year (or patent cohort) level.

The baseline regression shows that shareholder overlap SOL represents a statistically highly significant explanatory variable; the point estimate of 4.708 in Column 1 implies that an increase in shareholder overlap by one standard deviation (or 0.108) increases patent success in terms of overall log firm citation ($\ln[1+CITES]$) by 24% of its standard deviation of 2.084, suggesting that shareholder overlap has an economically large attenuation effect on patent success. Inclusion of firm fixed effects in Columns 2 limits the explanatory power of SOL to intertemporal variations in patent success within a firm; yet, the point estimate for the SOL coefficient remains large at 4.149. The three control variables seem largely redundant in the firm fixed-effect regression (reported in Column 2) because their exclusion in Column 3 yields a very similar regression coefficient of 4.266 for SOL.

Columns 4–6 repeat these regression for the three R&D intensive sectors. As expected, we find even stronger SOL effect in these sectors. The point estimate of 5.894 in Column 4 implies that increase of SOL by one standard deviation (or 0.106) increases the citation count by roughly 27.18% of its standard deviation of 2.299. The firm fixed effect regressions in Columns 5–6 identify only within firm variation in patent success, but shareholder overlap remains a statistically and economically highly significant covariate with or without inclusion of the other controls.

The control variables generally have the expected signs when firm fixed effects are excluded (Columns 1 and 4). A higher stock of cumulative R&D spending correlates positively with future patent success. Firm size measured by $\ln(1 + Sales)$ also 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. In the long run, it’s possible that these “control variables” are also be influenced by the patent process itself. For example, if shareholder overlap is positively related to patent success, the latter can also be reflected in larger Sales and a higher R&D Stock, thereby disqualifying both as exogenous conditioning control variables. In this case, their exclusion from the regression, as reported in Columns 3 and 6, represents the most reliable inference on the positive effect of shareholder overlap on patent success.
5.2 Intensive versus Extensive Margins

Shareholder overlap may affect the intensive and extensive margins differently. Moreover, separate specifications for them reveal different regression parameters. The specification for the intensive margin follows Eqs. (8) and (20), where the expectation term $E[cites(p)]$ is replaced by the empirical sample equivalent $\overline{cites}$. In the log transformation, we use $ln[1 + \overline{cites}]$ rather than $ln[cites]$ to include firms whose patents do not receive any citations. Specifically, we estimate the following equation:

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

(23)

where $\theta_1 = \gamma \delta > 0$ captures the positive effect of less patent holdup due to shareholder overlap. The parameter $\gamma$ measures the efficiency loss of patent holdup as opposed to the distributional loss through rent dissipation measured by $\delta$. 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 the upstream firms.

Table 3, Columns 1–3 summarize the effect of shareholder overlap on the intensive margin. Column 1 excludes firm fixed effects so that both cross- and within-firm variation in shareholder overlap is reflected in the point estimate of 2.423. The economic significance of the shareholder overlap measure for the intensive margin $ln[1 + \overline{cites}]$ is similar to that for the overall citation count $ln[1+CITES]$: An increase in shareholder overlap by one standard deviation (or 0.108) corresponds to an increase in the average citation count per patent by about 23% of its standard deviation. Inclusion of firm fixed effects in Column 2 restricts the identification of the shareholder overlap effect to intertemporal firm variation, but $SOL$ still yields an economically and statistically highly significant estimate. This finding is also robust to the exclusion of the control variables, as shown in Column 3.

The empirical specification for the extensive margin of patent production follows Eqs. (12) and (20) as

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

(24)

where $N$ corresponds to the $\overline{p}_L$ in our model description in Section 3.2. The model implied

$^{17}$Similar to Eq. (22) and Eq. (23), we use $ln[1+N]$ rather than $ln[N]$ in Eq. (24), but the results are qualitatively similar under both alternatives.
The regression results for the extensive margin are presented in Table 3, Columns 4–6. The point estimate of 1.576 again suggests strong economic significance for the shareholder overlap measure; a one-standard-deviation increase in SOL is associated with a 12% increase in the number of patents relative to its standard deviation of 1.435. The coefficient retains its economic and statistical significance in the specification with firm fixed effects in Columns 5. This finding is again robust to the exclusion of the three control variables, shown in Column 6.

Overall, the results suggest that shareholder overlap is strongly associated with both more citations for each granted patent (i.e., the intensive margin of patent success) and the number of granted patents (i.e., the extensive margin of patent production). Because both regression coefficients \( \theta_1 = \gamma \delta \) and \( \psi_1 = (1 + \gamma) \frac{1}{b} \delta \) are strictly positive, we conclude that both parameters \( \gamma \) and \( \delta \) are strictly positive. Overall, the result is consistent with the model presented in Section 3 that patent holdup not only redistributes rents \( (\delta > 0) \), but also compromises long-run patent success \( (\gamma > 0) \).

### 5.3 Shareholder Type and Overlap Concentration

In this subsection, we disaggregate the shareholder overlap into its components sorted by shareholder type. Ownership overlap by active shareholders should contribute more to the avoidance of cross-firm rent seeking than ownership overlap by passive shareholders. Following Agarwal, Jiang, Tang, and Yang (2013), we distinguish four types of institutional owners: (i) hedge fund (HF), (ii) bank and insurance companies (B&I) (iii) investment firms (IF) (including most mutual funds), and (iv) all other institutional investors such as pension funds, university endowments, etc.\(^{18}\) We can decompose the firm-level shareholder overlap in Eq. (19) by investor type as

\[
SOL_{s,t-1} = SOL_{HF_{s,t-1}} + SOL_{B&I_{s,t-1}} + SOL_{IF_{s,t-1}} + SOL_{other_{s,t-1}},
\]

and thus identify the attenuation effect of shareholder overlap separately for each investor category. Notwithstanding considerable heterogeneity of shareholder activism within each group, hedge funds can be expected to exercise (on average) the most effective shareholder power, whereas

---

\(^{18}\)We thank Prof. Vikas Agarwal for making the hedge fund list available to us and Prof. Brian Bushee for making the error-corrected investor classification data available at his research website at http://acct.wharton.upenn.edu/faculty/bushee/IIclass.html
the majority of pension funds (grouped under other) are likely to exert no influence on firm management and thus hold only formal shareholder power. Accordingly, we predict a larger holdup attenuation effect for SOL_HF compared to SOL_other.

The regression results reported in Table 4 confirm this intuition. Splitting shareholder overlap according to the four investor types in Columns 3–4 yields the highest coefficient for shareholder overlap of hedge funds (SOL_HF). Under firm fixed effects in Column 4, the coefficient for SOL_HF is 6.101, approximately 47% larger than the coefficient for SOL under the baseline regression reported in Column 2, suggesting that hedge funds are considerably more effective in exercising shareholder power than the average institutional owner. Shareholder overlap of bank and insurance companies (SOL_B&I) and investment firms (SOL_IF) produces somewhat lower coefficients, closer to the overall sample effect for SOL. By contrast, investors grouped under SOL_other (mostly pension funds and endowments) do not contribute to any holdup attenuation because their coefficient is not statistically significantly different from zero. This ranking of the SOL coefficient for different investor groups roughly corresponds to prior notions of investor activism expressed in the literature (see, e.g., Brav, Jiang, Tian, 2014; Solomon and Soltes, 2015).

Separately, we examine another hypothesis concerning the ownership concentration of shareholder overlap. If the downstream innovating firm and the upstream cited firms are jointly owned by a few relatively large shareholders, coordinated action might be easier to organize, and shareholders may have stronger incentives to resolve a potential holdup. To test this hypothesis, consider a downstream patent \( p \) filed by firm \( s \) in year \( t \) and a related upstream patent \( p_u \) owned by firm \( u \). Let \( i \in I_{p,p_u} \) denotes an overlapping investor who at the end of period \( t-1 \) 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-Hirschmann index (HHI) of shareholder overlap based on overlapping ownership shares \( w_i = \min[w_{i,s}, w_{i,u}] \) of all overlapping shareholders \( i \in I_{p,p_u} \). Then, we 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-Hirschmann index of ownership concentration of overlapping shareholders, defined as

\[
WHHI_{s,t-1} = \sum_{p=1}^{N_p} \sum_{u=1}^{N_u} w(p)w(p_u)HHI_{p,p_u,t-1},
\]

(26)
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 the end of year \( t - 1 \). \( WHHI \) describes how concentrated the overlapping ownership stakes are at the firm level and thus captures the coordination problem among the overlapping investors.

Table 4, Columns 5 and 6, include \( WHHI \) as a separate control variable. The estimated coefficient is statistically highly significant and positive, suggesting that a concentration of joint ownership shares by overlapping shareholders positively correlates with patent success beyond the shareholder overlap \( SOL \) itself. The coefficient estimate of 1.32 in Column 5 implies that an increase in the ownership concentration of shareholder overlap by one standard deviation (or 0.164) generates the same effect on patent success as raising \( SOL \) by 31% relative to its mean \((= [0.164 \times 1.32] / [4.915 \times 0.141])\). This suggests that coordination problems among dispersed overlapping institutional investors represent an important impediment to the exercise of effective shareholder power.

### 5.4 Endogeneity Concerns and Patent-Level Regression

The firm-level regressions in the previous section control for a variety of observable firm characteristics as well as firm fixed effects. Yet, time varying (unobservable) influences on both patent success and shareholder overlap may still pose a concern for our inference.\(^{19}\)

In this section, we include not only firm and year fixed effects separately, but also their interaction \( \epsilon_{s,t} \). Therefore, identification of the holdup attenuation effect on patent success relies entirely on the comparison of different patents filed by the same firm in the same year. Different patent filings by the same firm may list 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 the patent-level shareholder overlap \( sol_{p,t-1} \) in the following regression specification

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

where \( cites_{p,t} \) denotes the future citation count of patent \( p \) filed in year \( t \). Similar to the firm-level regressions, all independent variables lag the dependent variable by one year.

\(^{19}\)For example, media coverage may boost a firm’s citation count and simultaneously trigger stock purchases by investors with an investment bias towards technology stocks, thereby increasing the firm’s shareholder overlap measure.
The patent-level citation success $cites_{p,t}$ can only capture the intensive margin of patent production similar to Table 3, Columns 1-3. But in contrast to the firm-level regressions in Table 3, the patent-level data sample firms according to the number of patents filed. The estimated holdup attenuation effect therefore features a strong selection bias toward those firms with many patents. In addition, more than 25% of firms with only one patent filed per year are discarded from the patent level regression.

The result in Table 5 shows that within-firm variation of the patent success is statistically significantly related to the patent-specific shareholder overlap $sol$. A coefficient of 0.226 in Column 1 implies that an increase of shareholder overlap $sol$ by one standard deviation (0.168) is related to an increase in the patent-level citation count by 2.8% ($= 0.226 \times 0.168 / 1.361$). This modest economic effect represents the hold-up attenuation effect on the intensive margin of the patent intensive firms. The 3 percent most patent intensive firms account for roughly 50% of all patent filing. Columns 2 and 3 present separate regressions for the 50% patents attributable to the 97% least and 3% most patent intensive firms, respectively. Column 4 present results only for the 3 industries with the most patent filings. In each case the coefficient for $sol$ is statistically highly significant at the 1 percent level. Patent success within a firm is therefore also correlated to the patent specific shareholder overlap $sol$ which differentiates different patents within the same firm and filing year.

6 Related Evidence

6.1 R&D Expenditure

So far the analysis is focused on patent success as the main measure of the holdup attenuation effect of ownership overlap; yet, the model also predicts a positive effect of shareholder overlap on R&D investment. To test this prediction, we undertake a linear regression

$$
\ln[1 + R&D \ Exp_{s,t}] = \kappa_0 + \kappa_1 SOL_{s,t-1} + \kappa_2 Controls_{s,t-1} + \epsilon_s + \mu_t + \eta_{s,t}, \quad (28)
$$

where Eqs. (14) and (20) predict a positive a coefficient $\kappa_1 = (1 + \gamma)(1 + 1/b)\delta > 0$. We include the same control variables as in the previous regressions with the exception of $\ln(1 + R&D \ Stock)$, which summarizes past R&D expenditure.
Table 6 reports the regression results. The effect of shareholder overlap is statistically and economically highly significant in specifications without firm fixed effects (Column 1), with firm fixed effects and control variables (Column 2), and with firm fixed effects but without control variables (Column 3). For example, an increase in shareholder overlap by one standard deviation (or 0.108) in the last specification (Column 3) increases R&D expenditure by approximately 13% (= 0.108 × 1.222). The holdup attenuation effect of shareholder overlap on R&D investment is therefore economically important.

We also note that the advertisement channel of reciprocal patent citations evoked in Section 5.4 cannot easily account for the positive effect of shareholder overlap on R&D expenditure—Although inflated patent citations might help boost (short-term) equity valuations, it is unclear why firm management would increase R&D investment in parallel to its own citation manipulation.

6.2 Patent Litigation

[Section to be added]

6.3 Dynamic Adjustment of Shareholder Overlap

The optimal ownership structure under patent holdup should dynamically adjust towards a more efficient ownership structure that combines ownership in complementary assets. Under private information about future patent hold-up, investors have an incentive to achieve this joint ownership through shareholder overlap—thus internalizing the hold-up problem. We might therefore expect that the shareholder overlap between the downstream and upstream firms increases prior to public disclosure of the patent filings. Moreover, this dynamic adjustment in shareholder overlap should be strongest for activist investors such as hedge funds.

For each year cohort of patents filed between 1996 and 2002, we measure 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 \( \overline{SO\ell}(t, s) \) represent the average shareholder overlap measured based on ownership data at the end of year \( t + s \), where \( s = -4, -3, ..., 4 \). For example, \( \overline{SO\ell}(t, -3) \) denotes shareholder overlap between the downstream and upstream patent owning firms measured based on ownership in year \( t - 3 \) for all patents filed in year \( t \). We then calculate the growth rate of shareholder overlap and average over all filing cohorts \( t = 1996, 1996, ..., 2002 \).
to obtain the shareholder overlap growth at lag $s$ given by

$$\text{Overlap Growth}(s) = \sum_{t=1996}^{2002} \frac{\text{SOL}(t, s) - \text{SOL}(t, s - 1)}{\text{SOL}(t, s - 1)}$$

for $s = -3, -2, -1, \ldots, 3, 4$.

Figure 1 graphs this growth rate of shareholder overlap separately for all shareholders and for hedge funds representing more activist investors. As patent filings are publicly disclosed one year after the patent filing date, we mark the date $t + 1$ with a vertical line. Public disclosure of patent filings implies that the share price of both the up- and downstream firm should reflect any firm value change related to patent hold-up. Thus, the internalization of the hold-up through ownership overlap can only be achieved through ownership change prior to $t + 1$. The cumulative growth of shareholder overlap over the three year period from $t - 2$ to $t + 1$ is approximately 11% for hedge funds and only 3% for the overall shareholder overlap. After disclosure year $t + 1$, the shareholder overlap growth tends to decreases. Overall, the economic magnitude of the dynamic overlap adjustment appears small. Trading on private information about future hold-up rents might also expose the investors to insider trading charges and this may limit its scope. The static effect of increased shareholder overlap due to the long-run growth of institutional ownership appear to be qualitatively more important for the attenuation of patent hold-up.

7 Robustness

7.1 Alternative Hypotheses

This section discusses two alternative determinants of patent success and examines their empirical importance relative to shareholder overlap ($\text{SOL}$). Aghion, Van Reenen, and Zingales (2013) argue that R&D investments have a long time horizon, and a high share of institutional investors allows management to focus on the long-term return to investment. Following their specification, we measure the share of institutional ownership ($\text{IO}$) as the relevant proxy for investor patience. As institutional ownership also correlates with our shareholder overlap measure, it could potentially account for the firm-level evidence presented in Sections 5.1–5.3.

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

$$ IIF_{i,s,t} = \sum_{s' \neq s} x_{i,s',t} FIF_{s',t} $$

where $x_{i,s',t}$ represents the value weight of firm $s'$ in the portfolio of institutional investor $i$ at the end of year $t$. For any individual institutional shareholder primarily investing in innovative firms, the $IIF$ value should be high. In the third step, the shareholder innovation focus ($SIF$) for firm $s$ is defined as the value-weighted average of investor innovation focus for all shareholders $i$ in firm $s$,

$$ SIF_{s,t} = \sum_i w_{i,s,t} IIF_{i,s,t} $$

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

Table 7 presents the regression results for the two alternative hypotheses. Column 1 reproduces the benchmark regression reported in Column 1 of Table 2. Column 2 of the table replaces shareholder overlap $SOL_{s,t-1}$ with institutional ownership $IO_{s,t-1}$ as the key explanatory variable. Unlike Aghion, Van Reenen and Zingales (2013), we do not find that institutional ownership features a statistically significant correlation with the citation success of a firm. Including both shareholder overlap and institutional ownership in Column 3, we even find a negative effect of institutional ownership on patent success, whereas shareholder overlap retains its high positive level of economic and statistical significance. Table 7, Column 4 includes shareholder innovation focus $SIF_{s,t}$ as the third explanatory variable for patent success. We find that the general innovation
focus of a firm’s shareholders does not foster patent success of the respective firm after accounting for the SOL effect.

To verify that these results are robust, Columns 5–6 report analogous regressions based on the same (smaller) patent sample as that used by Aghion, Van Reenen, and Zingales (2013). We also mimic their specification by using $\ln[CITES_{s,t}]$ as the dependent variable and apply the same control variables. Column 5 reproduces their benchmark regression (reported in Table 1, Column 2 of their paper) with the same statistically significant positive coefficient of 0.546 for institutional ownership (IO). When we augment the regression with shareholder overlap SOL as an additional explanatory variable in Column 6, shareholder overlap remains statistically significant but not institutional ownership. The explanatory power of institutional ownership for patent success appears to arise from its correlation (0.385) with the shareholder overlap variable.

7.2 Measurement Issues

We subject a variety of measurement choices to a robustness analysis. The first robustness test concerns the truncation nature of patent citations. 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. The truncated nature of the dependent variable may imply a time dependent measurement error as the inferred patent success of later cohorts of patents is based on a shorter time span. Lerner, Sorensen and Stromberg (2011) propose a shorter, but more homogeneously truncated citations count over a three-year period immediately after a patent is granted. Following their approach, we define the relative citation count of a patent as

$$cites_{p,t}^{rel} = \frac{cites_{p,t}^{3y}}{\frac{1}{N_k} \sum_{p \in k} cites_{p,t}^{3y}}, (31)$$

where patent success is captured by 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 USPTO technol-

---

20 Aghion, Van Reenen, and Zingales (2013) use $\ln[CITES_{s,t}]$ as the dependent variable in their benchmark regression in Table 1, Column (2). They include $IO_{s,t}$, $\ln(R&D \ Stock_{s,t})$, $\ln(K/L_{s,t})$, and $\ln(Sales_{s,t})$ as the regressors. We use the exact same set of variables in our regressions reported in Columns 5–6. Their dataset is available at: https://www.aeaweb.org/articles.php?doi=10.1257/aer.103.1.277.

21 A detailed documentation on these robustness test is available as Web Appendix to this paper on our website www.haraldhau.com.
ogy class $k$. The firm-level measure $CITES_{rel_s,t}$ follows as the sum of $cites_{rel_{p,t}}$ over all $N_s$ successful patents filed by firm $s$ in year $t$. The importance weights $w(p)$ and $w(p_{u})$ in the calculation of shareholder overlap $SOL$ are also based on $cites_{rel_{p,t}}$. The modified shareholder overlap variable is denoted by $SOL_{rel}$. Notwithstanding these variable modifications, we still find qualitatively similar results for the holdup attenuation effect of shareholder overlap.

The second robustness test concerns the measurement of shareholder overlap itself. As patent projects might be initiated several years before the application year $t$, ownership overlap might also be accumulated earlier than year $t − 1$ (which is the measurement year for shareholder overlap used in our main analysis). We find that shareholder overlap measured based on equity stakes at the end of years $t − 2, t − 3$, and $t − 4$ still produces statistically highly significant point estimates for $SOL$, albeit with a smaller economic significance. This finding is consistent with a buildup of shareholder overlap prior to the filing of the downstream patent in year $t$.

Third, our benchmark measure of firm-level shareholder overlap $SOL$ uses importance weights based on citation count of patents. As an alternative measure, we replace the log citations count $\ln[1+cites_s(p)]$ in Eq. (21) with a rank measure of future citations $rank(p)$ to obtain a new shareholder overlap measure $SOL_{rank}$. This alternative measure of shareholder overlap generates very similar regression coefficients for all reported specifications. As another alternative measure, we suspend the importance weights altogether and aggregate all combinations of downstream and upstream patents under equal weights. The resulted shareholder overlap variable $SOL_{equal}$ is again statistically highly significant in the reported regression specifications, albeit at a slightly weaker level of economic significance. Overall, the use of patent importance weights based on either (i) log citation counts or (ii) rank statistics of citation counts slightly improves the explanatory power of shareholder overlap for patent success.

Fourth, we repeat Columns 1–3 of Table 2 but use $ln(CITES)$ as an alternative dependent variable. We discard all firm-year observations for which successful patent applications exist but these patents do not receive any citation (i.e., $CITES = 0$). The economic significance of $SOL$ remains high in this smaller sample; the point estimate suggests that an increase of $SOL$ by one standard deviation increases the citation count $CITES$ by roughly 20% in the firm fixed effect regression.
8 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. This should particularly be the case if such joint shareholder ownership comes from activist investors who exercise power over both firms and therefore mitigate the holdup problem in corporate innovation processes.

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 both at the firm level and the patent level; it positively correlates with both the intensive and extensive margin 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 year cohort filed by the same firm perform differently depending on their respective (patent-level) shareholder overlap. We also find that shareholder overlap coming from more active investors such as hedge funds tend to contribute more to the holdup attenuation—suggesting that the ‘extended boundary’ of the innovating firm depends also on the type of institutional shareholders. Furthermore, the ownership concentration of shareholder overlap matters independently of its level, suggesting coordination and free-rider problems among a large group of overlapping shareholders.

Our paper also documents economically and statistically significant effects of shareholder overlap on other variables of the patent process in accordance with the holdup attenuation hypothesis: R&D expenditure is positively correlated with the firm-level shareholder overlap, whereas the probability of patent litigation decreases in shareholder overlap.

Finally, we provide a dynamic perspective on efficient equity ownership: Ownership overlap
increases modestly prior to public disclosure of patent files for activist shareholders like hedge funds. Yet this dynamic integration of ownership is likely to play only a minor role for the attenuation of patent hold-ups. Quantitatively more important is the long-run increase in institutional ownership with its significant intertemporal increase in shareholder overlap.
References


Reported are the summary statistics for all regression variables in the sample period 1992–2007. Dependent firm level variables are (i) \( \text{CITES}_{s,t} \) as number of future citations received by the cohort of patents successfully filed by firm \( s \) in year \( t \), (ii) \( N_s,t \) as the number of successfully filed patents, (iii) \( \text{cites}_{s,t} \) as the average future citations per patent for the cohort of patents filed by firm \( s \) in year \( t \), and (iv) \( \text{R&D Exp} \) as R&D expenditure for firm \( s \) in year \( t \). At the patent level, (v) \( \text{cites}_{p,t} \) denotes the total number of future citations (exclusive of self-citations) received by patent \( p \) successfully filed in year \( t \), whereas (vi) \( \text{cites}^{F1}_{p,t} \) and (vii) \( \text{cites}^{F2}_{p,t} \) also exclude citations coming from all firms quoted by patent \( p \) (for filter \( F1 \)) and all firms that patent \( p \)'s innovator has ever quoted previously (for filter \( F2 \)), respectively. The explanatory variables \( \text{SOL}_{s,t-1} \) and \( \text{sol}_{p,t-1} \), refer to the shareholder overlap for firm \( s \) or a patent \( p \), respectively. We separate shareholder overlap by investor types which are hedge funds (\( \text{SOL}_{-HF} \)), bank and insurance companies (\( \text{SOL}_{-BkI} \)), investment firms (\( \text{SOL}_{-IF} \)), and others (\( \text{SOL}_{-Other} \)). \( \text{IO}_{s,t-1} \) is the aggregate institutional ownership of firm \( s \) as of the end of year \( t-1 \). The shareholder innovation focus \( SIF_{s,t-1} \) is defined as the investment bias of a firm’s shareholders towards firms with a large share of patents. \( \text{WHHI}^{SOL}_{s,t-1} \) represents the weighted HHI of shareholder overlap concentration for firm \( s \) in year \( t-1 \). The control variables include the (log of) lagged cumulative R&D investment, \( \text{ln}(1+\text{R&D Stock}_{s,t-1}) \); lagged capital to labor ratio, \( \text{ln}(1+\text{K/L}_{s,t-1}) \); and lagged sales, \( \text{ln}(1+\text{Sales}_{s,t-1}) \). The variable definitions are described in detail in the appendix.

### Table 1: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Median</th>
<th>STD</th>
<th>Skewness</th>
<th>Min.</th>
<th>P10</th>
<th>P90</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{ln}(1+\text{CITES}))</td>
<td>17,674</td>
<td>4.102</td>
<td>4.085</td>
<td>2.035</td>
<td>0.066</td>
<td>0.000</td>
<td>1.560</td>
<td>6.698</td>
<td>11.640</td>
</tr>
<tr>
<td>(\text{ln}(1+N))</td>
<td>17,674</td>
<td>2.056</td>
<td>1.609</td>
<td>1.346</td>
<td>1.284</td>
<td>0.693</td>
<td>0.693</td>
<td>4.025</td>
<td>8.395</td>
</tr>
<tr>
<td>(\text{ln}(1+\text{cites}))</td>
<td>17,674</td>
<td>2.428</td>
<td>2.495</td>
<td>1.122</td>
<td>-0.183</td>
<td>0.000</td>
<td>0.975</td>
<td>3.796</td>
<td>6.278</td>
</tr>
<tr>
<td>(\text{ln}(1+\text{R&amp;D Exp}))</td>
<td>17,083</td>
<td>2.779</td>
<td>2.745</td>
<td>1.883</td>
<td>0.426</td>
<td>0.000</td>
<td>0.000</td>
<td>5.236</td>
<td>9.408</td>
</tr>
<tr>
<td>(\text{ln}(1+\text{cites}))</td>
<td>489,895</td>
<td>1.954</td>
<td>2.020</td>
<td>1.361</td>
<td>0.085</td>
<td>0.000</td>
<td>0.000</td>
<td>3.707</td>
<td>6.969</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Median</th>
<th>STD</th>
<th>Skewness</th>
<th>Min.</th>
<th>P10</th>
<th>P90</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{SOL}_{-HF})</td>
<td>17,674</td>
<td>0.147</td>
<td>0.135</td>
<td>0.108</td>
<td>0.603</td>
<td>0.000</td>
<td>0.000</td>
<td>0.297</td>
<td>0.700</td>
</tr>
<tr>
<td>(\text{SOL}_{-BkI})</td>
<td>17,674</td>
<td>0.008</td>
<td>0.004</td>
<td>0.011</td>
<td>3.283</td>
<td>0.000</td>
<td>0.000</td>
<td>0.021</td>
<td>0.229</td>
</tr>
<tr>
<td>(\text{SOL}_{-IF})</td>
<td>17,674</td>
<td>0.068</td>
<td>0.063</td>
<td>0.052</td>
<td>0.659</td>
<td>0.000</td>
<td>0.000</td>
<td>0.139</td>
<td>0.357</td>
</tr>
<tr>
<td>(\text{SOL}_{-Other})</td>
<td>17,674</td>
<td>0.056</td>
<td>0.049</td>
<td>0.047</td>
<td>1.167</td>
<td>0.000</td>
<td>0.000</td>
<td>0.119</td>
<td>0.673</td>
</tr>
<tr>
<td>(\text{sol})</td>
<td>489,895</td>
<td>0.296</td>
<td>0.295</td>
<td>0.168</td>
<td>0.177</td>
<td>0.000</td>
<td>0.067</td>
<td>0.518</td>
<td>0.851</td>
</tr>
<tr>
<td>(\text{SIF})</td>
<td>17,674</td>
<td>0.201</td>
<td>0.200</td>
<td>0.057</td>
<td>1.735</td>
<td>0.000</td>
<td>0.136</td>
<td>0.269</td>
<td>1.810</td>
</tr>
<tr>
<td>(\text{IO})</td>
<td>17,674</td>
<td>0.485</td>
<td>0.507</td>
<td>0.265</td>
<td>-0.099</td>
<td>0.000</td>
<td>0.102</td>
<td>0.825</td>
<td>1.000</td>
</tr>
<tr>
<td>(\text{WHHI}^{SOL}_{s,t-1})</td>
<td>17,674</td>
<td>0.150</td>
<td>0.098</td>
<td>0.156</td>
<td>2.222</td>
<td>0.000</td>
<td>0.000</td>
<td>0.351</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Median</th>
<th>STD</th>
<th>Skewness</th>
<th>Min.</th>
<th>P10</th>
<th>P90</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{ln}(1+\text{R&amp;D Stock}))</td>
<td>17,674</td>
<td>3.889</td>
<td>4.021</td>
<td>2.217</td>
<td>0.008</td>
<td>0.000</td>
<td>0.000</td>
<td>6.649</td>
<td>10.714</td>
</tr>
<tr>
<td>(\text{ln}(1+\text{K/L}))</td>
<td>17,674</td>
<td>4.431</td>
<td>4.344</td>
<td>0.890</td>
<td>0.606</td>
<td>0.000</td>
<td>3.427</td>
<td>5.537</td>
<td>10.296</td>
</tr>
<tr>
<td>(\text{ln}(1+\text{Sales}))</td>
<td>17,674</td>
<td>5.523</td>
<td>5.526</td>
<td>2.464</td>
<td>-0.057</td>
<td>0.000</td>
<td>2.275</td>
<td>8.738</td>
<td>12.722</td>
</tr>
</tbody>
</table>
Table 2: Baseline Regressions

Reported are firm level OLS regressions of patent success (measured as the (log) future citation count, \(\ln(1 + \text{CITES}_{s,t})\) for all patents filed by firm \(s\) in year \(t\)) on the lagged shareholder overlap, \(\text{SOL}_{s,t-1}\), for the sample period 1992 – 2007. Column 1-3 report full sample results whereas column 4-6 results of subsample based on top 3 most innovative industries. Shareholder overlap measures the average shareholder ownership overlap between the innovating firm and other firms owning the precursory complementary patents. The control variables include the (log of) lagged cumulative R&D investment, \(\ln(1 + \text{R&D Stock}_{s,t-1})\); lagged capital to labor ratio, \(\ln(1 + K/L_{s,t-1})\); and lagged sales, \(\ln(1 + \text{Sales}_{s,t-1})\). Pre-sample mean cites count denotes the average future cites count the firm receives per pre-sample year. Industry fixed effects are based on four-digit SIC codes. All regressions report robust \(t\)-statistics clustered at firm and year levels in brackets. We denote by *, **, and *** the statistical significance at the 10%, 5%, and 1% level, respectively. The variable definitions are described in more detail in the appendix.

<table>
<thead>
<tr>
<th>Dependent Variables:</th>
<th>(\ln(1 + \text{CITES}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Sample</td>
</tr>
<tr>
<td>(\text{SOL})</td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>5.461**</td>
</tr>
<tr>
<td></td>
<td>[34.04]</td>
</tr>
<tr>
<td>Controls:</td>
<td></td>
</tr>
<tr>
<td>(\ln(1 + \text{R&amp;D Stock}))</td>
<td>0.325**</td>
</tr>
<tr>
<td></td>
<td>[26.24]</td>
</tr>
<tr>
<td>(\ln(1 + K/L))</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>[0.44]</td>
</tr>
<tr>
<td>(\ln(1 + \text{Sales}))</td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td>[-1.10]</td>
</tr>
<tr>
<td>Pre-sample cites</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
</tr>
<tr>
<td>Industry FE</td>
<td>YES</td>
</tr>
<tr>
<td>Firm FE</td>
<td>NO</td>
</tr>
<tr>
<td>Obs.</td>
<td>17,674</td>
</tr>
<tr>
<td>Adj. R^2</td>
<td>0.581</td>
</tr>
</tbody>
</table>
Table 3: Intensive versus Extensive Margin

Reported are OLS regressions for (i) the intensive margin, $\ln(1 + \overline{\text{cites}_{s,t}^s})$, and (ii) 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. We denote by $N_{s,t}$ the number of successful patents filed by firm $s$ in year $t$, and by $\overline{\text{cites}_{s,t}^s}$ the average future citations per patent for the cohort of patents successfully filed by firm $s$ in year $t$. Shareholder overlap measures the average shareholder ownership overlap between the innovating firm and other firms owning the precursory complementary patents. The control variables include the (log of) lagged cumulative R&D investment, $\ln(1 + R&D\text{ Stocks}_{s,t-1})$; lagged capital to labor ratio, $\ln(1 + K/L_{s,t-1})$; and lagged sales, $\ln(1 + Sales_{s,t-1})$. Pre-sample mean cites count denotes the average future cites count the firm receives per pre-sample year. Industry fixed effects is based on four-digit SIC codes. All regressions report robust $t$-statistics clustered at firm and year levels in brackets. We denote by *, **, and *** statistical significance at the 10%, 5%, and 1% level, respectively. The variable definitions are described in more detail in the appendix.

<table>
<thead>
<tr>
<th>Dependent Variables:</th>
<th>$\ln(1 + \overline{\text{cites}})$</th>
<th>$\ln(1 + N)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$SOL$</td>
<td>2.652***</td>
<td>2.322***</td>
</tr>
<tr>
<td></td>
<td>[28.33]</td>
<td>[22.64]</td>
</tr>
<tr>
<td>Controls:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(1 + R&amp;D\text{ Stock})$</td>
<td>-0.037***</td>
<td>-0.071***</td>
</tr>
<tr>
<td></td>
<td>[-4.79]</td>
<td>[-3.85]</td>
</tr>
<tr>
<td>$\ln(1 + K/L)$</td>
<td>-0.029**</td>
<td>-0.095***</td>
</tr>
<tr>
<td></td>
<td>[-2.37]</td>
<td>[-4.49]</td>
</tr>
<tr>
<td>$\ln(1 + Sales)$</td>
<td>-0.065***</td>
<td>-0.098***</td>
</tr>
<tr>
<td></td>
<td>[-10.56]</td>
<td>[-7.10]</td>
</tr>
<tr>
<td>Pre-sample cites</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Industry FE</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Firm FE</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Obs</td>
<td>17,674</td>
<td>17,674</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.472</td>
<td>0.626</td>
</tr>
</tbody>
</table>
Table 4: Hold-Up Resolution by Investor Types and Concentration of Shareholder Overlap

The baseline regression in Table 2 are repeated for a split of shareholder overlap (SOL_{s,t-1}) into four components which represent hedge funds (SOL_HF), bank and insurance companies (SOL_B&I), investment firms (SOL_IF), and others (SOL_Other) respectively. The regressions in Column 5-6 expand the baseline regressions by a direct measure of Herfindahl-Hirschman index of shareholder overlap, WHHI_{s,t-1}^{SOL}, instead of a variable decomposition. The control variables include the (log of) lagged cumulative R&D investment, ln(1 + R&D Stock_{s,t-1}); lagged capital to labor ratio, ln(1 + K/L_{s,t-1}); and lagged sales, ln(1 + Sales_{s,t-1}). Pre-sample mean cites count denotes the average future cites count the firm receives per pre-sample year. All regressions report robust t-statistics clustered at firm and year levels in brackets. The last row of the table reports p-values for the null hypothesis that the estimated regression coefficients are the same for SOL_HF, SOL_B&I, SOL_IF and SOL_Other. We denote by *, **, and *** the statistical significance at the 10%, 5%, and 1% level, respectively. The variable definitions are described in more detail in the appendix.

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>ln(1 + CITES)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SOL</td>
<td>5.461***</td>
<td>4.171***</td>
<td>5.374***</td>
<td>3.987***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[34.04]</td>
<td>[26.17]</td>
<td>[33.79]</td>
<td>[25.20]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOL_HF</td>
<td>8.378***</td>
<td>5.754***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[6.07]</td>
<td>[4.86]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOL_B&amp;I</td>
<td>6.485***</td>
<td>5.362***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[16.84]</td>
<td>[14.07]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOL_IF</td>
<td>5.059***</td>
<td>3.315***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[12.53]</td>
<td>[9.61]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOL_Other</td>
<td>0.979</td>
<td>1.400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.88]</td>
<td>[1.25]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHHI_{s,t-1}^{SOL}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.289***</td>
<td>1.095***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[15.80]</td>
<td>[13.28]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Controls:

| ln(1 + R&D Stock) | 0.325***       | 0.154*** | 0.326*** | 0.154*** | 0.347*** | 0.172*** |
|                  | [26.24]        | [5.32]   | [26.35]  | [5.28]   | [28.10]  | [5.95]   |
| ln(1 + K/L)      | 0.008          | -0.135*** | 0.009   | -0.135*** | 0.014   | -0.131*** |
|                  | [0.44]         | [-4.38]  | [0.46]   | [-4.37]  | [0.75]   | [-4.28]  |
| ln(1 + Sales)    | -0.010         | -0.015   | -0.008  | -0.015   | 0.019** | 0.014    |
|                  | [-1.10]        | [-0.76]  | [-0.79]  | [-0.75]  | [2.00]   | [0.69]   |

Pre-sample cites YES NO YES NO YES NO Year FE YES YES YES YES YES YES Industry FE YES NO YES NO YES NO Firm FE NO YES NO YES YES YES Obs 17,674 17,674 17,674 17,674 17,674 17,674 Adj. R^2 0.581 0.745 0.582 0.746 0.588 0.749 p-value 0.000 0.001
This table presents the correlation between patent success measured at the patent level and the lagged shareholder overlap for the sample period 1992 – 2007. Patent success is proxied by $ln(1 + cites_{p,t})$ as the (log) future citation count received by a patent $p$ filed in year $t$. Columns 1 reports the full sample. Columns 2-3 feature the subsample of patents from firms with smaller patent filing count that covers bottom 50% and from firms with larger patent filing count that covers top 50% of patents. Column 4 concerns the 3 industries with the most R&D expenditures. We drop firm-year where only one patent is filed since we have controlled for firm fixed effects. All regressions report robust $t$-statistics clustered at firm and year levels in brackets. We denote by *, **, and *** the statistical significance at the 10%, 5%, and 1% level, respectively. The variable definitions are described in more detail in the appendix.

### Table 5: Patent-Level Regressions

<table>
<thead>
<tr>
<th>Dependent Variables:</th>
<th>$ln(1 + cites)$</th>
<th>$ln(1 + cites)$</th>
<th>$ln(1 + cites)$</th>
<th>$ln(1 + cites)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Sample</td>
<td>Sort by Firm Patent Intensity</td>
<td>Top 3 industries</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>$sol$</td>
<td>0.226***</td>
<td>0.122***</td>
<td>0.286***</td>
<td>0.244***</td>
</tr>
<tr>
<td></td>
<td>[11.77]</td>
<td>[4.57]</td>
<td>[10.87]</td>
<td>[7.75]</td>
</tr>
<tr>
<td>Tech. FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year $\times$ Firm FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Obs</td>
<td>485,114</td>
<td>242,796</td>
<td>242,318</td>
<td>196,051</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.329</td>
<td>0.358</td>
<td>0.298</td>
<td>0.316</td>
</tr>
</tbody>
</table>
Table 6: R&D Expenditure and Shareholder Overlap

Reported are OLS regressions of R&D expenditure on the lagged shareholder overlap for the sample period 1992–2007. R&D expenditure is measured for every firm-year \((s, t)\) and \(SOL_{s,t-1}\) represents firm-level shareholder overlap with all cited firms in the successful patent applications of firm \(s\) in year \(t - 1\). The control variables include lagged capital to labor ratio, \(ln(1+K/L_{s,t-1})\); and lagged sales, \(ln(1+Sales_{s,t-1})\). Industry fixed effects is based on four-digit SIC codes. All regressions report robust \(t\)-statistics clustered at firm and year levels in brackets. We denote by *, **, and *** the statistical significance at the 10%, 5%, and 1% level, respectively. The variable definitions are described in more detail in the appendix.

<table>
<thead>
<tr>
<th>Dependent Variable: (ln(1 + R&amp;D\ Exp))</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOL</td>
<td>3.783***</td>
<td>0.767***</td>
<td>1.244***</td>
</tr>
<tr>
<td></td>
<td>[30.01]</td>
<td>[10.22]</td>
<td>[15.46]</td>
</tr>
<tr>
<td>Controls: (ln(1 + K/L))</td>
<td>0.198***</td>
<td>-0.040**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[11.95]</td>
<td>[-2.30]</td>
<td></td>
</tr>
<tr>
<td>(ln(1 + Sales))</td>
<td>0.427***</td>
<td>0.333***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[59.18]</td>
<td>[26.58]</td>
<td></td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Industry FE</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Firm FE</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Obs</td>
<td>17,083</td>
<td>17,083</td>
<td>17,083</td>
</tr>
<tr>
<td>Adj. R(^2)</td>
<td>0.666</td>
<td>0.933</td>
<td>0.925</td>
</tr>
</tbody>
</table>
Table 7: Alternative Explanatory Variables

We compare three potential determinants of innovation success, namely (i) *shareholder overlap* ($SOL_{s,t-1}$) between an innovating firm and upstream firms owning complementary patents as a proxy for a attenuation of a patent hold-up problem, (ii) *shareholder innovation focus* ($SIF_{s,t-1}$) as a proxy for a firm’s shareholders focus on research intensive portfolio investments and (iii) *institutional ownership* ($IO_{s,t-1}$) as advocated by Aghion, Van Reenen and Zingales (2013) as a proxy for investor patience. Columns 1–4 use full sample period 1992 – 2007, and Columns 5–6 use the sample of Aghion, Van Reenen and Zingales (2013), which spans the shorter period 1991 to 2002. The dependent variable $ln(1 + CITEST_{s,t})$ is the (log) number of total future citations received by the cohort of patents successfully filed by firm $s$ in year $t$. The first four regressions use the same control variables as in the previous tables. The last two regressions use the same dependent variable, $ln(CITEST_{s,t})$, and control variables as in Aghion, Van Reenen and Zingales (2013). Robust $t$-statistics are reported in brackets. We denote by *, **, and *** statistical significance at the 10%, 5%, and 1% level, respectively. The variable definitions are described in more detail in the appendix.

<table>
<thead>
<tr>
<th>Dependent Variables:</th>
<th>$ln(1 + CITEST)$ Full Sample</th>
<th>$ln(1 + CITEST)$ ARZ Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$SOL$</td>
<td>5.461***</td>
<td>5.517***</td>
</tr>
<tr>
<td></td>
<td>[34.04]</td>
<td>[34.15]</td>
</tr>
<tr>
<td>$IO$</td>
<td>0.013</td>
<td>-0.214***</td>
</tr>
<tr>
<td></td>
<td>[0.23]</td>
<td>[-3.89]</td>
</tr>
<tr>
<td>$SIF$</td>
<td>0.853***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[3.41]</td>
<td></td>
</tr>
<tr>
<td>Controls:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$ln(1 + R&amp;D Stock)$</td>
<td>0.325***</td>
<td>0.421***</td>
</tr>
<tr>
<td></td>
<td>[26.24]</td>
<td>[32.98]</td>
</tr>
<tr>
<td>$ln(1 + K/L)$</td>
<td>0.008</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>[0.44]</td>
<td>[1.21]</td>
</tr>
<tr>
<td>$ln(1 + Sales)$</td>
<td>-0.010</td>
<td>0.086***</td>
</tr>
<tr>
<td></td>
<td>[-1.10]</td>
<td>[8.30]</td>
</tr>
<tr>
<td>Pre-sample cites</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Industry FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Firm FE</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Obs</td>
<td>17.674</td>
<td>17.674</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.581</td>
<td>0.540</td>
</tr>
</tbody>
</table>

42
Figure 1: Plotted are the growth rate of average shareholder overlap for all institutional investor types (blue) and hedge funds (red) measured based on shareholding at the end of year $t+k$, where $k = -3, -2, ..., 4$. 