Speech Is Silver, but Silence Is Golden: Information Suppression and the Promotion of Innovation*

Gaurav Kankanhalli† Alan Kwan‡ Kenneth Merkley§

February 21, 2019

Abstract

Per Arrow (1962), an innovation is difficult to finance as mere disclosure reduces its value. We study strategic nondisclosure in a novel setting of mandatory filings of IP licensing deals made by public firms, which contain major redactions. We predict that nondisclosure conceals an innovation from rivals, signaling its value. Consistent with our prediction, redacting firms are received well by capital markets relative to disclosing firms. We also find that such firms innovate more in the future. A natural experiment and cross-sectional analyses show lower benefits to redaction yield reduced signaling value. Our results suggest public markets can balance innovation and disclosure; however, information frictions are only partly resolved.

Keywords: Innovation, Disclosure, Redaction, Licensing, Patents, R&D
JEL codes: G30, G38, O30, O32, O34

*We thank Ednaldo Silva and Bartosz Niejadlik of RoyaltyStat LLC (more information available at https://www.royaltystat.com/) for their generous assistance. We are grateful to Sumit Agarwal, Tania Babina (discussant), Murillo Campello, Thomas Hellmann, Po-Hsuan Hsu, Andrew Karolyi, Jiro Kondo, Yong Suk Lee (discussant), Abhiroop Mukherjee (discussant), Spencer Pierce (discussant), David Reeb, Gideon Saar, Raffaella Sadun, Manpreet Singh (discussant), Scott Stern, and seminar participants at the AFA, AIEA-NBER Conference on Innovation and Entrepreneurship, CICF, Cornell Finance Brownbag, FARS Midyear Meeting, Hong Kong University Brownbag, and the Hong Kong Joint Finance Research Workshop for their valuable comments.

†SC Johnson Graduate School of Management, Cornell University, 301I Sage Hall, Ithaca, NY 14853. Email: gak79@cornell.edu.

‡Faculty of Business and Economics, University of Hong Kong, 923 KK Leung Building, Pokfulam Road, Hong Kong 999077. Email: apkwan@hku.hk.

§Kelley School of Business, Indiana University, 1275 E 10th St, Bloomington, IN 47405. Email: kenmerk@iu.edu.

1 Introduction

It is well-understood that information frictions inhibit financial markets from efficiently allocating capital towards innovative firms and projects (Hall and Lerner, 2010). Disclosure to investors helps alleviate these frictions and is considered essential to modern capital markets. However, disclosure risks harm to the innovator if prospective competitors learn of the innovator's activity. In his seminal work on the welfare economics of knowledge production, Kenneth Arrow advanced the paradox that, in order to gauge the economic value of an idea, the inventor is forced to disclose information that inherently reduces the idea's value (Arrow, 1962). Arrow's disclosure paradox has implicitly been central to recent debates on the design of emerging forms of financing for startups, particularly crowdfunding, which prioritize transparency while potentially undermining entrepreneurs' competitive advantage.[1] The disclosure paradox also challenges policymakers' and scholars' calls for more disclosures by innovative firms seeking to participate in public capital markets.[2]

Studying the tension between disclosure and innovation is challenging, as innovation is often secretive and withheld disclosures are generally not observable. In this paper, we employ a novel setting: the intellectual property (IP) licensing agreement disclosures made by publicly traded firms. Our setting allows us to observe both the disclosures that companies make and those they withhold from public view. Public firms are required by regulators to disclose all material agreements they enter into.[3] However, wary that disclosures to investors may reveal sensitive information to competitors, the Securities and Exchange Commission (SEC) allows firms to file Confidential Treatment Requests (CTR), which enable firms to redact information that may constitute "competitive harm." We argue this is equivalent to nondisclosure of aspects of the underlying technology and the potential rents from licensing it. However, any member of the public, including a researcher, can request the full, uncensored document once the CTR expires, usually after the contract matures or at least ten years after the CTR is initially granted. Consequently, our setting enables us to observe both those IP licensing contracts firms choose to fully disclose as well as those they strategically redact.

Through a partnership with a consulting firm specializing in the valuation of intangibles, we obtain a comprehensive sample of both redacted and fully disclosed material IP licensing agreements filed with the SEC. The data cover royalties related to patents, trade secrets, trademarks, copyrights, and other types of IP. We assemble nearly four thousand intellectual property licensing agreements that meet our sampling restrictions. To tbe best of our knowledge, this sample of licensing agree-

[1] In a recent interview (see http://bit.ly/2Ay6Sw9), Josh Lerner notes "A lot of my doubts have to do with the inherent contradictions between the entrepreneurial process and disclosure requirements [...] but the very process of disclosing things is likely to destroy a lot of the competitive advantage that the entrepreneurs might have."

[2] In a recent open letter to the SEC (Chien et al., 2016), several legal and innovation scholars have requested the SEC to make data on IP disclosures, which we study in this paper, more available for public scrutiny. In 2009, the SEC started requiring firms to electronically file Form D, causing some practitioners to comment that the disclosure requirement is making it harder for young innovative firms to raise capital (see http://redeye.firstround.com/2008/06/the-death-of-st.html).

[3] We discuss the definition of materiality and its corresponding threshold in a later section.

ments is the largest and most representative sample of redacted disclosures to be used in academic research to date.[4] In our sample, approximately one-third of the agreements are redacted, thereby enabling us to analyze market and firm responses to strategic disclosure choices.

Beyond providing a setting for studying Arrow's disclosure paradox, IP licensing is an economically significant activity and merits examination on its own terms: Total IP licensing royalty revenues of U.S. corporations amounted to $195 billion in 2013.[5] IP is perhaps the primary asset for innovative firms. Licensing agreements represent the main form by which firms commercialize their IP outside their own boundaries by facilitating the trade of IP in the marketplace for ideas (Shapiro, 1985; Gans and Stern, 2003). Licensing allows for an efficient division of labor between upstream firms or non-commercial entities, such as universities, which are better suited for conducting research, and downstream firms with product market expertise, while avoiding duplicative investments for both parties.

Using the IP licensing setting, we examine investors' reactions to innovative firms' decisions to redact from contract disclosures important licensing and technology details, thus capturing the key tradeoff implied by Arrow (1962). The predominant view in the literature is that when firms redact, investors infer that the withheld information contains bad news, resulting in negative capital market reactions (Grossman and Hart, 1980; Grossman, 1981; Verrecchia, 1983; Verrecchia and Weber, 2006). However, when deciding whether to redact information, firms must balance the potential benefits of disclosure to investors with the potential costs of revealing information to rivals. We conjecture that, ceteris paribus, firms with more commercially valuable IP have greater incentives to redact. If the firm chooses to redact, then although the act of redaction itself mechanically deprives investors of information, it also has the potential to convey information to investors, who may reasonably construe that redacted agreements are, on average, of higher value. However, if the innovating firm does not redact, and discloses to rivals, a rival may enter this space and erode the innovator's rents. We document that rational investors recognize the tradeoff between disclosure and value in this context, and systematically reward, rather than punish, the reticence of certain innovative firms in the capital markets. This is consistent with redaction serving as a credible signal of the value of the underlying idea.[6]

[4] Verrecchia and Weber (2006) study a firm-year cross-section of redacted disclosures made by firms with lower than $100 million market cap. The sampling choice is made for tractability but creates a bias toward small firms. In addition, not all firms have similar types of information to redact. Boone et al. (2016) study firms at the stage of IPO, which creates similar sampling issues. These are the only two studies of the determinants and consequences of redaction. In contrast to our data, in neither of these studies do the authors observe the uncensored contents of the redacted agreements. Since we are able to observe the full, uncensored contents of the redacted documents, we control for contract characteristics of both redacted and unredacted contracts in our analysis. Hegde (2014) and Hegde and Luo (2017) study a sample of around 500 licensing agreements in the biotechnology industry that notably does not include redacted agreements.

[5] IRS Statistics of Income 2013, "Returns of Active Corporations, Table 6." This is the latest year for which this data is available.

[6] To rationalize this equilibrium, we present a simple theoretical framework in the Internet Appendix A.3. In this framework, when both types of firms (with high quality and lesser quality technology) redact, the market assesses their concealed innovations to have the unconditional average value. This leads to a penalty on liquidity as in Verrecchia and Weber (2006). However, the lowest quality firm has the incentive to unilaterally deviate and disclose, benefitting in the capital markets at the margin with lower costs of losing out in the product market. Therefore, this scenario cannot be sustained in equilibrium. This unravelling continues until the marginal firm is indifferent between disclosing and not

We focus on liquidity as our main outcome of interest. Decomposing liquidity into its commonly understood determinants, a firm's stock may be illiquid because of asymmetric information or low investor demand (Glosten and Milgrom, 1985; Kyle, 1985; O'Hara, 1995). In the absence of signaling value, we would expect redaction to worsen liquidity by exacerbating adverse selection problems (Verrecchia and Weber, 2006). However, the positive signal that redaction provides regarding the value of the licensed IP would, by our conjecture, improve liquidity by stimulating investor demand or proving otherwise positively informative. Our results are in line with the signaling prediction and are at variance with prior literature finding that firms use nondisclosure to conceal bad news. Compared to firms that fully reveal their IP licensing agreements, we find that redacting firms receive an increase to their liquidity over the short-, mid-, and long-term horizons. Our evidence therefore suggests that the market rewards, and does not punish, firms for nondisclosure, on average. The magnitude is economically significant: In the one year post-filing, redactors experience an increase in liquidity, as gauged by turnover, that represents at least 20% of the unconditional mean, relative to nonredacting firms; this result is echoed at the weekly horizon in event-time analysis. Complementing this result, we find that redaction also predicts greater equity returns, likelihood of equity issuance, and levels of institutional ownership, thus confirming our results across a range of capital market outcomes.

The notion that redaction allows firms to preserve innovation rents gives rise to our second prediction: Redacting firms should be hiding current innovations or innovate more in the future. At the margin, firms with more profitable technologies sustain a greater advantage from erecting a barrier to entry by redacting. The barrier is necessary despite legal protection of IP rights, as many licensing agreements cover inherently unprotectable IP such as trade secrets or as-yet-unpatented inventions (e.g., patent applications, impending collaborations, or joint-ventures between firms). Disclosure of a firm's IP licensing may encourage competitors to develop substitutive technologies or to preempt a firm's pipeline of future innovation activities. Secrecy over licensing terms may also enable a firm to retain a favorable position in future negotiations with other parties. The protection afforded by redaction preserves rents for the filing firm, consequently providing a greater incentive to innovate.

In terms of the redacting firms' innovation activity, we find evidence consistent with investors' positive capital market responses being rational. Among redacted agreements containing references to patents, those patents show superior technical quality as measured by higher raw and patent-class-year-adjusted citations. Redaction also predicts higher levels of future innovation and innovative activities as early as one year out and for several years subsequently. Here, the magnitude is strikingly large: In the first post-filing year, redactors produce a greater total dollar value of innovation, as implied by stock market reactions to patent grants (see Kogan et al. (2017)) on the order of 40% of the unconditional mean, relative to fully disclosing firms. This suggests that the ability to withhold commercially sensitive information about their IP portfolio enhances firms' ex ante incentives

based on the relative magnitude of the costs and benefits, with firms with higher value of innovation than this marginal firm redacting and those with lower value of innovation disclosing.

to innovate, which, in turn, justifies the observed capital market reactions. Our main takeaway is that redaction allows firms with high-value IP to signal this fact, while limiting exposure to competitive expropriation, thereby partially addressing Arrow's disclosure paradox.

Because we find that nondisclosure is on average rewarded, our results appear surprising in the context of both current corporate disclosure literature and Arrow's disclosure paradox. It is worth noting that while our average effect holds for our sample of publicly-listed firms filing licensing transactions, it may not characterize all strategic nondisclosure of innovations. What we can glean is that there exists a set of conditions sufficient to mitigate the concerns raised by Arrow's disclosure paradox. However, innovators' disclosure may be preferable where the conditions in our setting do not apply. Therefore, one contribution of this paper is the elucidation of certain of those conditions.

We present several cross-sectional tests to guide our interpretation and to illustrate the limits of the conditions under which our signaling equilibrium may apply. In particular, we consider instances when redaction would be a less credible signal to market participants. Notably, we find that redacting firms in more competitive industries, or those with a greater number of past redactions, face a marginally negative, not positive, capital market reaction relative to similar firms that fully disclose. In the case of competition, market participants encounter an inference problem, i.e., are firms in competitive industries redacting due to the value of their IP or due to the intensity of competition? In the case of firms that frequently redact, investors may reasonably doubt whether the marginal redaction is in fact a credible signal of another valuable unit of licensed IP. These signal-jamming tests show that not all firms are able at all times to credibly signal the value of their technology to investors through redaction; they may instead experience negative capital market effects post-filing relative to fully disclosing firms, similar to prior studies (Verrecchia and Weber, 2006; Boone, Floros, and Johnson, 2016). Crucially, this result obtains despite the fact the redactors are, on average, no less innovative in the future than are fully disclosing firms. The positive signal of redaction in these cases is, essentially, jammed due to the inference problem facing investors when evaluating the disclosure choices made by these firms.

Subsequently, we discuss a natural experiment in which we exploit plausibly exogenous variation in a firm's benefits of redaction. We note that a key implication of the signaling hypothesis is that the firm's decision to signal through redaction is endogenous to the value of the licensed IP. It is precisely this endogeneity that renders the signal credible, thereby precipitating a positive market response to redaction. As a result, attempts at identification require us not to exogenously vary the choice of redaction, but rather the incentive or ability to redact.

We study the passage of the Food and Drug Administration Amendments Act of 2007, which mandated heightened disclosure requirements for all firms, private or public, involved in clinical trials. Firms are required to publicly record the details of completed clinical trials, thereby substantially reducing, if not eliminating, the benefits of redacting information on the innovation. Importantly, this affects all firms engaged in clinical trials and is orthogonal to the quality of the underlying licensed innovations. For these firms, publication of clinical trial success and the ensuing licensing activity are a clear indication that a firm has a profitable opportunity. As competitors are tipped off, the benefits of redaction are substantially diminished, if not negated. Thus, as in Verrecchia and Weber (2006), redaction would increase the information asymmetry between investors and these firms, resulting in a negative effect on liquidity. Our results are consistent with the prediction that lower benefits accrue to redaction. Complementing this result, we also find evidence suggesting the set of redacting firms either stays the same or becomes composed of less innovative firms. This experiment provides suggestive evidence that the capital market benefits of redaction partially operate through concealment of innovative activity, thereby validating concealment as a signal. Redacting firms may still desire to conceal commercial contract terms, but our natural experiment estimates indicate that at least a significant portion of nondisclosure's observed benefits can be attributed to protecting innovative activity.

Finally, we consider several alternative explanations. First, while of questionable legality, private meetings between company management and large shareholders may substitute for public disclosure. However, we find contravening evidence that, for redacting firms, investor concentration decreases (implying the entry of new investors) and the positive effects on liquidity and institutional ownership are weaker for firms with a more concentrated investor base. If private disclosure to major investors substituted for redaction, we would expect to find the opposite.

A second alternative is that the positive investor reactions to redaction are consistent with superior future information releases from these firms, and are not due to redaction as a signal. Simply put, redacting firms have superior innovative capabilities, and therefore produce more and better innovation and release news of innovation in the future. In order to mitigate this concern, we first show that our results obtain in a matched sample of firms. We match on variables which are likely correlated with a firm's ex ante innovative capabilities, such as the prior intangible capital stock. Our matched sample results suggest that preexisting differences in innovative capabilities are unlikely to drive our main findings. Second, our results on short-term liquidity, in which positive effects are seen in the weeks post-filing, suggest the market reaction is immediate and even precedes the news on future innovation. Third, we reestimate our main specifications with firm-fixed effects and find qualitatively similar results, although these results often have reduced statistical significance. This is due to the fact that firm-fixed effects can be estimated only among the small subset of firms in our sample who have filed multiple agreements, limiting the power of the tests. Nevertheless, these findings remain constructive, as they indicate our main findings are unlikely to be driven by time-invariant unobserved firm attributes such as superior innovative capabilities.

A third alternative is that firms use redaction to protect their bargaining power. Firms may redact to protect negotiating power when they have an idea sufficiently innovative to command royalties from multiple licensees. We view this interpretation as benign, since our policy implication entails redaction's protecting innovators whether from external competitors or from competing against themselves in future negotiations. However, our natural experiment removes the incentive to redact information related to the innovation, and not the bargaining terms. Yet we find that, at the margin, capital market responses to redaction turn negative. Furthermore, the cross-sectional evidence on competition contradicts the bargaining power alternative. Firms that retain stronger bargaining power in more competitive industries should face even greater positive reactions to redaction, which is contrary to what we find.

In summary, our results suggest that Arrow's disclosure paradox is resolved, but only partially. For those firms who can credibly signal their innovation potential, capital markets function and rationally infer that firms with more valuable IP redact to protect against expropriation. This credibility may arise from reputation or the bedrock of investor protection provided by a regulator or the firm's board of directors. When such conditions arise, firms with something worth hiding choose to redact, and firms with lower value IP choose to disclose. As a result, through redaction, innovative firms can enhance their access to public markets and finance their future innovative activities while mitigating the costs that would incurred by public disclosure.

We emphasize three contributions made by this paper. First, by analyzing the effect of relaxing a key informational friction, i.e., the disclosure paradox, on the ability of innovative firms to raise financing, we contribute to the literature on the financing of innovation. In their review, Kerr and Nanda (2015) describe four key frictions to financing innovation.[7] Of these, the issue of disclosure has received considerably less attention despite its central importance. Our paper fills a gap between the literature of corporate disclosure and that of innovation. The former strain investigates disclosure and capital markets but does not address the nuanced context of innovative firms, while studies of empirical innovation have largely ignored the effect of disclosure on capital market outcomes. Although the topic has surfaced in recent theoretical work (Thakor and Lo, 2017), to the best of our knowledge, our study is the first to empirically examine the disclosure problem directly in the context of innovative firms.[8] The primary challenge to studying the disclosure friction in an empirical setting is the difficulty of observing the counterfactual of what was not disclosed. By comparing the contents of disclosed versus redacted documents, we provide a valuable template for future work on strategic disclosure.

Second, our study speaks to theoretical and empirical work arguing for the importance of a well-functioning marketplace for ideas to the promotion of entrepreneurship and innovation (Akcigit, Celik, and Greenwood, 2016; Gans, Stern, and Wu, 2016). Building upon prior empirical studies on trade in the market for ideas through licensing (Gans, Hsu, and Stern, 2008; Hegde, 2014; Hegde and Luo, 2017), we introduce an extensive database on IP licensing. To the best of our knowledge, it is

[7]First, the payoffs to innovation are uniquely uncertain. Second, the returns to successful innovation are highly skewed, rendering projects difficult to accurately value and diversify. Third, with heightened information asymmetry, the inability to write state-contingent dynamic contracts is particularly acute for innovative firms. Fourth and finally, the disclosures required in the process of raising capital can subject innovators to risk from competitive expropriation.

[8]In a cross-country study, Brown and Martinsson (2018) find that more transparent information environments are associated with greater levels of innovation.

the largest and most detailed of its kind to be used in academic research. Our results also speak to a theoretical debate concerning the plausibility of signaling models in the real world, thereby demonstrating the existence of important settings that provide sufficient conditions to sustain such equilibria (Connelly et al., 2011).

Finally, we provide new insights on the costs and benefits of corporate disclosure, particularly in the context of innovative firms. Prior empirical studies on redacted disclosure (Verrecchia and Weber, 2006; Boone, Floros, and Johnson, 2016) find a negative effect of redaction on capital market outcomes, which they argue is evidence for redaction's increasing information asymmetry problems between firms and investors.[9] Our results, which emphasize a positive capital market reaction, differ due to our focus on innovative firms (i.e., firms that engage in IP licensing),[10] for whom redaction per se is uniquely informative. Further, Verrecchia and Weber (2006) consider a sample of small firms (between \$50 and \$100 million market capitalization), and Boone, Floros, and Johnson (2016) consider a sample of IPO firms, which are by definition young. These are groups for whom information asymmetry problems are likely to be more acute and potentially outweigh the positive signaling effect.

Our findings on the benefits of nondisclosure for innovative firms complement the SEC's existing concerns that more disclosure is not always better for investors.[11] In 2017, SEC Chairman Jay Clayton stated that disclosure requirements are seen as being too onerous, and may deter growing firms from raising capital in public markets.[12] Amid increasing calls by academics and policymakers for greater transparency and disclosure, our findings should give pause for thought. As innovative firms compose a growing segment of economic activity, our results suggest that exceptions to mandatory disclosure requirements, such as CTRs, could be useful in striking the right balance between ensuring investor protection and encouraging firms to use public capital markets. This balance may serve as a bulwark against the trend of the declining number of listed firms in the United States (Doidge, Karolyi, and Stulz, 2017; Doidge, Kahle, Karolyi, and Stulz, 2018), thereby allowing a broader set of investors to participate in the economic gains of innovation (Ljungqvist, Persson, and Tåg, 2018). In our conclusion, we elicit comparisons between various forms of financing of innovative firms and how our results may apply.

[9] A contemporaneous study by Tian and Yu (2018) compares different types of redactions. They find redacting firms experience superior product market performance, particularly firms whose redactions are related to licensing or R&D. This corroborates our main finding, and also helps motivate our paper which investigates the underlying mechanism. Our paper explores the inherent tension between innovative firms' need for secrecy and competing motive to inform stakeholders.

[10] The SEC also permits redaction on other types of filings including credit, leasing, customer-supplier, employment, and equity related agreements. Our sample consists exclusively of licensing agreements, whereas the samples of Boone, Floros, and Johnson (2016) and Verrecchia and Weber (2006) consider redactions on all types of filings. Our sample thus allows us to study the role of redaction of specific, IP-related information, corresponding most closely to the tradeoff implied by Arrow's disclosure paradox.

[11] For more details, see https://www.sec.gov/News/Speech/Detail/Speech/1365171492408.

[12] See http://on.wsj.com/2jbmXmZ.

2 Institutional Context

As part of the Securities Acts of 1933 and 1934, companies with publicly sold securities are required to state all material facts relevant to investors who might buy their securities. In particular, firms are required to disclose "material definitive agreements not made in the ordinary course of business." The definition of materiality was sharpened through a variety of historical developments. The modern standard of materiality was refined in the 1976 case *TSC Industries, Inc. vs. Northway Inc.*, which defined materiality as a fact which, if omitted, would have significantly altered the "total mix of information available" in a shareholder's decision to purchase a security or vote.[13]

In spite of broad guidance provided by regulatory agencies, deciding whether a given corporate development crosses the materiality threshold is often subject to the discretion of the company. Some guidelines do exist: For example, the 5% rule of thumb suggests that reasonable investors would not be influenced in their investment decisions by actions that lead to fluctuations of net income of less than 5% (Vorhies, 2005). For a detailed discussion on the considerations involved in auditors' determination of materiality, see Choudhary, Merkley, and Schipper (2018). Regardless of the exact threshold, material disclosures represent significant events, the knowledge of which affects the decisions of current and prospective investors.

However, recognizing the potential commercial harm of disclosing proprietary information, the SEC has allowed firms to request confidential treatment of certain information in material filings, as stipulated in Rule 406 under the Securities Act of 1933 and Rule 24b-2 under the Securities Exchange Act of 1934. The Staff Legal Bulletin No. 1 issued in February 1997 and its Addendum issued in July 2001 on the subject of Confidential Treatment Requests (CTR) by the Division of Corporation Finance, contain guidance for filing firms requesting confidential treatment.[14] From this, we are able to infer details about the SEC's interpretation of confidential treatment requests. In particular, the guidance explicitly notes the proprietary cost argument for requesting redaction:

> *"Sometimes disclosure of information required by the regulations can adversely affect a company's business and financial condition because of the competitive harm that could result from the disclosure. This issue frequently arises in connection with the requirement that a registrant file publicly all contracts material to its business other than those it enters into in the ordinary course of business. Typical examples of the information that raises this concern include pricing terms, technical specifications and milestone payments."*

[13] In the words of Justice Thurgood Marshall, "An omitted fact is material if there is a substantial likelihood that a reasonable shareholder would consider it important in deciding how to vote... It does not require proof of a substantial likelihood that disclosure of the omitted fact would have caused the reasonable investor to change his vote. What the standard does contemplate is a showing of a substantial likelihood that, under all the circumstances, the omitted fact would have assumed actual significance in the deliberations of the reasonable shareholder. Put another way, there must be a substantial likelihood that the disclosure of the omitted fact would have been viewed by the reasonable investor as having significantly altered the 'total mix' of information available."

[14] Full text is available at https://www.sec.gov/interps/legal/slbcf1r.htm.

The guidance specifies five requirements for CTRs: (i) They should not be overly broad. Specifically, "the information covered by an application should include no more text than necessary to prevent competitive harm to the issuer;" (ii) applicants must provide written justification for the exemption; (iii) applicants must specify a particular duration; (iv) applicants must identify clearly the information that is the subject of the application; (v) applicants must consent to the release of the information to the SEC for official purposes, e.g., evaluation or audit.[15] In Appendix Figure A.1.1, we give an example of a portion of a redacted licensing agreement between *Genome Therapeutics Corp.* and *bioMérieux, Inc.* dated September 30, 1999. In the section of the agreement that we have reproduced in Figure A.1.1, payment terms including royalty rates, sublicensing fees, etc., are redacted and replaced with "*...*".

3 Data, Hypothesis Development, Methodology

3.1 Dataset

The main data used in this paper are drawn from a database of intellectual property licensing agreements from RoyaltyStat LLC. RoyaltyStat is one of a handful of companies that collect data on royalty rates of IP licensing transactions, and was the first company to come to market with a searchable online platform in 2000. The data are sourced from SEC-mandated material disclosures, wherein the licensing agreements are usually attached as exhibits in both scheduled and one-off filings. RoyaltyStat has developed a natural language processing engine to identify relevant filings and automatically populate certain standardized fields. Analysts then review and augment the output. The data are updated daily.

RoyaltyStat provided us access to every record and a rich set of fields in their databases. To preserve data confidentiality, we did not custody the data. We were provided access through a password-protected portal in which we uploaded *R* scripts and received textual output. Appendix Table A.2.1 enumerates some of the key fields we derived from their data. We use these fields to account for contract characteristics in our analysis.

The unit of observation is a licensing contract entered into by two parties, a licensor and a licensee. The filer must be a public firm and is either the licensor or the licensee. The counterparty may not be a commercial entity e.g., a nonprofit hospital, university, or individual inventor, and may be a private or publicly listed firm. In addition to identifying information, we observe when the fil-

[15] The latter point was later clarified: "If the remaining term of the contract is greater than 10 years from the date of the extension application, we generally will only grant confidential treatment for 10 years; if the remaining term of the contract is less than 10 years from the date of the extension application, we will consider a request for the remaining term of the contract; if the remaining term of the contract is less than five years from the date of the application, but there is a possibility that it will be extended beyond its stated term, we will consider granting confidential treatment for a period of up to five years."

ing was made, and whether or not the filing was retrieved through the Freedom of Information Act (FOIA); that is, whether or not its contents were redacted at the time of filing. Under FOIA, the public can request the entire uncensored contents of specific filings after the expiration of the CTR. The dataset used in this paper is partly assembled through thousands of FOIA requests issued to the SEC in most cases ten years after the date of the original filing. The remainder of the dataset consists of filings that were un-redacted as of the time of dissemination. The ability to compare information that was disclosed with contemporary counterfactual information that was withheld from the public eye provides a valuable setting in which to test the impact of nondisclosure.

In terms of the contract itself, we observe contract terms, payment terms, and intangible information. Contract information consists of the effective date of the contract, the duration, the geographic scope of the licensing agreement, and the rights granted by the licensing firm to the licensee. Payment terms consist of royalty rates, upfront fees, and additional payment terms such as tiered royalty rates that change as the quantity sold increases, or milestone payments for reaching certain contracted goals. Finally, intangible information includes descriptors of the technology, its applicable industry, and, when mentioned in filings, intangible numbers such as patent grant or application numbers.

Using this data, we then construct our sample. RoyaltyStat maintains a link to Compustat as a sublicensor of Compustat data. We assume their link is valid if available, and if missing, we augment their linkage with either the *CIK* or filer name using various linking tables. Appendix A.2 reports the sample size at each stage of the filtering process and the sample selection procedure.[16]

In most specifications, we have between 3,000 to 4,000 observations. This sample size justifies our claim for assembling the first large-sample analysis of licensing agreements, as it is larger than extant studies with access to licensing contract-level data. Our sample is approximately as large as a typical study on syndicated loan contracts or mergers. [17]

In addition to our workhorse licensing dataset, we utilize capital market and innovation data. We obtain data on capital market consequences from CRSP, DTAQ, SDC Platinum, and Thomson Reuters 13/F. We explain the variables we construct and their significance to our inferences in 4.1,

[16]The sample faces the following restrictions: First, the firm must be U.S.-domiciled and publicly exchange-listed at the time of filing as we look for equity market consequences. The first filter eliminates many observations as the filings are often pre-IPO or filed by foreign-domiciled firms. Second, the licensing agreement must refer to royalties assessed against a percentage quantity such as "net sales", as opposed to a unit quantity, such as "per ton" or "per yard." Net-sales are the most common type of royalty bases for such agreements. This excludes royalties on mineral rights (net smelter returns), land royalties and per-unit production or sale royalties. These agreements may not be comparable to the main set of agreements assessed against net sales and also are less likely to involve innovative firms. Third, the firm must have nonmissing balance sheet information, such as data on annual assets and net sales. Finally, the transactions should be arms-length and not between related parties. We have 7,972 license agreements whose filer can be linked to Compustat and can be disambiguated as either the licensor or licensee, and of them, 6,579 are US firms with publicly traded equity covered by CRSP at the time of filing. Some firms which do not make it into our sample are penny stocks not covered by Compustat, unlisted firms, or even Canadian or Mexican public firms. We also lose observations due to missing values for control variables such as assets, leverage, or various contract characteristics. However, our qualitative inferences are not sensitive to excluding either contract or firm characteristics.

[17]We have a similar sample size to, for instance, Campello and Gao (2017).

our methodology section. The main innovation outcomes we consider are related to patenting. We use the dataset provided by Kogan et al. (2017) through 2010 to obtain firm-level innovation measures. $FNPATS$ refers to the raw count of all patents granted to a given firm in a given year. TSM refers to the total dollar value of patents granted to a given firm in a given year, estimated from the stock market reaction to the patent grants. For robustness, we replicate the results using the NBER 2006 patent sample with citations extended through 2014 to adjust for citation truncation issues per Lerner and Seru (2017). The conclusions remain similar.

3.2 Summary Statistics

In this section, we provide summary statistics to familiarize the reader with our dataset, and, more broadly, the world of intellectual property licensing. In Figure 1 we plot the distribution of contracts across industries, where the industry classification scheme used is the 42 technology industries as defined by RoyaltyStat. Specifically, each contract is classified into one or more of the 42 technology industries based upon the technology space that best describes the underlying IP being licensed; this determination is made by RoyaltyStat analysts. As expected in a study focusing on IP-intensive firms, the vast majority of the licensing contracts have underlying IP that falls under the categories of Pharmaceutical Biotech, Medical Devices, and Software.

FIGURE 1 ABOUT HERE

In Figure 2 we plot the yearly distribution of total contracts filed and the proportion which are redacted. Note that because redacted contracts only enter into the RoyaltyStat database once analysts have been able to make an FOIA request to retrieve the original unredacted filing, this implies that the CTR status of these contracts must have expired. As discussed in Section 2.1, a CTR lasts approximately ten years on average. Therefore, post-2006 we are significantly undersampling redacted contracts. This is simply because for the average redacted contract filed post-2006, the CTR is yet to expire. Thus, the original unredacted contract cannot be retrieved via a FOIA request.[18] In the pre-2007 period, the rate of redaction appears to be growing over time, and over the 2000-2006 period in which we observe the majority of both redacted and public contracts, the proportion of total contracts that are redacted ranges between 30% and 45%.

FIGURE 2 ABOUT HERE

[18]Since in the post-2006 period we are significantly undersampling redacted contracts, we rerun our entire analysis in the pre-2007 sample, and aside from a drop in sample size of around 10% to 15%, our results are qualitatively and quantitatively similar. To address this issue in our main analysis, we include year-fixed effects and in our matched sample, we require matched redacted and unredacted contracts to be within a three-year filing window of each other.

The panels of Table 1 display summary statistics of various contract characteristics. Panel A presents summary statistics on three key contract features: the license fee, the royalty rate, and the contract duration. The summary statistics confirm that these are generally economically significant contracts. The median license fee of $250,000 represents 2.3% of the median annual sales of a firm in our sample. These contracts are of a relatively long duration, with the median contract having a maturity of 17 years. Also, the distributions of rates and fees are highly skewed, with a long right tail.

In Table 1 Panel B, we show the distribution of agreements by the type of underlying IP. Note that these categories are not mutually exclusive since the IP underlying a single agreement often encompasses several types. Hence, the percentages sum to greater than 100. By far the most common form of IP is patents, followed by technology, know-how, and trademarks.

Finally, we obtain a set of 64 randomly sampled redacted contracts, and manually analyze these documents to determine the specific contract terms redacted by the filing firms. The extensive margin proportion of these contracts in which each specific contract term was redacted is given in the first column of Table 1 Panel C. In almost all cases the royalty rate is redacted, followed by payment-related terms and terms which identify the underlying IP.

However, since not all contracts explicitly state these contract terms, we also calculate these percentages restricting only to those redacted contracts which include each clause. In the second column of Table 1 Panel C, we present these intensive margin proportions, that is, the rate at which each clause is redacted conditional on it being explicitly mentioned in a contract. The statistics tabulated in Table 1 Panel C indicate that information related to intangibles is also redacted in the vast majority of cases when it is explicitly mentioned. In fact, three of the five most commonly redacted terms are related to the licensed IP.[19] Taken together, these results suggest that firms seem to redact both the payment terms of the licensing deal and the nature of the licensed IP. This evidence supports the idea that redaction signals the value of the underlying IP while concealing it, given that the economic value of IP is a function of both its identity and its commercialization terms.

TABLE 1 ABOUT HERE

In Table 2 we present summary statistics describing the firm-level characteristics of filers in our sample. We also compare the firm characteristics of our sample to the overall sample of public firms in Compustat. We expect that, since we are studying a sample of innovative firms, i.e., firms which have at least one material IP licensing agreement, the average firm in our sample will be smaller than the average firm in Compustat (mean assets of $311M in our sample versus $2.53B overall), the average firm will be more R&D intensive (mean R&D/Assets of 33.7% in our sample versus 6.6% overall), and the firm will be younger (mean age of 13 years in our sample versus 20 years overall).

[19] These statistics may understate how frequently IP information is redacted, as the absence of an explicit mention may itself be an implicit redaction.

TABLE 2 ABOUT HERE

Next, we present graphical evidence that firms with greater proprietary costs of disclosure, namely firms with superior technology or in more competitive product markets, are more likely to strategically redact.[20] This condition is essential for redaction to be a credible signal to investors, and hence justifies the positive effect of redaction on capital market and innovation outcomes as illustrated by our model described in A.3 of the Internet Appendix. In Figure 3 Panel A, we plot the fraction of redacted contracts among the total set of filed contracts across quintiles of product market competitiveness proxied by the Hoberg and Phillips (2016) measure of product market fluidity. This measure captures the tendency of new words in a firm's product description to also appear in competitors' documentation in following years, with higher values corresponding to a greater intensity of product market rivalry. The redaction rate increases with product market competitiveness, suggesting that firms facing greater competitive threats are measurably more likely to redact. Firms in the highest quintile of product market competitiveness, who are more likely to face the threat of potential entry, redact 62% of their filings, whereas firms in the lowest quintile redact only approximately 10% of their filings.

FIGURE 3 ABOUT HERE

In Figure 3 Panel B, we plot the redaction rate across quintiles of total prior five-year patent citations for the subset of contracts whose underlying IP consists of patents. Patent citations are customary proxies for the economic value of patents (Hall, Jaffe, and Trajtenberg, 2005). Thus, contracts that license more economically valuable patents are more likely to be redacted. This effect is particularly skewed among the highest quintile of contracts, which have a 9% higher redaction rate than those in the second highest quintile. The graphical evidence supports the notion that redaction is a signal used by firms who have more economic value to lose from being forced to disclose the value of their IP.

3.3 Hypothesis Development and Methodology

The impact of redaction on capital market variables is unclear ex ante. Generally, a firm's choice to not disclose information can be met with a negative capital market response due to increased information asymmetry. In particular, rational investors may assume that firms use redaction to conceal bad news, which aggravates adverse selection problems between managers and investors, in turn generating negative capital market effects (Grossman and Hart, 1980; Grossman, 1981; Verrecchia, 1983; Verrecchia and Weber, 2006).

[20] Although we do not report these results, we also obtain similar inferences with econometric analysis modeling the decision to redact as a function of product market competition and patent citations.

In contrast, redaction can be met with a positive capital market response if that redaction signals positive future prospects. Innovative firms, in particular, may choose not to disclose their IP because doing so reveals information to competitors whose use of that information reduces the innovative firm's rents. Redaction will only be a credible signal of the value of IP if firms with high- (low-) valued IP have incentive to redact (fully disclose) in equilibrium, i.e., the equilibrium is separating. In this equilibrium, the firm with lower value technology will disclose to deter entry.

To illustrate, consider the case of a firm with either a high-value licensed IP or a low-value licensed IP. Managers possess this information, which is inaccessible to either investors or potential competitors unless disclosed. The firm learns of a potential entrant which, if it learns that the licensed IP is of high value, will have a greater incentive to enter the technology space, invest in developing a substitutive technology, and compete away the rents. The incumbent firm then has to decide whether or not to redact the contents of its material licensing agreement filed with the SEC, taking into account the effect this decision may have on the potential entrant's incentive to enter.

An incumbent with a high-value IP has a greater incentive to redact because disclosing that its licensed IP is of high value increases the likelihood of a rival's entering with a substitutive technology and competing away its presumably high profits. If the incumbent redacts, the entrant will face a barrier to entry and have to invest in market research and R&D to determine the nature of the technology and its value before deciding to enter. We posit that the higher the value of the IP, the greater the implicit cost barrier for the entrant due to redaction by the incumbent.

For a firm with low-value IP, the decision is more ambiguous. If it redacts, then the rival will be unable to distinguish whether the IP is low or high value. The rival may decide to enter if it believes that at the expected value of the IP, the potential rents, are worth the expected cost of researching and developing a competing technology. If, however, the low-value IP firm discloses its IP, the rival will have a lower incentive to enter and compete, due to the possibility that the projected lower profits may not exceed the fixed costs of entry.

Therefore, provided that the potential expected rents are high enough, relative to the expected costs of search and entry, incumbents with high-value (low-value) IP have incentives to redact (disclose). In doing so, the firm with high-value IP effectively constructs an insurmountably costly barrier to entry by redacting, while the firm with low-value IP "hides in plain sight" by disclosing that its technology space is not worth entering. Thus, redaction serves as a credible signal of the value of the licensed IP. While investors would ordinarily react negatively to redaction, inferring that the suppressed information contains bad news, rational investors are able to disentangle the mechanical deprivation of information from the differential incentives of high (low) value IP firms to redact (disclose). Therefore, the market interprets redaction as a positive signal of the value of the licensed IP, and reacts positively to it. We formalize this intuition in a simple stylized game, presented in section A.3 of the Internet Appendix. We show that when both types of firms (high quality and lesser quality technology) redact, the market assesses their concealed innovations to have the unconditional average value. This leads to a penalty on liquidity as in Verrecchia and Weber (2006). However, the lowest quality firm has the incentive to unilaterally deviate and disclose, benefitting in the capital markets at the margin with lower costs of losing out in the product market. Therefore, this scenario cannot be sustained in equilibrium. This unravelling continues until the marginal firm is indifferent between disclosing and not based on the relative magnitude of the costs and benefits, with firms with higher value IP than this marginal firm redacting and those with lower value IP disclosing.

Our intuition corresponds to several theoretical models of voluntary disclosure. This literature, starting with Verrecchia (1983), models the tradeoff faced by firms making the decision to voluntarily disclose information to investors, which is presumed to improve the precision with which the market can value their stock, yet which incurs proprietary costs. Verrecchia (1983) establishes the existence of a partial disclosure equilibrium, and shows that the greater the proprietary cost of disclosure, the less negative will be traders' reactions to the withholding of information. Further studies have endogenized the proprietary cost of disclosure as a function of product market rivals' responses to the disclosed information (Bhattacharya and Ritter, 1983; Darrough and Stoughton, 1990; Feltham and Xie, 1992; Wagenhofer, 1990). The signaling equilibrium underlying our hypotheses is also consistent with Anton and Yao (2002), which shows that partial disclosure of informaton about an innovation can serve as a signal of its value.

Given the discussion above, we formulate two empirical hypotheses we take to the data. The first concerns the effect of redaction on capital market outcomes. We argue that for innovative firms, the redaction of licensing contract terms is a credible signal of the value of the licensed IP. The fact that firms with high-value IP are more likely redact, while firms with low-value IP are more likely to disclose, gives rise to our first hypothesis.

Hypothesis 1. *Firms which redact their material IP licensing agreements experience a positive capital market reaction post-filing relative to firms which do not.*

Since firms with high-value licensed IP, through redaction, are able to block potential competitors from developing substitutive technologies, thereby preventing the expropriation of rents, they have greater incentives to conduct further innovative activities in this area. This line of reasoning leads to our second hypothesis, on the real effects of redaction.

Hypothesis 2. *Firms which redact their material IP licensing agreements perform a greater level of innovation activity post-filing relative to firms which do not.*

To test the above hypotheses, we first estimate annual contract-level panel regressions of the form:

$$Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t} \tag{1}$$

where Y refers to an outcome variable measured at the horizon $t+k$, where t is the contract filing date, and k refers to months (years) post-filing for monthly (annual) outcome variables. $\mathbf{1}\{REDACT\}$ refers to a dummy variable for whether a given contract i, filed at date t, by firm j, with underlying technology belonging to SIC 3-digit industry l, was redacted. **X** is a vector of firm- and contract-level controls. We include fixed effects for the SIC 3-digit industry code assigned to the technology underlying a given contract, α_l, and year fixed effects, α_t.

Note that in the above specifications, we do not distinguish between the role of the filing firm as a licensor or licensee. We expect our hypotheses to hold regardless of whether the firm is a licensor or licensee. Both types of firms have an incentive to signal the value of the licensed IP through redaction. A licensee filing firm may continue downstream innovation, and therefore wish to maintain secrecy. A licensor would not want to reveal its licensee's future plans as that may reduce the licensor's future revenues, and therefore may continue to support the licensee through their own R&D.[21]

The set of firm-level controls that we include in our regression specification consists of the logarithm of total assets($\log(Assets)$), the logarithm of total intangible capital as defined by Peters and Taylor (2017) ($\log(K_{int})$), the ratio of cash to assets $\left(\frac{Cash}{Assets}\right)$, leverage ratio $\left(\frac{Debt}{Assets}\right)$, and the $SmallYoung$ index value which collapses the age and size of the firm into a single index.[22] The set of contract-level controls includes dummy variables which take the value of 1 if the filing firm is the licensor($FilerLicensor$), if the licensor is an individual ($LicensorIsIndividual$), a non-profit ($LicensorIsNonProfit$), a university ($LicensorIsUniversity$), if the given contract is an amendment of a previous agreement ($IsAmendment$), if the underlying IP consists of know-how[23] ($IsKnowHow$), and show-how[24] ($IsShowHow$). We also include (i) the logarithm of the duration of the contract ($\log(Duration)$); (ii) a dummy for whether the licensing rights extend worldwide ($Worldwide$); (iii)a dummy for whether the agreement confers the licensee the right to sublicense the IP ($Sublicense$); and (iv) a dummy for whether the IP is being licensed exclusively to that particular licensee ($Exclusive$). We include these variables to capture common economic determinants of licensing contract terms (as discussed in Varner (2011) and Hegde (2014)).

Our coefficient of interest in Eq. (1) is β_1, and it can be interpreted as the average difference in the post-filing value of Y between firms which redact and those which do not. Specifically, when $k = 1$, the coefficient β_1 captures the average difference in variable Y at time $t+1$ (1 year post-filing), between firms which redact and firms which do not, controlling for the prior level of Y. A positive estimate of β_1 indicates that Y_{t+1} for redacting firms is on average greater than Y_{t+1} for nonredacting firms. We control for the prior level (more precisely, the value of Y one month prior to filing) of the dependent variable, which is denoted as Y_{t-1}, to mitigate scale effects not already captured by

[21] Nevertheless, in some specifications, we allow for differential effects for both types of filers, naturally finding stronger results for the licensor.

[22] The $SmallYoung$ index is calculated as minus one times the $SizeAge$ index as computed by Hadlock and Pierce (2010). The $SmallYoung$ index is calculated as $-1 \times SizeAge = -1 \times [(-0.744 \times Size) + (0.042 \times Size^2) - (0.075 \times Age) + (0.001 \times Age^2)]$, where $Size$ is the log of inflation-adjusted book assets, and Age is the number of years the firm is listed with a nonmissing stock price on Compustat. As in Hadlock and Pierce (2010), when we construct the index, $Size$ is winsorized at $\log(4.5 \times 10^9)$, and Age is winsorized at 37 years.

[23] Know-how encompasses unpatentable proprietary information including blueprints, drawings, data etc.

[24] Show-how consists of proprietary assistance rendered by the licensor to the licensee.

controlling for total assets and intangible capital, and persistence in Y. Finally, we control for other possible determinants of Y which may confound our interpretation, in vector $\mathbf{X}$, and for unobserved time- and industry-invariant factors that may affect Y through the use of fixed effects.

Second, we employ matching techniques in order to improve comparability of ex ante innovative capabilities between the sample of firms which redact and those which do not. In particular, our matching aims to address the argument that firms which have ex ante superior innovative capabilities are more likely to license more valuable IP, which they have a greater incentive to redact — and it is this inherent difference in capabilities that drives our results rather than redaction as a signal. By matching firms on variables that we contend are correlated with ex ante innovative capabilities, we are essentially comparing the market reaction to filing a redacted versus unredacted contract among firms which are substantively similar on this dimension. We match redacted and unredacted contracts on the basis of a propensity score which is a function of the $SmallYoung$ index (size and age being important unconditional determinants of a firm's likelihood to file patents) and the level of intangible capital ($\log(K_{int})$). In addition, we require an exact match on the 1-digit SIC industry code of the licensed technology, and we require that the treated and untreated contracts fall within a three-year block of each others' filing dates. Finally, we require an exact match on whether the underlying IP consists of patents or not, as firms which patent ex ante are much more likely to continue patenting.

We evaluate the improvement in comparability of the samples generated by matching by considering the percent improvement in balance for each of the matching variables, defined as $100 \times \frac{(|a|-|b|)}{|a|}$, where a is the balance statistic before and b is the balance statistic after matching. The balance statistics considered include means of each matching variable, as well as mean, median and maximum value of the empirical distribution of each matching variable across the treatment and control samples. In Table A.5.1 of the Internet Appendix, we show the mean improvement in balance for each of the variables on which we construct the propensity score, and the score itself, as well as improvement in the mean, median, and maximum of the empirical distribution of each matching variable across the treatment and control samples. The score itself (labeled as Distance), improves by 99% on average, the balance of the $SmallYoung$ index improves by 70% on average, and the balance of intangible capital improves by 94% on average. As noted by Imai, King, and Stuart (2008), this is a superior method of comparing balance of covariates pre- and post-matching.[25] Finally, in the matching analysis, we cluster the standard errors at the level of the matches, as recommended by Abadie and Spiess (2016).[26]

[25] The usual method of computing t-statistics on the difference in means of the variables of interest is highly misleading for four reasons described in their paper.

[26] The broad inferences are robust to alternative matching schemes, but we choose a somewhat coarse industry definition and three-year time intervals to avoid oversampling a subset of contracts in the control group.

4 Outcomes of Strategic Redaction

In this section, we present the results of the empirical tests of our two hypotheses. After establishing the main result, we perform additional analyses to demonstrate the limits of the conditions under which our signaling equilibrium holds, and to argue against conceivable alternatives. In order to demonstrate these limits, we present evidence that for some subsets of firms, redaction is associated with a negative capital market response despite redacting firms' being no weaker on future fundamentals than nonredacting firms. We show that the positive signal is jammed, and investors face an inference problem, since firms cannot credibly signal the value of the licensed IP. For example, redacting firms in more competitive industries, or those which have redacted more in the past, face a marginally negative capital market reaction relative to similar firms which fully disclose, even though the redactors are no less innovative in the future. For firms in more competitive industries, investors' inference is clouded as they are unable to distinguish whether redacting firms are in fact concealing high-value IP or are simply avoiding competition for even their low-value IP, which rivals are unlikely to imitate anyway. In the case of firms which redacted more frequently in the past, the inability to credibly signal may be due to a "crying wolf" effect, in which investors deduce that there are diminishing returns to the ability of a firm to produce consistently valuable IP. In both cases, investors fears are immaterial as redacting firms produce the same or greater levels of innovation in the future, relative to nonredactors. This evidence concerning signal jamming justifies the argument that redaction is a signal, wherein investors at some times face an inference problem. We next provide evidence from a natural experiment exploiting exogenous variation in the benefit firms receive from redacting information related to their innovations.

Throughout the discussion of our results we address alternative explanations. The primary alternative to our story is that redaction is not a signal, but rather it is common among high-type firms and capital market outcomes are driven by superior fundamentals of these firms in the future. We take several approaches to dispatching this alternative. First, we show that in a short horizon the capital market response is positive, years before any future fundamentals are realized. Second, our results obtain under matched sample analysis which enforces comparisons between fundamentally similar firms. Finally, we present robustness checks in which we re-estimate our main specifications under firm-fixed effects. Our estimates are directionally consistent, but with reduced significance, owing to the fact that the majority of firms in our sample file only a single contract.

4.1 Redaction and Capital Market Outcomes

In Table 3, we consider the effect of redaction on future capital market outcomes measured in three ways. For each contract i filed by firm j at time t, we compute daily turnover for firm j's stock, which is the end-of-day price multiplied by the volume of shares traded. We average this across all days in the three, six, and 12 months after filing time t to obtain the variable $Turnover$. Next we compute daily bid-ask spread for firm j's stock, which is the closing ask minus the closing bid scaled by the end-of-day price, and average this across all days in the three and six months after filing time t to obtain the variable $Spread$. Finally, we calculate daily values of Amihud's illiquidity measure (Amihud, 2002) for each day in the next three and six months after time t. We do this by computing the ratio of the daily absolute return to the (dollar) trading volume on that day, for firm j's stock. We average this value across all days in the three and six months after filing time t, to obtain the variable $Illiq$. We interpret increases in turnover and decreases in the latter two measures as improvements to liquidity.[27] We take the logarithm of the latter two variables to form $\log(Spread)$ and $\log(Illiq)$ to control for outliers, consistent with the convention in the literature. With these three outcome variables, we estimate the panel regression in Eq. (1). Results are reported in Table 3.

TABLE 3 ABOUT HERE

Panel A contains OLS coefficient estimates for specification (1) with firm- and contract-level controls, as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology industry level. Panel B contains estimates from a matched sample. The coefficient of interest is the coefficient on $\mathbf{1}\{REDACT\}$, which measures the difference in the level of the dependent variable in the three, six, and 12 months post-filing, between firms which file redacted and un-redacted contracts. For $Turnover$, across all three horizons, the coefficients on $\mathbf{1}\{REDACT\}$ are positive and significant, indicating redaction is associated with improved liquidity post-filing. For $\log(Spread)$ and $\log(Illiq)$, for both horizons, the coefficients on $\mathbf{1}\{REDACT\}$ are negative and significant, also suggesting improved liquidity. These magnitudes are economically significant. For instance, the mean value of $Turnover_{t+3}$ in our sample is 6.356, and thus the marginal effect of redaction of 1.990 from the first column indicates that firms filing redacted contracts are associated with a post-filing improvement in turnover of 31% of the mean in the subsequent three months, relative to firms which file unredacted contracts. Results from Panel B provide further support for the positive effect of redaction on liquidity. The estimates from the matched sample are similar in significance and magnitude to the OLS estimates, suggesting that the effect is not driven by differences in observable covariates between redacting and nonredacting firms.

Thus, liquidity improves post-filing for firms filing redacted contracts relative to those filing unredacted ones. This suggests that contrary to the findings of Verrecchia and Weber (2006), investors

[27] Turnover could reflect disagreement owing to bearishness of the market as opposed to purely liquidity. Thus, turnover is important for validation with respect to prior literature (Verrecchia and Weber, 2006), but leaves itself open to interpretation. Our second variable is spread. Spreads embed two components, an increase in adverse selection concerns versus an increase in demand. To the extent that redaction mechanically deprives investors of information, a reduction in spread must mean a net reduction in the change of adverse selection (which weakly increased) and demand (whose increase is enough to offset this). Thus a decreased spread is in our view unambiguously associated with improved liquidity. Similarly, the Amihud illiquidity measure is a measure of price impact, which can similarly be thought to unambiguously reflect a net increase of investor demand and information.

do not penalize firms which strategically redact disclosures related to IP licensing agreements through reduced liquidity.[28] These results provide support for our first hypothesis claiming that redaction serves as a credible, investor-recognized signal of the value of a firm's licensed IP, and is associated with a positive capital market response.

We next analyze the market reaction to redaction in terms of liquidity in event time. In the event time analysis, we estimate the following regression:

$$Y_{i,t} = \beta_1 \mathbf{1}\{REDACT\}_{i,t} \times Post_{i,t} + \beta_2 Post_{i,t} + \sum_i \alpha_i + \epsilon_{i,t} \tag{2}$$

This regression is estimated with weekly data and i indexes the filing event. The analysis aims to show that redaction is indeed a signal, and the positive liquidity response is not solely a response to the realization of the superior future innovations of redacting firms relative to firms which do not redact. If the positive liquidity response to redaction is a result of investor reaction to redacting firms' performing more future innovation after the filing date, we should not expect to see any short-term impact in response to the redaction signal. Innovation investments are inherently long term and unlikely to bear fruit at a shorter horizon. We compute high frequency measures of liquidity from intraday data and regress these on indicators for redaction and post-filing in an event window of 15 weeks pre- and post-filing, excluding the week of the filing event to mitigate the effect of concurrent announcements. We use three liquidity measures as outcome variables that form analogues to our longer-horizon analysis.[29] *Post* is a variable that takes the value of 1 for all observations that fall in the 15-week window following the filing date of the disclosure (excluding the week of the event itself, to control for confounding filings), and 0 otherwise. All specifications include event fixed effects (essentially demeaning by event), and the standard errors are clustered by event.

[28] This can be partially attributed to the fact that Verrecchia and Weber (2006)'s sample differs from ours in two key ways. First, they focus on small firms (between \$50 and 100 million market capitalization). In contrast, the average firm in our sample has a market capitalization of \$990 million, and the largest firm in our sample has a market capitalization of \$279 billion. Second, they consider all varieties of redacted material agreements, including debt, equity, employment, customer-supplier, and IP licensing related agreements. Our sample focuses exclusively on IP licensing agreements, and thus includes relatively innovative firms, for whom the benefits of nondisclosure are likely to be more pronounced. Consistent with their findings, however, we show that small and young firms that redact in our sample, also suffer a marginally negative effect on their stock liquidity post-filing as compared to firms which do not redact.

[29] The quoted spread is defined as the difference between the prevailing market quoted bid and ask prices, divided by the midpoint between these two prices. The effective spread is defined as twice the absolute value of the difference between the actual trade price and the midpoint of the market quote (i.e., between the quoted bid price and the quoted ask price), divided by the midpoint between these two prices based. The price impact of trades, also known as Kyle's Lambda, is measured by performing regressions of the absolute return over a given time period on dollar volume. The price impact is defined as the slope from this regression. These slopes are averaged to calculate daily and annual measures of price impact in an analogous way to the spread and turnover obtained from CRSP. Higher values of price impact correspond to a given magnitude of trading being associated with a greater movement in returns, and thus price impact is seen as a measure of stock illiquidity (in fact, this is an intra-day high frequency analogue of the (Amihud, 2002) measure which we compute from CRSP).

The results are tabulated in Table 4. The coefficient on the interaction term $\mathbf{1}\{REDACT\} \times Post$ represents a "difference-in-differences," and is negative and statistically significant in all but one cases, suggesting redacting firms experience improved liquidity relative to nonredacting firms in the 15 weeks post-filing, relative to the 15 weeks prefiling. To account for noise and outliers in these intraday measures of liquidity, we winsorize them at the 1% level, z-score them within-firm, and take their logarithm: we find that the effect persists across all measures.[30] The short-term liquidity response is unlikely to be driven only by the realization of greater future innovation activity by redacting firms, which would transpire over longer horizons.

TABLE 4 ABOUT HERE

We next consider the effect of redaction on post-filing equity returns, likelihood of equity issuance, and institutional ownership. In Table 5 Panel A, we consider the effect of a firm's redacting its filing on its post-filing stock returns, and on the likelihood that it conducts a seasoned equity offering (SEO) in the subsequent one, two, and three years. We compute equity returns by compounding monthly raw returns, starting from the filing date, for the subsequent three and six months to form the variable Ret_{t+3} and Ret_{t+6} respectively. We also compute the monthly alphas net of the Fama-French five-factor model, and compound these over the subsequent three and six months starting from the filing date to form the variables α_{t+3}^{FF5} and α_{t+6}^{FF5} respectively. We construct the variable $\mathbf{1}\{Issuance\}$ by assigning a value of 1 for firms which conducted a secondary equity offering in the subsequent one, two, and three years post-filing and 0 otherwise.

TABLE 5 ABOUT HERE

The first four columns of Panel A contain OLS coefficient estimates of the effect of redaction on equity returns three and six months out, and alphas for the same time horizons. In all four specifications we find directionally positive effects, and one out of the four coefficient estimates are statistically significant. Thus, there is some evidence that redaction is associated with higher equity returns for redacting firms in the subsequent two quarters post-filing relative to fully disclosing firms. The coefficients on $\mathbf{1}\{REDACT\}$ in the last three columns indicate that redacting firms have a 2.5% higher likelihood of conducting an SEO in the next year, 3.5% higher in two years, and 4.7% higher in three years, relative to firms which do not redact their filings. Thus, not only are redacting firms rewarded by investors with improved liquidity, these firms also have higher equity returns and are more likely to issue equity in the future. These findings further support the hypothesis that investors regard redaction as a positive credible signal by investors.

[30] In untabulated results, we consider shorter event windows (two, five, and 10 weeks), and relax our choice to exclude the event week and surrounding week before and after. The results, sometimes weaker and sometimes stronger in statistical significance than those tabulated, largely obtain in direction and economic magnitudes.

In Table 5 Panel B, we measure the effect of redaction on institutional ownership. We use three proxies for institutional ownership. $InstOwn$ refers to the percentage of outstanding shares owned by institutions (entities that file 13/Fs), $\log(InstOwn)$ refers to the logarithm of $InstOwn$, and $\#Block$ refers to the number of institutional block-owners (13/F filers who own a stake of 5% or more) that own shares in the filing firm. The OLS estimates in Panel B indicate that redaction is associated with significantly greater levels of institutional ownership for filing firms relative to those who do not redact in the six and 12 months post-filing.[31] Thus, rather than deterring institutional owners, redaction signals the presence of a valuable technology and is associated with an expansion in their holdings.

The results on institutional ownership, however, raise the concern that while redaction may prevent retail investors, and other parties who rely on public disclosures, from learning about the value of the licensed technologies, redaction may not prevent institutional owners from accessing this information. In particular, institutional owners may have access to private shareholder meetings with management in which the contents of redacted filings is divulged. By this logic, redaction is no longer a signal of value given that existing institutional investors may have other means of accessing the information withheld from the public. The probable higher economic value of redacted contracts would explain why existing investors expand their holdings, which in turn would be reflected in higher overall levels of institutional ownership post-filing.

We rule out this explanation in two ways. First, we show that the positive liquidity and ownership effects of redaction are partially offset for firms with more concentrated institutional ownership. The rate of leakage of redacted information through private shareholder meetings is likely to be higher for firms with more concentrated institutional ownership. Second, we find a zero to negative effect of redaction on the concentration of institutional ownership post-filing. We measure institutional ownership concentration by computing the Herfindahl-Hirschman Index of institutional ownership and we denote this variable as HHI. The results from the first three columns of Table 6 show that, if anything, liquidity worsens at the margin for firms which redact and have more concentrated institutional ownership relative to those which do not redact, post-filing. The statistically insignificant coefficient on $\mathbf{1}\{REDACT\} \times HHI$ in the fourth column indicates that there is no particular rise in institutional ownership post-filing for firms with more concentrated ownership. Finally, the fifth column shows that redaction is associated with a decline in HHI, and the sixth column suggests that this decline is, if anything, driven by firms with ex ante more concentrated ownership. Thus, it does not appear to be the case that pursuant to information obtained in private shareholders meetings, existing owners expand their holdings in redacting firms. Instead, new institutional owners without access to these meetings buy into redacting firms at a greater rate than do existing investors, which leads to a rise in the overall level of institutional ownership, but a decline in its concentration.

[31] We repeat this analysis in the matched sample. While we do not report these results for sake of brevity, they are almost identical in statistical and economic significance.

TABLE 6 ABOUT HERE

The combined results from this section provide counter-evidence to the logic that redaction is perceived negatively by investors. While redaction mechanically deprives investors of information, the ability of firms to signal the value of their IP by redacting is recognized and rewarded by investors through improved stock liquidity, higher stock returns, greater likelihood of equity issuance, and increased institutional ownership.

4.2 Redaction and Future Innovation

The next set of consequences we consider relates to future innovation activity. In Table 7 Panel A, we display regressions in which the dependent variables are various measures of patenting at a one-year-ahead horizon with respect to the filing date.

TABLE 7 ABOUT HERE

The first four columns of Panel A contain full-sample OLS estimates, and the last four columns contain matched sample estimates. We use two measures of patenting, borrowed from Kogan et al. (2017). $FNPATS$ refers to the total count of all patents granted to a given firm in a given year. TSM refers to the total dollar value of patents granted to a given firm in a given year, estimated from the stock market reaction to the patent grants. We take the logarithm of one plus $FNPATS$ and TSM respectively at horizons of one, two, and three years post-filing as the outcome variables in our regressions. The three-year horizon is significant as it suggests that the firm is not merely hiding completed and imminent innovations, but also a future pipeline of work that is not yet realized. Our outcome of interest is, therefore, the logarithm of a firm's innovation output, which can be interpreted as an intensive margin test. However, untabulated extensive margin tests also produce statistically reliable and directionally similar results.[32] In the regressions of patenting outcomes, we also control for the total prior patent stock of the filing firm to account for the fact that patenting activity is highly persistent.

Across all specifications, $\mathbf{1}\{REDACT\}$ loads significantly and positively, demonstrating that, relative to nonredacting firms, those firms which strategically redact a given contract experience increased patenting in the one year post-filing, measured in terms of both a raw count of patents granted, and stock market value generated by these patents grants. The effect obtains in both the matched sample and through OLS estimates, controlling for contract- and firm-level characteristics, including the firm's prior patenting and intangible capital stock. Furthermore, when we allow for asymmetric effects between the licensor and the licensee, the licensor seems to have a stronger marginal effect in terms of yet higher future patenting, as seen in the even-numbered columns. This reflects the role of

[32] Additionally, we rerun our innovation regressions scaling both measures, $FNPATS$ and TSM by total assets of the filing firm. Results are substantively similar in both significance and magnitude.

the licensor as the entity generating the IP. A licensee may build follow-on-IP, or simply utilize this innovation in a downstream application.

The magnitudes of the estimated effects are striking in their economic significance. The mean value of $\log(FNPATS)_{t+1}$ is 0.611, and the estimate from the first column of 0.257 implies that redacting firms exhibit an increase in patenting representing around 42% of the mean, relative to firms which do not redact in the one year post-filing. Similarly, the estimated effect of redaction from the third column is approximately 45% of the mean value of $\log(TSM)_{t+1}$. The magnitudes from the matched sample are similar to those obtained from OLS estimates. In unreported analysis, we also rerun our tests scaling the total stock market value accumulation (TSM) by assets, controlling for its lag. The inferences are similar in terms of statistical precision and quantitative magnitude. We do this in accordance with Kogan et al. (2017) who posit that the total amount of patenting activity may simply be a correlate of the size of the firm. Although we control for many orders of size through prior patent stock, assets, and the cumulant of intangible capital K_{int}, as a final direct robustness check we find our results are not mechanically driven by a partial correlation between redaction and size.

In Table 7 Panel B, we show regression estimates that measure the effects of strategic redaction on patenting at a longer horizon of two and three years after the filing date. The estimates of the coefficients on $\mathbf{1}\{REDACT\}$ are of similar statistical and economic significance as those in Table 7 Panel A. The OLS estimates suggest that redacting firms' stock market implied values of innovation in the two and three years post-filing are greater than those of fully-disclosing firms' by 29% and 38%, respectively, of the unconditional means. Redacting firms thus seem to follow a sustained higher innovation path post-filing when compared with nonredacting firms. This pattern points to a clear motive for firms to redact: They may protect their future pipeline of innovations from being expropriated.

To the extent that patenting is an incomplete measure of firm-level innovation activity, we also consider the effect of strategic redaction on the growth of various forms of intangible capital as defined in Peters and Taylor (2017). We consider two measures. First, the stock of knowledge capital ($KKnow$), and second, the stock of off-balance sheet capital ($KOffBS$). These variables are provided through WRDS, and are defined in Peters and Taylor (2017) as disaggregated components of their intangible capital stock measure. Knowledge capital, or $KKnow$, is calculated as the accumulation of R&D expenses using the perpetual inventory method, and off-balance sheet capital, or $KOffBS$, includes the accumulation of R&D expenses and a portion of SG&A expenses through the perpetual inventory method. These measures incorporate that portion of a firm's intangible assets that does not appear on the balance sheet. Results shown in Table 8 indicate a positive and significant coefficient of $\mathbf{1}\{REDACT\}$ on both measures, suggesting that, post-filing, redacting firms accumulate more intangible capital than do nonredacting firms. Thus, redaction is associated with significantly positive real effects in terms of both innovation effort and output.

<u>TABLE 8 ABOUT HERE</u>

On the whole, strategic redaction of an IP licensing agreement predicts significantly greater future innovation post-filing based on both patent and nonpatent measures. These combined results provide compelling evidence in favor of our second hypothesis, i.e., allowing high-quality technology firms to redact enables them to preserve comparative advantage and prevent competitive entry into their technology space, thereby protecting and expediting future innovation activity.

4.3 Signal Jamming

To illustrate the limits of the conditions under which our signaling equilibrium holds, we study the cross-sectional effects of variables that may reduce ("jam") the credibility of the redaction signal. These tests also help us rule out the alternative that superior firm fundamentals are driving the market's positive response to redaction. Under the signal jamming story, because of the inference problem faced by investors, less credible firms which redact will have a negative marginal capital market effect of redaction despite being no worse in terms of future realizations of its innovation activity. Under the fundamentals story, however, credibility should not be associated with a negative capital market reaction to redaction. Since the future innovation levels of less credible redacting firms are not lower than those of firms which choose not to redact, neither should be the capital market response to their redaction decision.

We consider two proxies for credibility, one, product market competition, and two, prior redaction tendency. In highly competitive industries, investors face greater difficulty in determining whether it is the desire to protect valuable IP that is driving the firm's decision to redact or whether it is a desire to avoid competition; therefore, product market competition affects the credibility of a firm's redaction signal. In other words, since competition and IP value are strategic substitutes from the perspective of the signaling firm, a given investor's ability to infer high technology value from the redaction signal is confounded. With no reason to reward low-value IP in a competitive industry, the stronger the competition the greater the likelihood that the redaction signal is jammed, and that redaction will be associated with a negative capital market reaction.

Prior redaction tendency will affect the credibility of a firm's marginal redaction signal if investors believe that the supply of high-value IP a given firm can produce is finite. Thus, if investors observe the same firm repeatedly redacting, their subjective assessment of the likelihood of the marginal redacted IP being truly valuable would decrease and the likelihood that the firm is "crying wolf" would increase.

First, we analyze whether the degree of product market competition that a firm faces affects its ability to credibly signal through redaction. We proxy for product market competition employing two text-based measures, textual similarity ($TSimm$) introduced by Hoberg and Phillips (2016) and product market fluidity ($ProdMktFluid$) developed by Hoberg, Phillips, and Prabhala (2014). Textual similarity captures the degree of closeness between a firm's product descriptions and those of its rivals. Product market fluidity captures the tendency of new words in a firm's product description to also appear in competitors' descriptions in following years, with higher values corresponding to a greater intensity of product market rivalry. In Table 9, we measure competition through textual similarity.

TABLE 9 ABOUT HERE

The results in Table 9 Panel A show that the coefficients on the interaction term ($\mathbf{1}\{REDACT\} \times Post \times TSimm$ in the case of event-time analysis and $\mathbf{1}\{REDACT\} \times TSimm$ in the case of the panel regressions) are directionally negative in the cases of turnover and institutional ownership, and positive in the cases of price impact, percentage spread, and Amihud's illiquidity measure, showing that redaction is associated with a marginally negative effect on liquidity and institutional ownership. Table 9 Panel B reinforces the signal-jamming explanation by demonstrating that redacting firms in more competitive industries produce similar levels of innovation in the future relative to firms which do not redact. In Table A.5.2 of the Internet Appendix, we measure competition through product market fluidity. The results generally confirm that redacting firms in more competitive industries face worse liquidity relative to nonredactors post-filing and are not worse in terms of future innovation.

Second, we analyze whether firms lose their ability to credibly signal high-value technology through redaction if they have frequently redacted in the past. We measure prior redaction tendency by counting the number of distinct occasions prior to the current filing date on which the filing firm filed at least one redacted IP licensing agreement. We define the number of distinct occasions as $\#PriorRedactOccasions$.[33] The results are presented in Table 10.

TABLE 10 ABOUT HERE

In Panel A, we consider the effect on capital market outcomes. The coefficients on the interaction of $\#PriorRedactOccasions$ with $\mathbf{1}\{REDACT\}$ are directionally consistent for all four measures of liquidity and statistically significant for three out of four measures, suggesting that serial redactors experience deteriorated liquidity post-filing, when they redact, when compared with currently redacting firms which had not redacted as frequently in the past. In addition, the marginal effect of redaction on $InstOwn$ for firms with higher past redaction tendency is negative and statistically significant. Thus, the effect of redaction on liquidity and institutional ownership is concave, as firms can credibly signal high value technology only up to a point. Repeated, indiscriminate redaction seems to be viewed negatively by investors.

[33] We use this as our preferred measure of prior redaction tendency to avoid issues arising from the fact that often, firms file multiple IP licensing agreements (some of which may be redacted) on a single filing occasion. An alternative measure, would count the number (or proportion) of total redacted filings across these occasions. However, such a measure runs the risk of double counting (and thus overstating) the firm's true redaction propensity. Nevertheless, untabulated analysis using variations of the prior redaction tendency measure, such as those which count individual filings and not occasions yield substantially similar results in both sign and significance.

Crucially for the signal jamming story, less credible firms that redact should see no declines in future innovation than would less credible firms which do not redact. Table 10 Panel B confirms this conjecture. The coefficient on the interaction terms are statistically insignificant in three out of four measures and close to zero in magnitude across all measures of innovation. Prior redaction history therefore appears to have little or no effect on the future innovation of redacting firms relative to those firms which fully disclose. This evidence is consistent with the signal jamming mechanism and inconsistent with a story driven by fundamentals. It also illustrates the limits of the conditions under which firms can credibly signal the value of the licensed IP through redaction.

4.4 Incentives to Redact: Evidence From a Natural Experiment

We employ the passage of the Food and Drug Administration Amendments Act (FDAAA) in 2007 as a natural experiment to test the effect of a reduction in firms' incentives to redact information as a signal of valuable IP. Section 801 of the FDAAA expanded the requirements of firms to disclose information regarding their clinical trials to the government portal ClinicalTrials.gov, which consequently mandated the disclosure of their technologies irrespective of quality.[34] Our use of this experiment helps address two potential endogeneity concerns. First, the benefits to redaction could arise from protecting contents that are not IP, such as contract terms. Second, the response of investors could be a reaction to the firms' type, not their signaled technology.

We make one plausible assumption in this analysis: The FDAAA does not affect the filing of material agreements for firms producing pharmaceutical innovations requiring clinical trials relative to other pharmaceutical innovations, in the post-2007 period.[35] Now that firms can no longer fully conceal their clinical trials, their licensing agreements transmit their plans to commercialize these innovations further. Thus, if we observe capital market responses turning negative on the margin, we can infer the average positive effect of redaction is partly due to concealment of a specific innovation.

Our identification strategy relies on classifying pharmaceutical licensing contracts (as labeled by RoyaltyStat's analysts) into those involving clinical trials and those not involving clinical trials. While this information is not directly observable, pharmaceutical contracts that license IP whose de-

[34] For more information, see https://clinicaltrials.gov/ct2/manage-recs/fdaaa.

[35] One concern may be the FDAAA coincides with the financial crisis. It may be that firms suffer financially during this period, and their financial constraints result in lower quality innovations, and this constraint applies particularly to conducting clinical trials. Our experimental design helps to sidestep this concern because we study only material agreements. Financial constraints would affect the hurdle rate for starting a clinical trial, i.e., whether an agreement enters into our sample, but it is only a material agreement if it crosses this hurdle rate. Thus, we are only looking at the choice made by firms of disclosure among these sets of innovations that are significant enough to commercialize in a material agreement. Moreover, we are not aware of any evidence for clinical trials that pass the hurdle rate for commercialization being more capital intensive than any other pharmaceutical innovations *that also pass the hurdle rate for commercialization*, which we use as a control group. Finally, we argue this bias may go against the marginal effect we find in the following sense: The hurdle rate during bad times is higher (Campello, Graham, and Harvey, 2010), therefore project quality across the board should be higher. Therefore, redacted trials should of higher quality, resulting in more positive marginal effects.

velopment require that clinical trials customarily set up milestone payments timed to the completion of various stages of such trials (Giordano-Coltart and Calkins, 2007; Stewart, Allison, and Johnson, 2001). We label as "treated" those pharmaceutical contracts which include milestone payments, and our "control" group consists of pharmaceutical contracts not including milestone payments. We then estimate the following quadruple-differences specification:

$$\begin{aligned} Y_{i,j,l,t+k} = \quad & \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} \times \mathbf{1}\{PHARMA\}_{i,j,l,t} \times \mathbf{1}\{MILESTONE\}_{i,j,l,t} \times Post_t^{2007} \\ & + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t} \end{aligned} \tag{3}$$

where Y refers to an outcome variable measured at the horizon $t+k$, where t is the contract filing date, and k refers to months (years) post-filing for monthly (annual) outcome variables. $\mathbf{1}\{REDACT\}$ refers to a dummy variable for whether a given contract i, filed at date t, by firm j, with underlying technology belonging to SIC 3-digit industry l, was redacted, $\mathbf{1}\{PHARMA\}_{i,j,l,t}$ is a dummy variable for whether contract i is classified as being in the pharmaceutical industry by RoyaltyStat analysts, $\mathbf{1}\{MILESTONE\}_{i,j,l,t}$ is a dummy variable for whether contract i contains milestone payments, and $Post_t^{2007}$ is a dummy variable that takes the value of 1 in all years after, but not including, 2007. $\mathbf{X}$ is a vector of firm- and contract-level controls, and includes all stand-alone and interacted combinations of the variables described in the previous lines. We include fixed effects for the SIC 3-digit industry assigned to the technology underlying a given contract, α_l, and year-fixed effects, α_t.[36]

Our coefficient of interest is β_1, which represents a DIDIDID estimator of the marginal difference in outcome Y between firms filing redacted pharmaceutical contracts in the post-2007 period with milestone payments relative to firms filing redacted pharmaceutical contracts in the post-2007 period without milestone payments. We report results from the above specification in Table 11.

TABLE 11 ABOUT HERE

Results in the first four columns indicate that in the post-2007 period, redaction leads to comparatively worse capital market outcomes in terms of liquidity and institutional ownership for treated firms filing pharmaceutical licenses involving milestone payments relative to those firms filing pharmaceutical licenses which do not contain milestone payments. The economic magnitude of the difference in outcomes between firms filing redacted pharmaceutical contracts with and without milestone payments are indicated by the values in red, below the main regression coefficient estimates in Table 11. They are computed by summing up the coefficient estimates for all permutations of the interactions between $\mathbf{1}\{REDACT\}_{i,j,l,t}$, $\mathbf{1}\{PHARMA\}_{i,j,l,t}$, $\mathbf{1}\{MILESTONE\}_{i,j,l,t}$, and $Post_t^{2007}$ (which are omitted for brevity), and assuming that all continuous control variables take their sample mean value. We

[36] We are able to estimate the specification in the presence of industry fixed effects as these are based upon the SIC 3-digit industry of the technology, whereas the dummy $\mathbf{1}\{PHARMA\}$ is an independent classification of the technology as being in the pharmaceutical industry by RoyaltyStat analysts (and does not perfectly overlap with any single SIC 3-digit industry).

do this for the treated and control groups for each outcome variable, compute the difference between the two groups, and scale this by the sample mean of each variable, to obtain the percentage values displayed in the row titled "Treated–Control (% of Sample Mean)."

The effect magnitudes are similar to those of the difference between redacted and un-redacted contracts estimated in Table 3. For instance, the difference in turnover between treated and control firms implied by the estimate in the first column of Table 11 is –2.733, or 29% of the sample mean turnover. Similarly, the difference in turnover between redacting and fully-disclosing firms in the second column of Table 3 Panel A is 1.798 (or 19% of the sample mean turnover). This finding is consistent with our signaling argument that an exogenous decrease in the incentive to redact leads to a reduction in the signaling value of redaction, and therefore translates to a negative capital market response. The results reported in Table 11 are robust to various choices of capital market outcome measures, outcome horizons, and fixed effects.

We can conjecture two possible interpretations for a relation between a reduced capital market response and firms with high-type technology opting to redact. If the same high-type firms redact on average, a dampened capital market response indicates that the market expects lower future cash flows from the licensed innovation. However, a different interpretation would be that, since the benefits to redaction are smaller vis-à-vis the costs, fewer firms choose to redact when they possess a high-type technology. Both interpretations are consistent with the claim that the signal conveys the value of an innovation, and its signaling value being diminished. However, the latter interpretation would be consistent with lower future innovation on the margin as firms who choose to redact possess lower quality innovations. In line with this, we show in Table 11 Column 5 that treated firms have significantly lower patenting activity after filing a redacted contract relative to control firms in the post-2007 period. Taken together, the findings from this natural experiment provide causal evidence for the signaling interpretation of redaction, and corroborate the associations established in the previous two subsections.[37]

4.5 Robustness Tests

The first set of robustness tests addresses the possibility that time-varying industry trends may be driving redaction as well as the positive capital market effects accruing to those firms relative to nonredactors. We therefore rerun our main empirical specifications with industry-by-year-fixed effects and add additional controls for liquidity outcomes including market capitalization, squared market capitalization, and turnover. Results in Table A.5.3 of the Internet Appendix show that our inferences on the positive effect of redaction on liquidity, institutional ownership, and future innovation obtain with these additional fixed effects and controls.

[37]We test the identifying assumption underlying this analysis by checking whether treated and control firms display parallel trends in all outcome variables. Results from these tests, which are omitted for the sake of brevity, confirm that the parallel trends assumption is not violated for any of the outcome variables reported in reported in Table 9.

In the second set of tests, we vary the main specification. Instead of using the level of the outcome variable Y as our dependent variable, we use the change in the outcome (e.g. ΔY). Our results continue to obtain under this specification, as seen in Table A.5.4 of the Internet Appendix.

In the third set of tests we include firm-fixed effects. This test helps dispel the concern that inherently better firms are driving both market recognition of the firm being of superior type, redaction and future innovation. By including firm-fixed effects, we are essentially comparing the market reaction to redaction in the subset of firms with multiple IP licensing agreement filings for which some are redacted and some disclosed. The results in Table A.5.5 of the Internet Appendix suggest that, holding time-invariant firm-specific factors constant, redaction continues to be associated with improved liquidity and superior future innovation within-firm for licensing firms. We acknowledge, however, that these results cannot be conclusive, as many firms do not appear in the sample more than a handful of times.

5 Conclusion

In this paper, we argue that nondisclosure serves as a credible signal of the value of a firm's technology. By developing a novel licensing database, we present evidence through multiple capital market outcomes to demonstrate that strategic nondisclosure is rewarded rather than penalized by investors. Redacting firms subsequently increase their innovation activities relative to the firms that fully disclose. Cross-sectional variation indicates that signaling equilibrium can be sustained only if the market can cleanly infer the firm has a technology worth hiding, which is confounded by repeated redaction or in highly competitive industries, where the modus operandi is to always conceal information. The natural experiment confirms that one benefit of signaling is the ability to hide innovation.

We emphasize that external validity for the signaling equilibrium we observe is uncertain. We posit no conclusion regarding the average tradeoff entrepreneurs make between innovation and disclosing to prospective stakeholders, whether investors, business partners, or employees. Our findings argue for an average effect observed in filings of significant commercialization deals of IP made by public corporations, which are disciplined by their boards of directors, auditors, the SEC, and ultimately by long-term shareholder responses. In another setting, the absence of one or more of these protections may result in nondisclosure's being punished rather than rewarded. Entrepreneurs raising capital through crowdfunding, for example, are likely to lack auditors and a reputable board of directors, and are certain to be subjected to weaker SEC oversight. In contrast, venture capital fund general partners may be able to solicit funding from limited partners based on their reputation and internal governance alone, without divulging the proprietary information of their principal investments.

The contrast between these three settings indicates that certain fixed costs may be required to establish a framework in which strategic nondisclosure of innovative activity is consistently rewarded. Public markets and, potentially, venture capital firms currently pay these costs, while crowdfunding entities do not. As a result of these conditions, however, the penalties of nondisclosure may be more severe in this setting.

Finding the right balance is an urgent concern. Our results invite discussion regarding a potential solution: Between the extremes of voluntary disclosure and mandatory disclosure, which in principle favor the firm and the investor respectively, mandatory disclosure with judicious exceptions emerges as a viable resolution.

References

Abadie, Alberto, and Jann Spiess, 2016, Robust post-matching inference, *Working Paper* .

Akcigit, Ufuk, Murat Alp Celik, and Jeremy Greenwood, 2016, Buy, keep, or sell: Economic growth and the market for ideas, *Econometrica* 84, 943–984.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Anton, James J, and Dennis A Yao, 2002, The sale of ideas: Strategic disclosure, property rights, and contracting, *The Review of Economic Studies* 69, 513–531.

Arrow, Kenneth, 1962, Economic welfare and the allocation of resources for invention, in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, 609–626 (Princeton University Press).

Bhattacharya, Sudipto, and Jay R Ritter, 1983, Innovation and communication: Signalling with partial disclosure, *The Review of Economic Studies* 50, 331–346.

Boone, Audra L, Ioannis V Floros, and Shane A Johnson, 2016, Redacting proprietary information at the initial public offering, *Journal of Financial Economics* 120, 102–123.

Brown, James R, and Gustav Martinsson, 2018, Does transparency stifle or facilitate innovation?, *Management Science* .

Campello, Murillo, and Janet Gao, 2017, Customer concentration and loan contract terms, *Journal of Financial Economics* 123, 108–136.

Campello, Murillo, John R Graham, and Campbell R Harvey, 2010, The real effects of financial constraints: Evidence from a financial crisis, *Journal of Financial Economics* 97, 470–487.

Chien, Colleen V, Jorge L Contreras, Carol A Corrado, Stuart JH Graham, Deepak Hegde, Arti K Rai, and Saurabh Vishnubhakat, 2016, Comment to the SEC in support of the enhanced disclosure of patent and technology license information, *Available at SSRN 2815618* .

Choudhary, Preeti, Kenneth J Merkley, and Katherine Schipper, 2018, Auditors' quantitative materiality judgments: Properties and implications for financial reporting reliability .

Connelly, Brian L, S Trevis Certo, R Duane Ireland, and Christopher R Reutzel, 2011, Signaling theory: A review and assessment, *Journal of Management* 37, 39–67.

Darrough, Masako N, and Neal M Stoughton, 1990, Financial disclosure policy in an entry game, *Journal of Accounting and Economics* 12, 219–243.

Doidge, Craig, Kathleen M Kahle, G Andrew Karolyi, and René M Stulz, 2018, Eclipse of the public corporation or eclipse of the public markets?, *Journal of Applied Corporate Finance* 30, 8–16.

Doidge, Craig, G Andrew Karolyi, and René M Stulz, 2017, The US listing gap, *Journal of Financial Economics* 123, 464–487.

Feltham, Gerald A, and Jim Z Xie, 1992, Voluntary financial disclosure in an entry game with continua of types, *Contemporary Accounting Research* 9, 46–80.

Gans, Joshua, Scott Stern, and Jane Wu, 2016, The Foundations of Entrepreneurial Strategy.

Gans, Joshua S, David H Hsu, and Scott Stern, 2008, The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays, *Management Science* 54, 982–997.

Gans, Joshua S, and Scott Stern, 2003, The product market and the market for 'ideas': Commercialization strategies for technology entrepreneurs, *Research Policy* 32, 333–350.

Gertner, Robert, Robert Gibbons, and David Scharfstein, 1988, Simultaneous signalling to the capital and product markets, *The RAND Journal of Economics* 173–190.

Giordano-Coltart, Jennifer, and Charles W Calkins, 2007, Best practices in patent license negotiations, *Nature Biotechnology* 25, 1347.

Glosten, Lawrence R, and Paul R Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.

Grossman, Sanford J, 1981, The informational role of warranties and private disclosure about product quality, *The Journal of Law and Economics* 24, 461–483.

Grossman, Sanford J, and Oliver D Hart, 1980, Disclosure laws and takeover bids, *The Journal of Finance* 35, 323–334.

Hadlock, Charles J, and Joshua R Pierce, 2010, New evidence on measuring financial constraints: Moving beyond the kz index, *Review of Financial Studies* 23, 1909–1940.

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

Hall, Bronwyn H, and Josh Lerner, 2010, The Financing of R&D and Innovation, *Handbook of the Economics of Innovation* 1, 609–639.

Hegde, Deepak, 2014, Tacit knowledge and the structure of license contracts: Evidence from the biomedical industry, *Journal of Economics & Management Strategy* 23, 568–600.

Hegde, Deepak, and Hong Luo, 2017, Patent publication and the market for ideas, *Management Science* 64, 652–672.

Hoberg, Gerard, and Gordon Phillips, 2016, Text-based network industries and endogenous product differentiation, *Journal of Political Economy* 124, 1423–1465.

Hoberg, Gerard, Gordon Phillips, and Nagpurnanand Prabhala, 2014, Product market threats, payouts, and financial flexibility, *The Journal of Finance* 69, 293–324.

Imai, Kosuke, Gary King, and Elizabeth Stuart, 2008, Misunderstandings among experimentalists and observationalists about causal inference, *Journal of the Royal Statistical Society, Series A* 171, 481–502.

Kerr, William R, and Ramana Nanda, 2015, Financing innovation, *Annual Review of Financial Economics* 7, 445–462.

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

Kyle, Albert S, 1985, Continuous auctions and insider trading, *Econometrica* 1315–1335.

Lerner, Josh, and Amit Seru, 2017, The use and misuse of patent data: Issues for corporate finance and beyond, Technical report, National Bureau of Economic Research.

Ljungqvist, Alexander, Lars Persson, and Joacim Tåg, 2018, The incredible shrinking stock market: On the political economy consequences of excessive delistings, Technical report.

O'Hara, Maureen, 1995, *Market Microstructure Theory*, volume 108 (Blackwell Publishers Cambridge, MA).

Peters, Ryan H, and Lucian A Taylor, 2017, Intangible capital and the investment-Q relation, *Journal of Financial Economics* 123, 251–272.

Shapiro, Carl, 1985, Patent licensing and R&D rivalry, *The American Economic Review* 75, 25–30.

Stewart, Jeffrey J, Peter N Allison, and Ronald S Johnson, 2001, Putting a price on biotechnology, *Nature Biotechnology* 19, 813.

Thakor, Richard T, and Andrew W Lo, 2017, Optimal financing for R&D-intensive firms, Technical report, National Bureau of Economic Research.

Tian, Xiaoli Shaolee, and Miaomiao Yu, 2018, Redact when competitors act–examining the threat of rivals' product portfolio modifications .

Varner, Thomas R, 2011, An economic perspective on patent licensing structure and provisions, *Business Economics* 46, 229–238.

Verrecchia, Robert E, 1983, Discretionary disclosure, *Journal of Accounting and Economics* 5, 179–194.

Verrecchia, Robert E, and Joseph Weber, 2006, Redacted disclosure, *Journal of Accounting Research* 44, 791–814.

Vorhies, James Brady, 2005, The new importance of materiality, *Journal of Accountancy* 199, 53.

Wagenhofer, Alfred, 1990, Voluntary disclosure with a strategic opponent, *Journal of Accounting and Economics* 12, 341–363.

Figure 1: Distribution of Contracts by Technology Industry

This figure plots the distribution of contracts across industries, where the industry classification scheme used is the 42 technology industries defined by RoyaltyStat analysts. Specifically, each contract is classified into 1 (or more) of the 42 technology industries based upon the technology space that best describes underlying IP being licensed; this determination is made by RoyaltyStat analysts.

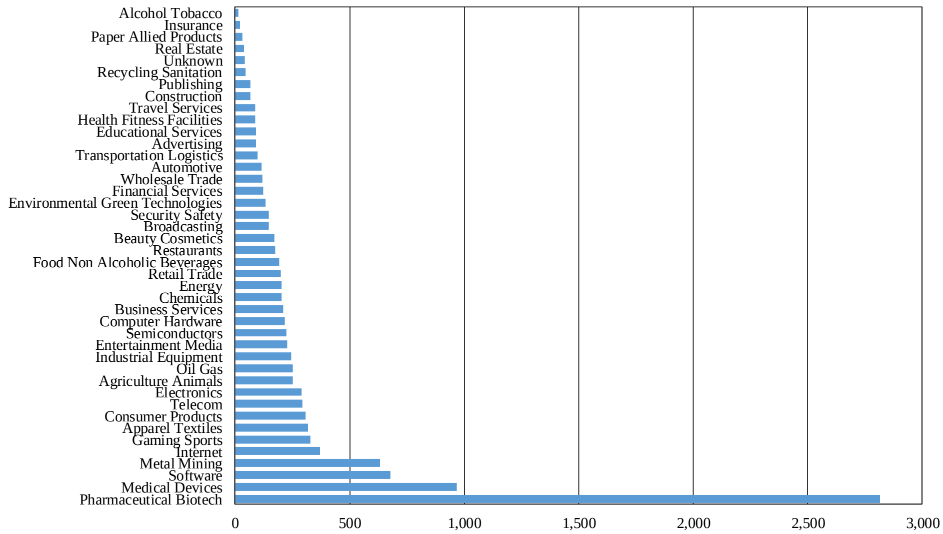


Figure 2: Distribution of Redacted vs. Unredacted Contracts Over Time

This figure plots the yearly (time-series) distribution of total contracts and the proportion which are redacted.

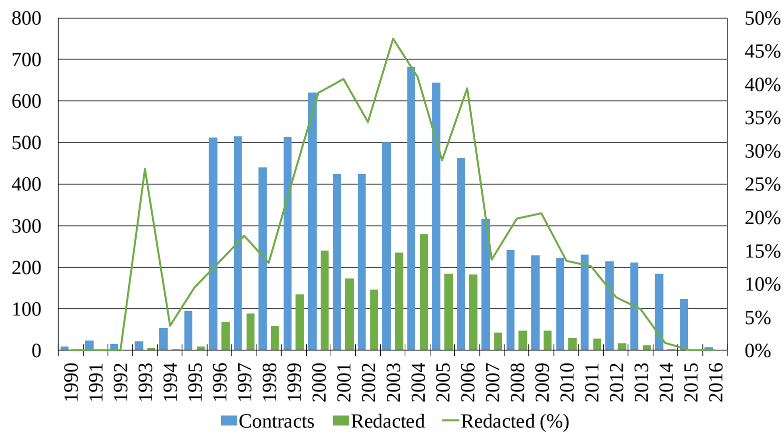
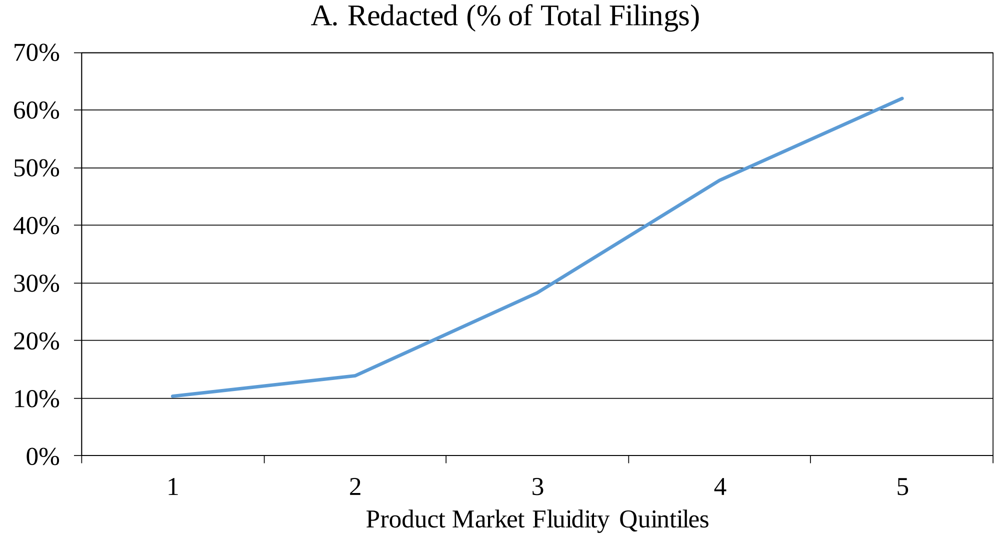


Figure 3: Relationship Between Redaction Rate and Proprietary Costs of Disclosure

Panel A plots the relation between quintiles of product market fluidity (i.e., the tendency of new words in a firm's product description to also appear in competitors' in following years), as calculated by Hoberg and Phillips (2016) and the fraction of redacted contracts among total filed contracts in each quintile. Panel B plots the relation between quintiles of total five-year patent citations for contracts whose underlying IP consists of patents and the fraction of redacted contracts among total filed contracts in each quintile.



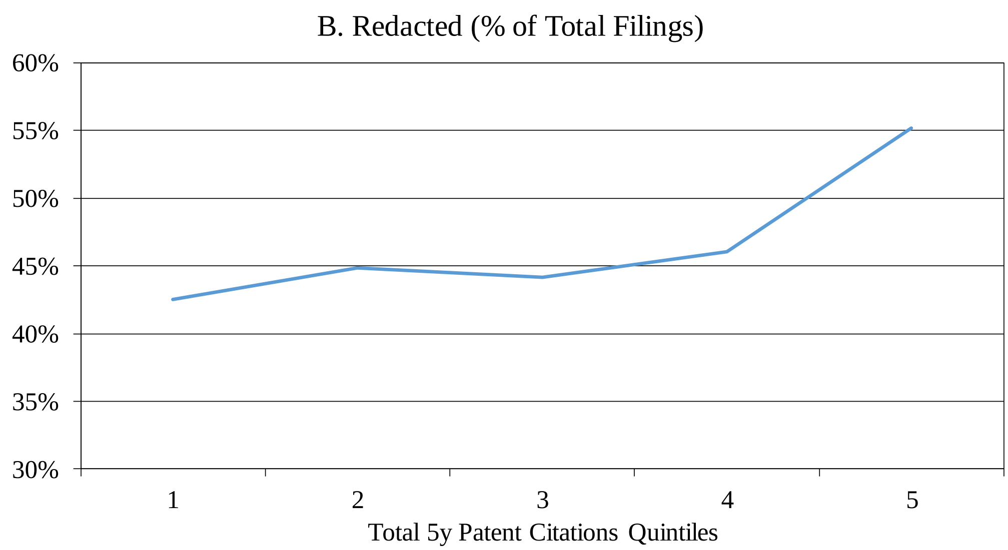


Table 1: Contract-Level Summary Statistics

Panel A: Distribution of Key Contract Terms

	License Fee ($)	Royalty Rate (%)	Duration (years)
N	2,418	7,972	6,625
Mean	4,280,652	11.49	19.744
St. Dev.	23,279,623	14.363	26.049
Min	0	0.005	0
Q1	2	0.5	0
Q5	5,000	1	1
Q25	45,792	3	5
Q40	100,000	5	10
Median	250,000	6	17
Q60	500,000	8	20
Q75	1,500,000	12	20
Q95	15,000,000	50	99
Q99	60,000,000	70	99
Max	550,000,000	100	500

Panel B: Distribution of Contracts by IP type

IP Type	Frequency (%)	IP Type	Frequency (%)
Is Amendment	13.40%	Is Patent	54.00%
Is Asset Purchase	0.00%	Is Process	14.70%
Is Consulting	0.90%	Is Proprietary	2.30%
Is Copyrights	15.70%	Is Research	6.30%
Is Cross-License	2.10%	Is Services	3.30%
Is Distribution	4.20%	Is Shares	1.40%
Is Franchise	3.50%	Is Show-How	16.80%
Is Know-How	30.50%	Is Software	14.70%
Is Liabilities	2.20%	Is Sublicense	2.90%
Is Marketing	1.30%	Is Supply	3.60%
Is Mineral	9.30%	Is Technology	30.80%
Is Option	4.60%	Is Trademark	26.00%

Table 1: Contract-Level Summary Statistics (continued)

Panel C: Proportion of Redacted Clauses		
	Redacted (%)	
Contract Term	Extensive Margin	Intensive Margin
Royalty Rate	98.44%	98.44%
Identity of Patents (Applications)	41.18%	80.95%
Identity of Patents (Grants)	46.51%	74.29%
License Fee	60.00%	73.33%
Identity of Non-Patent Intangible	55.32%	65.91%
Milestone Payments	55.81%	61.54%
Other Fees	55.81%	60.00%
Identity of Product	38.33%	46.55%
Termination Clauses	39.62%	39.62%
Duration	29.82%	29.82%
Territory of License	11.11%	12.96%
Signatories	6.35%	6.35%

Table 2: Firm-Level Summary Statistics

In this table, we compare the summary statistics of the firms appearing in our sample with the summary statistics of firms appearing in the overall Compustat sample over the same sample period, excluding financial and utility firms.

	RoyaltyStat Sample						Full Compustat Sample					
Variable	**N**	**Mean**	**St. Dev.**	**Q25**	**Median**	**Q75**	**N**	**Mean**	**St. Dev.**	**Q25**	**Median**	**Q75**
Firm Characteristics												
$Tobin's\ Q$	4,902	2.505	2.995	0.461	1.279	3.296	83,056	3.172	4.051	1.097	1.886	3.479
$Leverage$	6,368	0.302	0.532	0	0.091	0.362	112,754	0.220	0.258	0.0231	0.154	0.331
$Assets$	6,439	310.732	2036.404	7.048	27.625	98.180	113,262	2534.086	8090.788	48.507	242.123	1116.498
$MktCap$	3,914	928.116	5915.451	42.633	128.147	371.521	80,530	2088.369	12203.585	39.173	153.819	709.580
$Tangibility$	6,381	0.191	0.218	0.042	0.105	0.251	112,309	0.224	0.232	0.039	0.139	0.333
$R\&D/Assets$	6,439	0.337	1.272	0	0.100	0.337	113,262	0.066	0.341	0	0	0.044
$CAPEX/Assets$	6,310	0.060	0.089	0.009	0.028	0.067	104,260	0.0554	0.0681	0.0132	0.0335	0.0693
$NI/Assets$	6,407	–0.777	1.527	–0.786	–0.262	0.013	112,910	–0.06311	0.309	–0.039	0.016	0.063
ROA	6,370	0.708	0.849	0.074	0.429	1.047	112,888	0.950	0.847	0.270	0.787	1.363
$NetSales$	6,407	182.009	1063.144	0.762	10.507	55.166	112,920	1,353.099	3,854.712	29.579	134.642	710.303
Age	6,642	12.897	8.380	7	10	16	114,458	19.752	14.040	9	15	26
Innovation Outcomes												
$FNPATS$	3,731	4.280	35.475	0	0	2	82,143	2.926	12.998	0	0	0
TSM	3,731	38.933	393.682	0	0	1.804	82,143	33.175	198.044	0	0	0
K_{int}	6,591	238.965	3209.708	5.999	24.980	80.053	114,379	623.01560	2,075.893	11.534	50.443	254.877
$KKnow$	6,591	113.918	1485.622	0	6.655	39.264	114,379	74.30018	74.300	0	0	16.903
$KOffBS$	6,591	163.392	1990.594	4.954	20.176	66.941	114,379	316.031	1,034.283	7.373	32.142	145.431
Capital Market Outcomes												
$Turnover$	3,970	9.548	26.146	1.578	3.938	9.165	82,141	4.680	8.212	0.93729	2.596	5.659
$PctSpread$	3,921	0.012	0.019	0.002	0.006	0.014	81,422	0.017	0.033	0.002	0.008	0.019
$Illiq$	3,741	5.147	21.903	0.056	0.402	2.413	82,141	3.850	11.827	0.020	0.020	1.817
$InstOwn$	3,614	32.981	28.082	8.395	25.937	52.902	78,742	0.367	0.289	0.106	0.312	0.600
$\#Block$	3,623	1.356	1.501	0	1	2	79,310	1.374	1.444	0	1	2
$InstOwnHHI$	3,623	0.254	0.246	0.078	0.159	0.341	79,310	0.240	0.255	0.061	0.135	0.318
Competition Measures												
$ProdMktFluid$	3,672	9.295	4.155	5.975	8.977	12.412	82,549	7.395	3.767	4.557	6.727	9.579
$TSimm$	3,613	8.628	8.812	1.507	3.960	15.342	90,698	9.531	18.453	1.303	2.331	6.475

Table 3: Redaction and Future Liquidity

Table 3 displays estimated regression coefficients and standard errors (in brackets) corresponding to the effects of redaction on future liquidity. The baseline specification takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. The Y variables consist of $Turnover$, $Spread$, and $Illiq$. For each contract i filed by firm j at time t, we compute we compute daily turnover for firm j's stock, which is the end-of-day price multiplied by the volume of shares traded. We average this across all days in the three, six, and 12 months after filing time t to obtain the variable $Turnover$. Next we compute daily bid-ask spread for firm j's stock, which is the closing ask minus the closing bid scaled by the end-of-day price, and average this across all days in the three and six months after filing time t to obtain the variable $Spread$. Finally, we calculate daily values of Amihud's illiquidity measure (Amihud, 2002) for each day in the next three and six months after time t. We do this by by computing the ratio of the daily absolute return to the (dollar) trading volume on that day, for firm j's stock. We average this value across all days in the three and six months after filing time t, to obtain the variable $Illiq$. We take the logarithm of the latter two variables to form $\log(Spread)$ and $\log(Illiq)$. Panel A contains OLS coefficient estimates from specifications which include firm- and contract-level controls (specified in Appendix A.4), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level.

Panel A: OLS Estimates

	$Turnover_{t+3}$	$Turnover_{t+6}$	$Turnover_{t+12}$	$\log(Spread)_{t+3}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+3}$	$\log(Illiq)_{t+6}$
$\mathbf{1}\{REDACT\}$	1.990***	1.798***	1.252**	–0.042*	–0.072***	–0.205***	–0.158***
	[0.631]	[0.558]	[0.480]	[0.023]	[0.019]	[0.043]	[0.042]
$Level_{t-1}$	0.158	0.133	0.099	0.737***	0.668***	0.635***	0.675***
	[0.117]	[0.097]	[0.076]	[0.026]	[0.026]	[0.030]	[0.028]
Industry+Year FE	Y	Y	Y	Y	Y	Y	Y
Contract Controls	Y	Y	Y	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y	Y	Y	Y
N	3,668	3,665	3,637	3,618	3,616	3,668	3,665
R^2	0.152	0.148	0.191	0.849	0.844	0.775	0.775

Panel B contains estimates from a matched sample. Matching is done with replacement based on a propensity score which is a function of $SmallYoung$ and $\log(KInt)$, plus exact matching on SIC1 of the technology, filing dates within a three-year window of each other, and underlying IP consisting of patents. Standard errors in the matching analysis are clustered at the level of the matches as recommended by Abadie and Spiess (2016).

Panel B: Matching Estimates

	$Turnover_{t+3}$	$Turnover_{t+6}$	$Turnover_{t+12}$	$\log(Spread)_{t+3}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+3}$	$\log(Illiq)_{t+6}$
$\mathbf{1}\{REDACT\}$	1.140***	0.960***	1.109***	–0.048**	–0.087***	–0.210***	–0.229***
	[0.374]	[0.320]	[0.277]	[0.021]	[0.023]	[0.043]	[0.045]
$Level_{t-1}$	0.262**	0.211**	0.159**	0.844***	0.782***	0.735***	0.705***
	[0.114]	[0.090]	[0.071]	[0.013]	[0.019]	[0.021]	[0.022]
Contract Controls	N	N	N	N	N	N	N
Firm Controls	Y	Y	Y	Y	Y	Y	Y
N	2,412	2,411	2,403	2,387	2,387	2,412	2,411
R^2	0.248	0.247	0.238	0.871	0.840	0.777	0.761

Table 4: Redaction and Future Liquidity in Event Time

Table 4 displays regression coefficients, and standard errors in brackets, of weekly liquidity measures around filing events. We use three liquidity measures as outcome variables. The quoted spread is defined as the difference between the prevailing market quoted bid and ask prices, divided by the midpoint between these two prices. The effective spread is defined as twice the absolute value of the difference between the actual trade price and the midpoint of the market quote (i.e., between the quoted bid price and the quoted ask price), divided by the midpoint between these two prices based. The price impact of trades, also known as Kyle's Lambda, is measured by performing regressions of the absolute return over a given time period on dollar volume. The price impact is defined as the slope from this regression. These slopes are averaged to calculate daily and annual measures of price impact in an analogous way to the spread and turnover obtained from CRSP. Higher values of price impact correspond to a given magnitude of trading being associated with a greater movement in returns, and thus price impact is seen as a measure of stock illiquidity (in fact, this is an intra-day high frequency analogue of the (Amihud, 2002) measure which we compute from CRSP). In the first three columns, these liquidity variables are winsorized at the 1% level. In the second three columns, they are z-scored within each event. In the final set of columns, we take the logarithm of these variables. *Post* is a variable that takes the value of 1 for all observations that fall in the 15-week window following the filing date of the disclosure (excluding the week of the event itself, to control for confounding filings), and 0 otherwise. All specifications include event fixed effects (essentially demeaning by event), and the standard errors are clustered by event.

	$QSpread_w$	$EffSpread_w$	$PrcImpact_w$	$\overline{QSpread}$	$\overline{EffSpread}$	$\overline{PrcImpact}$
$\mathbf{1}\{REDACT\} \times Post$	–7.301**	–0.036	–2.969***	–9.431***	–11.660***	–5.506**
	[3.609]	[0.029]	[1.132]	[3.021]	[3.490]	[2.164]
$Post$	1.879	–0.024	0.484	–2.683	–13.111***	–1.617
	[2.471]	[0.020]	[0.807]	[1.816]	[2.106]	[1.318]
Event FE	Y	Y	Y	Y	Y	Y
N	99,864	99,504	97,209	99,842	99,479	97,167
R^2	0.732	0.810	0.453	0.005	0.011	0.004

	$\log(QSpread)$	$\log(EffSpread)$	$\log(PrcImpact)$
$\mathbf{1}\{REDACT\} \times Post$	–4.515***	–1.696**	–8.548***
	[1.422]	[0.733]	[2.713]
$Post$	–1.568*	–1.408***	1.300
	[0.853]	[0.484]	[1.705]
Event FE	Y	Y	Y
N	99,864	99,504	97,209
R^2	0.876	0.889	0.255

Table 5: Redaction and Future Equity Returns, Issuance and Institutional Ownership

Table 5 displays estimated regression coefficients and standard errors (in brackets) corresponding to the effects of redaction on equity returns, equity issuance and institutional ownership. The baseline specification takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. The Y variables consist of Ret, α^{FF5}, $\mathbf{1}\{Issuance\}$, $InstOwn$, and $NumBlockOwn$. We compute equity returns by compounding monthly raw returns, starting from the filing date, for the subsequent three and six months to form the variables Ret_{t+3} and Ret_{t+6} respectively. We also compute the monthly alphas net of the Fama–French five-factor model, and compound these over the subsequent three and six months starting from the filing date to form the variables α^{FF5}_{t+3} and α^{FF5}_{t+6} respectively. $\mathbf{1}\{Issuance\}_t$ is a dummy variable that takes the value of 1 if a given firm conducts a secondary equity offering (SEO), in which capital is raised through an equity issuance, in a given year t. $InstOwn$ refers to the percentage of outstanding shares owned by institutions (entities that file 13/Fs), and $\#Block$ refers to the number of institutional block-owners (13/F filers who own a stake of 5% or more) that own shares in the filing firm. We take the logarithm of $InstOwn$ to form $\log(InstOwn)$. Panels A and B contain OLS coefficient estimates from specifications with firm- and contract-level controls (specified in Appendix A.4), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level. In the returns regressions, we additionally control for the logarithm of the lagged market capitalization of the filing firm, squared market capitalization, and turnover. In the issuance regressions, in lieu of controlling for the prior level of the Y variable, Y_{t-1}, we control instead for the prior level of the logarithm of the price-to-book ratio.

Panel A: Redaction and Equity Returns and Equity Issuance

	Ret_{t+3}	Ret_{t+6}	α^{FF5}_{t+3}	α^{FF5}_{t+6}	$\mathbf{1}\{Issuance\}_{t+1}$	$\mathbf{1}\{Issuance\}_{t+2}$	$\mathbf{1}\{Issuance\}_{t+3}$
$\mathbf{1}\{REDACT\}$	0.040***	0.057	0.005	0.006	0.025**	0.035*	0.047**
	[0.0169]	[0.0502]	[0.0083]	[0.0160]	[0.013]	[0.019]	[0.021]
$Level_{t-1}$	–0.202***	–0.206**	0.103**	0.148***			
	[0.0360]	[0.0566]	[0.0488]	[0.0525]			
$\log\left(\frac{P}{B}\right)_{t-1}$					0.019***	0.027***	0.033***
					[0.007]	[0.008]	[0.007]
Industry + Year FE	Y	Y	Y	Y	Y	Y	Y
Contract Controls	Y	Y	Y	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y	Y	Y	Y
N	3,125	3,091	3,113	3,081	3,400	3,361	3,318
R^2	0.174	0.192	0.132	0.124	0.149	0.177	0.184

Panel B: Redaction and Institutional Ownership

	$InstOwn_{t+6}$	$\log(InstOwn)_{t+6}$	$\log(InstOwn)_{t+12}$	$\#Block_{t+6}$
$\mathbf{1}\{REDACT\}$	0.405	0.089***	0.146***	0.433***
	[0.540]	[0.03]	[0.045]	[0.112]
$Level_{t-1}$	0.913***	0.728***	0.589***	0.671***
	[0.008]	[0.037]	[0.033]	[0.026]
Industry + Year FE	Y	Y	Y	Y
Contract Controls	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y
N	3,477	3,477	3,377	3,484
R^2	0.894	0.786	0.698	0.611

Table 6: Redaction and Institutional Ownership Concentration

Table 6 displays estimated regression coefficients and standard errors (in brackets) corresponding to the effects of redaction on the concentration of institutional ownership. The baseline specification takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. The Y variables consist of $Turnover$, $Spread$, $Illiq$, $InstOwn$ and HHI. For each contract i filed by firm j at time t, we compute we compute daily turnover for firm j's stock, which is the end-of-day price multiplied by the volume of shares traded. We average this across all days in the three, six, and 12 months after filing time t to obtain the variable $Turnover$. Next we compute daily bid-ask spread for firm j's stock, which is the closing ask minus the closing bid scaled by the end-of-day price, and average this across all days in the three and six months after filing time t to obtain the variable $Spread$. Finally, we calculate daily values of Amihud's illiquidity measure (Amihud, 2002) for each day in the next three and six months after time t. We do this by computing the ratio of the daily absolute return to the (dollar) trading volume on that day, for firm j's stock. We average this value across all days in the three and six months after filing time t, to obtain the variable $Illiq$. We take the logarithm of the latter two variables to form $\log(Spread)$ and $\log(Illiq)$. $InstOwn$ refers to the percentage of outstanding shares owned by institutions (entities that file 13/Fs). We take the logarithm of $InstOwn$ to form $\log(InstOwn)$. HHI is the the Herfindahl-Hirschman Index of institutional ownership (with lower values representing less concentration). This table contains OLS coefficient estimates from specifications which include firm- and contract-level controls (specified in Appendix A.4), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level.

	$Turnover_{t+6}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+6}$	$\log(InstOwn)_{t+12}$	HHI_{t+6}	HHI_{t+6}
$\mathbf{1}\{REDACT\}$	2.405***	–0.171***	–0.207***	0.0167	–0.010*	0.011
	[0.884]	[0.043]	[0.071]	[0.078]	[0.005]	[0.011]
$\mathbf{1}\{REDACT\} \times HHI$	–8.167***	0.589***	0.428	0.581		–0.103
	[3.005]	[0.181]	[0.261]	[0.410]		[0.063]
HHI	–1.132	0.254***	1.392***	–0.694***	0.707***	0.726***
	[1.188]	[0.079]	[0.155]	[0.173]	[0.023]	[0.023]
$Level_{t-1}$	0.126	0.656***	0.614***	0.545***		
	[0.093]	[0.024]	[0.029]	[0.036]		
Industry + Year FE	Y	Y	Y	Y	Y	Y
Contract Controls	Y	Y	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y	Y	Y
N	3,547	3,514	3,547	3,377	3,484	3,385
R^2	0.225	0.850	0.792	0.703	0.735	0.737

Table 7: Redaction and Future Patenting

Table 7, Panel A displays estimated regression coefficients and standard errors (in brackets) corresponding to the effects of redaction on future patenting at the 1 year ahead horizon. The baseline specification takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. We use two measures of patenting, borrowed from Kogan et al. (2017). $FNPATS$ refers to the raw count of all patents granted to a given firm in a given year. TSM refers to the total dollar value of patents granted to a given firm in a given year, estimated from the stock market reaction to the patent grants. We take the logarithm of one plus each of these variables to form $\log(FNPATS)$ and $\log(TSM)$ respectively. In these regressions, we also control for the total prior patent stock of the filing firm. The first four columns contain OLS coefficient estimates from specifications with firm- and contract-level controls (specified in Appendix A.4), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level. The second three columns contain estimates from a matched sample. Matching is done with replacement based on a propensity score which is a function of $SmallYoung$ and $\log(KInt)$, plus exact matching on SIC1 of the technology, filing dates within a three-year window of each other, and underlying IP consisting of patents. Standard errors in the matching analysis are clustered at the level of the matches as recommended by Abadie and Spiess (2016).

Panel A: Redaction and Future Innovation (1 year ahead)

	$\log(FNPATS)_{t+1}$		$\log(TSM)_{t+1}$		$\log(FNPATS)_{t+1}$		$\log(TSM)_{t+1}$	
$\mathbf{1}\{REDACT\}$	0.257***	0.142***	0.349***	0.194***	0.241***	0.166***	0.416***	0.287***
	[0.049]	[0.049]	[0.079]	[0.068]	[0.035]	[0.043]	[0.052]	[0.062]
$\mathbf{1}\{REDACT\} \times FilerLicensor$		0.267***		0.361***		0.129*		0.258***
		[0.058]		[0.079]		[0.070]		[0.106]
$FilerLicensor$		0.056**		0.076**		0.156***		0.203***
		[0.028]		[0.038]		[0.054]		[0.076]
$FilerLicensor \times \log(TotalPriorPatents)$						0.046**		0.022
						[0.021]		[0.034]
$\log(TotalPriorPatents)$	0.307***	0.301***	0.425***	0.416***	0.339***	0.309***	0.458***	0.435***
	[0.016]	[0.016]	[0.030]	[0.028]	[0.013]	[0.017]	[0.019]	[0.024]
Analysis	OLS	OLS	OLS	OLS	Matching	Matching	Matching	Matching
Industry + Year FE	Y	Y	Y	Y	N	N	N	N
Contract Controls	Y	Y	Y	Y	N	N	N	N
Firm Controls	Y	Y	Y	Y	Y	Y	Y	Y
N	3,696	3,696	3,696	3,696	2,436	2,436	2,436	2,436
R^2	0.563	0.571	0.562	0.568	0.413	0.434	0.411	0.426

Table 7 (continued): Redaction and Future Patenting

Panel B: Redaction and Future Innovation (2—3 years ahead)				
	$\log(TSM)_{t+2}$	$\log(TSM)_{t+3}$	$\log(TSM)_{t+2}$	$\log(TSM)_{t+3}$
$\mathbf{1}\{REDACT\}$	0.230**	0.318***	0.316***	0.412***
	[0.113]	[0.120]	[0.058]	[0.063]
$\log(TotalPriorPatents)$	0.399***	0.352***	0.496***	0.474***
	[0.030]	[0.037]	[0.022]	[0.025]
Analysis	OLS	OLS	Matching	Matching
Industry + Year FE	Y	Y	N	N
Contract Controls	Y	Y	N	N
Firm Controls	Y	Y	Y	Y
N	3,301	2,948	2,131	1,930
R^2	0.547	0.524	0.417	0.375

Table 8: Redaction and Future Intangible Capital Accumulation

Table 8 displays estimated regression coefficients and standard errors (in brackets) corresponding to the effects of redaction on future intangible capital accumulation. The baseline specification takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. We consider the following two outcome variables, Y. First, the level of knowledge capital ($KKnow$), and second, the level of off-balance sheet capital ($KOffBS$). These variables are provided through WRDS, and are defined in Peters and Taylor (2017) as follows. $KKnow$, is calculated as the accumulation of R&D expenses using the perpetual inventory method, and $KOffBS$, includes the accumulation of R&D and a portion of SG&A expenses through the perpetual inventory method. The first two columns contain OLS coefficient estimates from specifications with firm- and contract-level controls (specified in Appendix A.4), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level. The latter two columns contain estimates from a matched sample. Matching is done with replacement based on a propensity score which is a function of $SmallYoung$ and $\log(KInt)$, plus exact matching on SIC1 of the technology, filing dates within a three-year window of each other, and underlying IP consisting of patents. Standard errors in the matching analysis are clustered at the level of the matches as recommended by Abadie and Spiess (2016).

	$KKnow_{t+2}$	$KOffBS_{t+2}$	$KKnow_{t+2}$	$KOffBS_{t+2}$
$\mathbf{1}\{REDACT\}$	0.125***	0.114***	0.197***	0.199***
	[0.029]	[0.023]	[0.019]	[0.017]
Analysis	OLS	OLS	Matching	Matching
Industry + Year FE	Y	Y	N	N
Contract Controls	Y	Y	N	N
Firm Controls	Y	Y	Y	Y
N	3,811	5,037	2,606	2,967
R^2	0.389	0.372	0.312	0.312

Table 9: Cross-Sectional Analysis Conditional on Competition

Table 9 displays estimated regression coefficients and standard errors (in brackets) corresponding to the cross-sectional effects of firms' product market competition. In the first two columns of Panel A, we display regression coefficients (and standard errors in brackets) of weekly liquidity measures around filing events. The outcome variables are as defined in previous tables. In the next four columns of Panel A and all columns of Panel B, we present results from the cross-sectional variant of our baseline OLS panel regression specification. The regression takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \beta_3 Z_{i,j,l,t} + \beta_4 \mathbf{1}\{REDACT\}_{i,j,l,t} \times Z_{i,j,l,t} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. Textual similarity, or $TSimm$ captures the degree of closeness between a firm's product descriptions and those of its rivals, and was downloaded from the Hoberg-Phillips data library. For details on its construction, see Hoberg and Phillips (2016). All variables are as previously defined. The regressions are estimated by OLS with firm- and contract-level controls (specified in Appendix A.4), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level.

Panel A: Redaction and Capital Market Outcomes

	$\overline{PrcImpact}$	$\log(PrcImpact)$	$Turnover_{t+6}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+6}$	$\log(InstOwn)_{t+12}$
$\mathbf{1}\{REDACT\}$			1.762***	–0.137***	–0.369***	0.160**
			[0.642]	[0.047]	[0.078]	[0.068]
$\mathbf{1}\{REDACT\} \times Post$	–11.059***	–12.830**				
	[4.162]	[5.566]				
$\mathbf{1}\{REDACT\} \times Post \times TSimm$	0.548*	0.421				
	[0.286]	[0.361]				
$\mathbf{1}\{REDACT\} \times TSimm$			–0.109***	0.005*	0.017***	–0.004
			[0.027]	[0.003]	[0.004]	[0.003]
Analysis	Event-Time	Event-Time	OLS	OLS	OLS	OLS
Event FE	Y	Y	N	N	N	N
Industry + Year FE	N	N	Y	Y	Y	Y
Contract + Firm Controls	N	N	Y	Y	Y	Y
N	70,069	70,088	3,242	3,202	3,242	3,041
R^2	0.012	0.257	0.355	0.852	0.780	0.720

Panel B: Redaction and Innovation Outcomes

	$\log(FNPATS)_{t+1}$	$\log(FNPATS)_{t+2}$	$\log(TSM)_{t+1}$	$\log(TSM)_{t+2}$
$\mathbf{1}\{REDACT\}$	0.262**	0.250*	0.298***	0.279
	[0.115]	[0.141]	[0.159]	[0.183]
$\mathbf{1}\{REDACT\} \times TSimm$	–0.004	–0.006	–0.005	–0.010
	[0.005]	[0.007]	[0.007]	[0.008]
Analysis	OLS	OLS	OLS	OLS
Industry + Year FE	Y	Y	Y	Y
Contract + Firm Controls	Y	Y	Y	Y
N	2,459	2,166	2,459	2,166
R^2	0.590	0.593	0.604	0.604

Table 10: Cross-Sectional Analysis Conditional on Prior Redaction

Table 10 displays estimated regression coefficients and standard errors (in brackets) corresponding to the cross-sectional effects of firms' past redaction tendency. In the first two columns of Panel A, we display regression coefficients (and standard errors in brackets) of weekly liquidity measures around filing events. The outcome variables and empirical specification are as in previous tables. In the next four columns of Panel A and all columns of Panel B, we present results from the cross-sectional variant of our baseline OLS panel regression specification. The regression takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \beta_3 Z_{i,j,l,t} + \beta_4 \mathbf{1}\{REDACT\}_{i,j,l,t} \times Z_{i,j,l,t} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. We measure prior redaction tendency by the variable $\#PriorRedactOccassions$, which counts the number of distinct occasions, prior to the current filing date, on which at least one redacted licensing agreement was filed by the filing firm. All other variables are as previously defined. The regressions are estimated by OLS with firm- and contract-level controls (specified in Appendix A.4), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level.

Panel A: Redaction and Capital Market Outcomes

	$\overline{PrcImpact}$	$\log(PrcImpact)$	$Turnover_{t+6}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+6}$	$\log(InstOwn)_{t+12}$
$\mathbf{1}\{REDACT\}$			1.925***	–0.073**	–0.229***	0.126***
			[0.632]	[0.029]	[0.052]	[0.028]
$\mathbf{1}\{REDACT\} \times Post$	–6.400**	–12.950***				
	[2.847]	[3.707]				
$\mathbf{1}\{REDACT\} \times Post \times \#PriorRedactOccasions$	1.523	3.755***				
	[1.501]	[1.449]				
$\mathbf{1}\{REDACT\} \times \#PriorRedactOccasions$			–0.255*	0.012	0.041**	–0.044***
			[0.139]	[0.014]	[0.018]	[0.010]
Analysis	Event-Time	Event-Time	OLS	OLS	OLS	OLS
Event FE	Y	Y	N	N	N	N
Industry + Year FE	N	N	Y	Y	Y	Y
Contract + Firm Controls	N	N	Y	Y	Y	Y
N	84,662	84,700	3,665	3,616	3,665	3,477
R^2	0.012	0.259	0.148	0.845	0.775	0.787

Panel B: Redaction and Innovation Outcomes

	$\log(FNPATS)_{t+1}$	$\log(FNPATS)_{t+2}$	$\log(TSM)_{t+1}$	$\log(TSM)_{t+2}$
$\mathbf{1}\{REDACT\}$	0.254***	0.198**	0.357***	0.254
	[0.056]	[0.088]	[0.083]	[0.110]
$\mathbf{1}\{REDACT\} \times \#PriorRedactOccassions$	–0.002	–0.004	–0.014	–0.047*
	[0.024]	[0.020]	[0.028]	[0.025]
Analysis	OLS	OLS	OLS	OLS
Industry + Year FE	Y	Y	Y	Y
Contract + Firm Controls	Y	Y	Y	Y
N	3,696	3,301	3,696	3,301
R^2	0.563	0.549	0.562	0.547

Table 11: Incentives to Redact — Evidence From the FDAAA (2007)

Table 11 displays estimated regression coefficients and standard errors (in brackets) corresponding to the DIDIDID specification $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} \times \mathbf{1}\{PHARMA\}_{i,j,l,t} \times \mathbf{1}\{MILESTONE\}_{i,j,l,t} \times Post_t^{2007} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. $\mathbf{1}\{REDACT\}$ refers to a dummy variable for whether a given contract i, filed at date t, by firm j, with underlying technology belonging to SIC 3-digit industry l, was redacted, $\mathbf{1}\{PHARMA\}_{i,j,l,t}$ is a dummy variable for whether contract i is classified as being in the pharmaceutical industry by RoyaltyStat analysts, $\mathbf{1}\{MILESTONE\}_{i,j,l,t}$ is a dummy variable for whether contract i contains milestone payments, and $Post_t^{2007}$ is a dummy variable that takes the value of 1 in all years 2008 onwards. All outcome variables are as previously defined, and we include firm- and contract-level controls (specified in Appendix A.4), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level.

	$Turnover_{t+6}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+6}$	$\log(InstOwn)_{t+12}$	$\log(FNPATS)_{t+2}$
$\mathbf{1}\{REDACT\}$	2.671***	–0.059	–0.222**	0.079	0.273***
	[0.811]	[0.050]	[0.092]	[0.071]	[0.100]
$\mathbf{1}\{REDACT\} \times \mathbf{1}\{PHARMA\} \times \mathbf{1}\{MILESTONE\} \times Post^{2007}$	–10.456**	0.785***	2.017***	–0.289	–1.226***
	[5.189]	[0.217]	[0.627]	[0.380]	[0.386]
$Treated - Control$ (% of Sample Mean)	–29%	3%	9%	–1%	–8%
Industry + Year FE	Y	Y	Y	Y	Y
Contract Controls	Y	Y	Y	Y	Y
Firm Controls	Y	Y	Y	Y	Y
N	3,665	3,616	3,665	3,377	3,686
R^2	0.156	0.845	0.775	0.700	0.670

Internet Appendix

Appendix A.1: Supplementary Figures

Figure A.1.1: Sample Redacted Licensing Agreement

This figure is an example of a portion of a redacted licensing agreement between Genome Therapeutics Corporation and bioMérieux Incorporated dated September 30, 1999. In the section of the agreement that we reproduced in this figure (the total length of such agreements often exceeds 30 pages), payment terms including royalty rates, sublicensing fees etc. are redacted.

6.5 ROYALTIES PAYABLE BY EACH PARTY AND ITS AFFILIATES AND SUBLICENSEES.

6.5.1 ROYALTIES ON NET SALES OF BMI PRODUCTS AND GTC PRODUCTS.

(a) BMI shall pay royalties to GTC based on the Net Sales of a BMI Product at the rate, *...*, specified in EXHIBIT G hereto. *...*.

In the event a BMI Product should subsequently become a BMI Blockbuster Product, the Parties shall determine, within *...* days following the BMI Product's change of status, a new royalty rate to be applied to the BMI Blockbuster Product, it being understood that such rate shall be *...*.

(b) GTC, its Affiliates and sublicensees shall pay royalties to BMI based on the Net Sales of a GTC Product at a rate to be determined in good faith by mutual agreement of the parties with respect to each GTC Product within *...* days following the First Commercial Sale of the GTC Product. The royalty rate for Net Sales of each GTC Product shall be determined *...*.

6.5.2 SUBLICENSES. (i) BMI and GTC shall, with respect to any sublicenses by BMI or GTC hereunder, pay to the other party *...* percent (*...*) of all consideration received

For the purposes of calculating the royalty payments to be made in United States dollars, GTC shall multiply the sales made in countries other than the United States by the appropriate royalty rate before converting such sum into United States dollars at the exchange rate used by GTC for reporting such sales for United States financial statement purposes.

BMI shall use, for the purposes of calculating the royalty payments to be made in United States dollars, the following rules:

(a) For the period from January 1, 1999 until *...*:

For the conversion of all currencies, other than United States dollars, into United States dollars, the amount shall be first converted into the Euro using the official spot rate published by the European Central Bank on the last business day of the period to which the payment of royalties relates and then, the amount so obtained in the Euro shall be converted into United States dollars applying the official rate published by the European Central Bank on the last business day of the period to which the payment of royalties relates.

(b) For the period from *...* until the expiration of the contract:

(i) For the conversion of the Euro into United States dollars, the amount shall be converted applying the official rate published by the European Central Bank on the last business day of the period to which the payment of royalties relates.

(ii) For the conversion of currencies other than the Euro into United States dollars, the amount shall be first converted into the Euro using the official spot rate published by the European Central Bank on the last business day of the period to which the payment of royalties relates and then, the amount so obtained in the Euro shall be converted into

Appendix A.2: Sample Construction and Selection

In this appendix, we provide details on the sample construction and selection procedure. We first concord the data to financial databases, then disambiguate the economic role of the filer in the transaction (i.e. classify the filer as the licensor or the licensee). Finally, we filter observations according to our sample requirements. Appendix Table A.2.2 reports the sample sizes at each stage of the filtering process.

To concord the data with financial databases, we first use RoyaltyStat's provided *gvkeys*, followed by SEC *CIK* codes and filer names. RoyaltyStat maintains filer names and *CIK* codes as reported in the EDGAR log, and the point-in-time *gvkey*. The remainder of firms, for which such identifiers are unavailable, are firms that RoyaltyStat could not link to Compustat due to omissions in their coverage, or because while they do have publicly filed securities, they are not covered by Compustat. The company suggests that this is most often the case with penny stocks. This translates to about 8,500 contracts with a filer *gvkey* referring to a firm in Compustat North America.

Second, we augment RoyaltyStat's coverage of filers using the filer name and the filer *CIK*. Given a filer *CIK*, we coalesce possible names for the company using point-in-time *CIK* and the Capital IQ "Helper" table available on WRDS. Given a filer name, we standardize frequently occurring business words ("company" to "co", "limited" to "ltd") and perform exact name matching on company names coalesced from Compustat, CRSP, Capital IQ and other financial databases that we have at our disposal. This yields approximately 1,000 additional matches for the identifiers not covered by RoyaltyStat.

Third, we concord each contract to the identifier of the *filer* and then identify whether the filer was the licensor or the licensee in the transaction. In order to ensure data integrity, we keep only contracts in which the filer can clearly be disambiguated as a licensor or licensee in the transaction. To perform this disambiguation, we first match the filer with the name, *CIK* or *gvkey* provided by RoyaltyStat for the licensor or licensee. Through RoyaltyStat-provided data of 9,678 records, we are initially unable to assign the filer as the licensor or licensee in 2,824 cases. These remainder cases are either due either to ambiguity, or to RoyaltyStat's efforts to identify the licensor and licensee to this date still being in progress.

To match the remainder records, RoyaltyStat provided us a list of company names in its database. We first apply programmatic rules to match the transaction parties to the filers and then hand-match the remainder.[38]

Fourth, we exclude contracts that have a per-unit royalty, which RoyaltyStat flagged as a related party transaction, or which we flag as ambiguous. The final result set has 7,972 observations, though sample size varies across analyses due to vastly different data coverage across the various data sources that we use.

[38] The steps involved are as follows. First, we flag the counter-party as an individual, university or government institute; in these cases, the filer must be the opposite party. Second, we write down a list of nearly 1,500 stem words, such that if the licensor contains the stem word and the filer contains the stem word, it is very unlikely for the filer to be the licensee, and vice versa. Third, we perform exact name matching the subsidiaries of the firm based on CorpWatch. Fourth and finally, for the remainder cases, RoyaltyStat allowed us to export a list of filer names and we disambiguated (where possible) the role of the filer in relation to the economic transaction when possible. For 1,061 of 2,824, we were not able to match the filer to the licensor or the licensee. A substantial fraction of these cases are due to ambiguity and limited information particularly for transactions occurring, in some cases, over twenty years ago.

Table A.2.1: Data Fields Extracted From RoyaltyStat Data

Identifying Information	**Description**
Licensor (name, CIK, *gvkey*)	The party in possession of the intangible asset contracting to allow another party to use it in exchange for payment.
Licensee (name, CIK, *gvkey*)	The party paying the possessor of the intangible asset in exchange for the right to use the intangible for proprietary purposes.
Filer (name, CIK, *gvkey*)	The party which submitted a filing to the SEC documenting the existence of a contract. The filer is usually (but not always) identifiably the licensor or licensee.
$IsRelatedParty$	Are the parties subsidiaries of the same company? For the purpose of transfer pricing, this distinction is made because the contract does not represent open-market pricing.
Filing Information	
Filing date	The date of the filing in EDGAR.
FOIA	Was the record originally redacted, and later recovered, through the Freedom of Information Act?
Intangible information	
SIC of intangible	Based on the technology described in the SEC filing, RoyaltyStat's analysts assign the technology covered in a contract a Standard Industrial Classification code.
RoyaltyStat Industry	RoyaltyStat has a classification of 42 industries. Possible values include: Pharmaceutical Biotech, Medical Devices, and Software.
RoyaltyStat Subindustry	Each industry is classified into a number of sub-industries. For instance, the Pharmaceutical Biotech industry contains Cancer, Genetic, and Cardiovascular sub-industries.
Intangible numbers	Intangible numbers specifically mentioned in a filing. The vast majority are patents, including both applications and grants.
Description	A textual description of the contract and intangible being licensed.
Intangible property type	RoyaltyStat's proprietary categorization of intangibles in a contract. Possible values include patents, trademarks, copyrights, trade secrets, software, process know-how, land and mineral rights.
Contract features	
Duration	The number of years, if known, a royalty contract is active for. In some contracts, the duration is unknown, perpetual, or is stated but the contract contains an option to renew. In econometric analysis, if it is perpetual, it is coded as 100 years.
Agreement type	Whether the agreement can be classified as a cross-license, joint-venture, patent-related, trademark-related, and so forth. From this we create booleans for each agreement type.
Rights granted	For instance, whether the contract gives the licensee the right to sublicense
Territory agreement applies	The countries or legal jurisdictions over which the company is granted a license to use the technology. Possible values include specific countries, worldwide, cyberspace, or unknown.
Effective date	The date the contract terms become effective, if known. In some cases, the effective date was immediate and the filing date occurred with a delay, and in some cases the filing date precedes the effective date
Exclusivity	Is the contract exclusive to the party?
Payment information	
Royalty base	The accounting identity against which any royalty rate is assessed. Most commonly it is net sales. Others include cost-of-goods sold, net profits, per-yard, per square meter, per unit.
Royalty rate	In the units of the royalty base, this is the numerical payment made by the licensee.
Royalty rate currency	The currency against which per-unit royalties are assessed.
Tiered royalty	Many royalty contracts have incentive schemes whereby the rate changes at variable quantities.
Royalty rate low	When the rate is tiered, this is the lowest rate that can be charged, if stated.
License fee amount	The currency amount of the license fee.
License fee currency	The currency in which the license fee is to be paid.
Fee type	The vast majority of cases are upfront fees, but in some cases there are scheduled fees, or renewal fees.
Other payments	Booleans that flag additional payment terms (upfront fees, renewal fees, milestone payments, etc).

Table A.2.2: Sample Selection Procedure

	#	Comment
# Observations to start.	17,487	As of June 2016, this was the entire number of records available on RoyaltyStat's database.
# Obs with filer *gvkey* matched to Compustat North America.	8,774	Based on RoyaltyStat's maintained *gvkeys*, we can identify the following number of firms.
# Obs of US contracts with *gvkey* after programmatically imputing using *CIK* or name.	9,678	We augment RoyaltyStat's *gvkeys* using programmatic matching of *CIK* codes to *gvkeys* based on Capital IQ Helper. We also match based on names. If a punctuation-cleaned name is an exact match for a Capital IQ Company name, CRSP name, Compustat name, Worldscope name, which can be linked back to a *gvkey*, it is in our database.
# Obs where the identity of the filer can be disambiguated (i.e. the filer can be assigned the status of being either the licensor or the licensee), excluding related party filings.	7,972	This is our baseline sample.
# Obs of US contracts with *PERMNOs*.	6,579	Not all companies have *PERMNOs* because they are not exchange-listed, the observation is prelisting or post delisting.

Appendix A.3: Theoretical Framework

In order to motivate our empirical hypotheses, we present a simple stylized model that demonstrates the strategic role of redaction as a signal of technology quality. The setup consists of three firms, an innovative firm (whose choice of redaction we are modeling), an uninformed rival whose decision to enter the innovative firm's technology space and compete is a function of the innovative firm's redaction decision, and an investor, whose preferences and beliefs determine the market price of the innovative firm's stock. The innovative firm, which we denote as A, is endowed with a technology type $t \in \{\bar{t}, \underline{t}\}$, with $\bar{t} > \underline{t}$, where the unconditional probability $\mathbb{P}(t = \bar{t}) = \theta$ is known by the uninformed potential rival, which we denote as B, and the investor, whom we denote as I. However, the realization of the technology type, $\tilde{t}$, is known only to A.

Firm A has the choice to disclose $\tilde{t}$, and we denote this choice by $D \in \{0, 1\}$. The choice by firm A to disclose (which is exactly equivalent to the choice to not redact), $D = 1$, results in both B and I observing the reported value of t, $\hat{t}$. In order to narrow the strategy space, we assume that firm A's disclosure of $\tilde{t}$ is truthful, i.e. the reported $\hat{t} = \tilde{t}$. Thus, for the sake of convenience we simply use $\tilde{t}$ to denote firm A's disclosure. The assumption of truthful reporting is reasonable given that the disclosures in question are SEC filings, subject to auditor approvals as well as SEC review through the Division of Corporation Finance, and costly shareholder litigation if proven to be false. The choice by A to not disclose (which is exactly equivalent to the choice to redact), $D = 0$, means that both B and I are unable to observe $\tilde{t}$. In case of complete nondisclosure regardless of A's type, they both believe that $\tilde{t}$ takes its unconditional expected value $t' = \theta\bar{t} + (1 - \theta)\underline{t}$.

In addition, firm A is associated with an ex ante degree of information asymmetry, $\beta \in \{\beta_G, \beta_B\}$ where $\beta_G < \beta_B$. Firm A has $\beta = \beta_G$ with probability λ, and this is independent of its technology type t. Let the unconditional expected value of β be defined as $\beta' = \lambda\beta_G + (1 - \lambda)\beta_B$.

Our analysis proceeds in two steps. First, we consider the effect of $A's$ disclosure decision on B's entry decision, and the resulting effect on A's profits. Second, we analyze the investor's response to the product-market equilibrium. For the first step, we assume that the economic rents available in the market occupied by A, which B is considering entering, are a function of A's technology type $\tilde{t}$. Specifically, we assume that the rents are normally distributed with mean $\tilde{t}$ and variance $\frac{\sigma^2}{\tilde{t}}$, i.e. we denote the total rents available in the market as π and $\pi \sim N\left(\tilde{t}, \frac{\sigma^2}{\tilde{t}}\right)$. The rents (net of costs) earned by each firm are indicated by π^A and π^B.

We now consider Firm B. Firm B incurs a cost $c > 0$ of entry, if it chooses to enter the market. B also incurs s, a cost of search, if it wishes to enter the market when firm A does not disclose its type, i.e $D = 0$. The search cost s is a function of B's belief of A's type, which we denote t^*. We do not specify a functional form for s, only that $s\left(\underline{t}\right) = \underline{s}$ and $s\left(\bar{t}\right) = \bar{s}$, with $\bar{s} > \underline{s} > 0$, and $\frac{\partial s}{\partial t^*} > 0$. The search cost may, for example, correspond to the fact that the company B must conduct research to determine the nature of the undisclosed technology so as to imitate it. Also, it must develop its own substitutive technology whose cost increases according to the type of the technology; i.e., it is more costly to replicate a good technology than a bad technology.

How does Firm B's decision affect payoffs for A? Since A is already in the industry, it does not incur any cost. Both firms A and B are assumed to be risk-neutral profit maximizers. If B does not enter, we assume that the entire rents are received by A, that is $\pi^A = \pi$. If B enters, the two firms split the rents evenly. So, if B enters, A receives $\pi^A = \frac{\pi}{2}$. For B, if it enters when A has selected $D = 1$ (has chosen to disclose), its expected profit is $\mathbb{E}\left[\pi^B\right] = \mathbb{E}\left[\frac{\pi}{2}\right] - c$. If B enters when A has selected $D = 0$, its expected profit is $\mathbb{E}\left[\pi^B\right] = \mathbb{E}\left[\frac{\pi}{2}\right] - c - s$. Firm B only enters if, based on its beliefs on $\tilde{t}$, $\mathbb{E}\left[\pi^B\right] \geq 0$, namely $\mathbb{E}\left[\frac{\pi}{2}\right] \geq c + s$. We further assume that $\frac{\underline{t}}{2} < c$, implying that when B knows A's type is low $\left(\tilde{t} = \underline{t}\right)$, it does not enter as $\mathbb{E}\left[\frac{\pi}{2}\right] = \frac{\underline{t}}{2} - c - s < 0$. We also assume that $\frac{\bar{t}}{2} < \bar{s}$, indicating that the search cost incurred by B when A's technology is high-type is prohibitive.

To ascertain the equilibrium outcome, consider Firm A's strategies, which can be represented by pairs $\left\{D\left(\underline{t}\right), D\left(\bar{t}\right)\right\}$. For example, $\{1,1\}$ implies that firm A discloses its type both when $\tilde{t} = \underline{t}$ and $\tilde{t} = \bar{t}$.

The first strategy is $\{1,1\}$ where A discloses $\tilde{t}$ regardless of its type. In this case, B enters only when it observes $\tilde{t} = \bar{t}$. Even though B does not incur the search cost, it is still not worth entering as $\frac{\underline{t}}{2} - c < 0$. So, A receives $\underline{t}$. When $\tilde{t} = \bar{t}$, B enters and receives (in expectation) $\frac{\bar{t}}{2} - c$ and A receives $\frac{\bar{t}}{2}$. Thus, the payoffs for A to this strategy are $\left\{\underline{t}, \frac{\bar{t}}{2}\right\}$. The second strategy is $\{0,1\}$, where A discloses if its type is high $\tilde{t} = \bar{t}$, but not if its type is low. Again, in this case B only enters when it observes $\tilde{t} = \bar{t}$. B does not enter when $D = 0$ as this indicates $t = \underline{t}$ with probability 1. This is because B's potential expected profit from entry is $\frac{\underline{t}}{2} - c - \underline{s} < 0$, so it does not enter. So, A's profit is $\underline{t}$. When B observes $\tilde{t} = \bar{t}$, it enters and receives (in expectation) $\frac{\bar{t}}{2} - c$ and A receives $\frac{\bar{t}}{2}$. A's payoffs are thus $\left\{\underline{t}, \frac{\bar{t}}{2}\right\}$.

The next strategy is $\{1,0\}$, where A discloses if its type is low, $\tilde{t} = \underline{t}$, but not if its type is high. In this case, B does not enter regardless of type. The rents in the low-type market are not attractive enough, as $\frac{\underline{t}}{2} < c$, and the search cost is prohibitively high in the high-type market, as $\frac{\bar{t}}{2} < \bar{s}$. Thus, A's profits are $\left\{\underline{t}, \bar{t}\right\}$. The final strategy is $\{0,0\}$. In this case, firm B's belief on A's type is $t' = \theta\bar{t} + (1-\theta)\underline{t}$. Since $\underline{s} < s\left(t'\right) = s' < \bar{s}$, we require:

$$\theta > \underline{\theta} = \frac{2(c + s') - \underline{t}}{\bar{t} - \underline{t}}$$

such that B enters. This imposes an lower bound on the probability of A's technology being high-type, $\theta > \underline{\theta}$. This condition is more likely to hold if, first, $\left(\bar{t} - \underline{t}\right)$ is large, or, second, if the entry and search costs are not too high relative to $\underline{t}$. Specifically, since θ is a probability, we need $\underline{\theta} \geq 0$. This implies that we need $c + s' \geq \frac{\underline{t}}{2}$. We know that $\frac{\underline{t}}{2} - c < 0$ by assumption, and $s' > 0$, so, $s' > \frac{\underline{t}}{2} - c \implies c + s' > \frac{\underline{t}}{2}$ as required. If $\theta > \underline{\theta}$, B enters and A's expected payoffs are $\left\{\frac{\underline{t}}{2}, \frac{\bar{t}}{2}\right\}$. We summarize the payoffs of the game in the following table:

		$\tilde{t}=\underline{t}$	
		$D=1$	$D=0$
$\tilde{t}=\bar{t}$	$D=1$	$\left\{\frac{\bar{t}}{2},\underline{t}\right\}$	$\left\{\frac{\bar{t}}{2},\underline{t}\right\}$
	$D=0$	$\left\{\bar{t},\underline{t}\right\}$	$\left\{\frac{\bar{t}}{2},\frac{\underline{t}}{2}\right\}$

There are two Nash equilibria in pure strategies, $\{0,1\}$ and $\{1,0\}$, however only $\{1,0\}$ survives iterated elimination of weakly dominated strategies.

This analysis shows that as long as the probability θ exceeds a lower-bound value (i.e. there exists sufficiently large probability that A's technology is high type), the game has two Nash equilibria in pure strategies. By iterated elimination of weakly dominated strategies the unique surviving Nash equilibrium strategy is $\{1,0\}$, where A discloses if its technology is low type and redacts if its technology is high type. This is an essential condition for nondisclosure (redaction) to be a credible signal of type. The equilibrium is separating, and as in Gertner, Gibbons, and Scharfstein (1988), it is the nature of the product-market competition that determines whether the equilibrium is separating or pooling.

We now turn to the second part of our analysis, determining the investor's response to the product-market equilibrium. In equilibrium, it is optimal for A to redact only if $\tilde{t}=\bar{t}$ and disclose only if $\tilde{t}=\underline{t}$. The investor thus infers that $\tilde{t}=\bar{t}$ if she observes nondisclosure and directly observes that $\tilde{t}=\underline{t}$ when the technology is disclosed. The equilibrium payoff to firm A from its optimal strategy of $\{1,0\}$ is $\left\{\underline{t},\bar{t}\right\}$. We assume that the world ends after the product market competition takes place, and the investor receives firm A's total profit π^A as a liquidating dividend. As in Verrecchia (1983), the price of the firm's stock is determined by the following expression reflecting the investor's mean-variance preferences:

$$P=\mathbb{E}\left[\pi^A\right]-\beta Var\left[\pi^A\right]$$

Since the equilibrium is separating, the investor prices A's stock as:

$$P=\begin{cases}\bar{t}-\beta_G\frac{\sigma^2}{\bar{t}} & if\ \ D=0\ \ and\ \ \beta=\beta_G\\ \bar{t}-\beta_B\frac{\sigma^2}{\bar{t}} & if\ \ D=0\ \ and\ \ \beta=\beta_B\\ \underline{t}-\beta_G\frac{\sigma^2}{\underline{t}} & if\ \ D=1\ \ and\ \ \beta=\beta_G\\ \underline{t}-\beta_B\frac{\sigma^2}{\underline{t}} & if\ \ D=1\ \ and\ \ \beta=\beta_B\end{cases}$$

The expected price conditional on observing redaction, i.e. $D=0$, is $\mathbb{E}[P|D=0]=\left[\lambda\times\left(\bar{t}-\beta_G\frac{\sigma^2}{\bar{t}}\right)\right]+\left[(1-\lambda)\times\left(\bar{t}-\beta_B\frac{\sigma^2}{\bar{t}}\right)\right]=\bar{t}-\beta'\frac{\sigma^2}{\bar{t}}$, and the expected price conditional on observing $D=1$, i.e. full disclosure, is $\mathbb{E}[P|D=1]=\underline{t}-\beta'\frac{\sigma^2}{\underline{t}}$. The variance of the price conditional on observing $D=0$ is $Var[P|D=0]=\tilde{\beta}\frac{\sigma^4}{\bar{t}^2}$ and the variance of the price conditional on observing $D=1$ is $Var[P|D=1]=\tilde{\beta}\frac{\sigma^4}{\underline{t}^2}$, where $\tilde{\beta}=\lambda(1-\lambda)\left(\beta_G-\beta_B\right)^2$. As $\tilde{\beta}>0$ and $\bar{t}>\underline{t}$, this implies that $Var[P|D=0]<Var[P|D=1]$.

Our model generates two main implications (corresponding to the two hypotheses), and one cross-sectional implication (which guides the cross-sectional analysis) that we take to the data.

Implication 1 The first key implication of our stylized model is that redacting firms have a lower conditional variance of stock price as $Var[P|D = 0] < Var[P|D = 1]$. To the extent that the conditional variance of a stock's price is a proxy for the stock's illiquidity, this implies that redacting firms experience better stock liquidity than nonredacting firms.

Implication 2 The second key implication is that redaction is associated with greater innovation, as firms with high-value technology are able to preserve the economic rents from potential competitors and avoid expropriation. This is consistent with the fact that in equilibrium, the rival firm B does not enter A's market and compete.

Cross-sectional Implication If investors are able to observe β, then following a similar argument that we used to demonstrate that $Var[P|D = 0] < Var[P|D = 1]$, it can be shown that $Var\left[P|\beta_G\right] < Var\left[P|\beta_B\right]$. Thus, in the cross-section, firms which are perceived as having a worse reputation, or a greater degree of information asymmetry, by investors (i.e. firms with $\beta = \beta_B$) should experience worse stock liquidity conditional on a given disclosure choice.

Appendix A.4: Control Variable Definitions

The firm-level controls are the logarithm of total assets($\log(Assets)$), the logarithm of total intangible capital as defined by Peters and Taylor (2017) ($\log(K_{int})$), the ratio of cash to assets $\left(\frac{Cash}{Assets}\right)$, leverage ratio $\left(\frac{Debt}{Assets}\right)$, and the $SmallYoung$ index value which collapses the age and size of the firm into a single index. The contract-level controls are dummy variables which take the value of 1 if the filing firm is the licensor($FilerLicensor$), if the licensor is an individual ($LicensorIsIndividual$), a nonprofit ($LicensorIsNonProfit$), a university ($LicensorIsUniversity$), if the given contract is an amendment of a previous agreement ($IsAmendment$), if the underlying IP consists of know-how ($IsKnowHow$), and show-how ($IsShowHow$). We also include the logarithm of the duration of the contract ($\log(Duration)$), a dummy for whether the licensing rights extend worldwide ($Worldwide$), whether the agreement confers the licensee the right to sublicense the IP ($Sublicense$), and if the IP is being licensed exclusively ($Exclusive$).

Appendix A.5: Supplementary Tables

Table A.5.1: Percentage Improvement in Balance as a Result of Matching

In this table, we evaluate the improvement in comparability of the samples generated by matching by considering the percent improvement in balance for each of the matching variables, defined as $100 \times \frac{((|a|-|b|)}{|a|}$, where a is the balance statistic before and b is the balance statistic after matching for a given variable. This table shows the improvement in balance for the mean difference pre- and post-matching, as well as the median, mean, and maximum of the empirical distributions pre- and post-matching of each of the variables on which we construct the propensity score, and the score itself . $Distance$ is the propensity score. All other variables are as defined in the text.

	Mean Diff.	eQQ Med	eQQ Mean	eQQ Max
$Distance$	99.988	86.656	87.084	42.903
$SmallYoung$	70.312	30.984	48.566	16.833
$\log(KInt)$	94.277	93.664	88.578	59.638

Table A.5.2: Cross-Sectional Analysis Conditional on Competition

This table displays estimated regression coefficients and standard errors (in brackets) corresponding to the cross-sectional effects of firms' product market competition. In the first two columns of Panel A, we display regression coefficients (and standard errors in brackets) of weekly liquidity measures around filing events. The outcome variables and empirical specification are as described in Table 7. In the next four columns of Panel A and all columns of Panel B, we present results from the cross-sectional variant of our baseline OLS panel regression specification. The regression takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \beta_3 Z_{i,j,l,t} + \beta_4 \mathbf{1}\{REDACT\}_{i,j,l,t} \times Z_{i,j,l,t} + \mathbf{X}\gamma + \sum_l \alpha_l + \sum_t \alpha_t + \epsilon_{i,j,l,t}$. Product market fluidity, or $ProdMktFluid$ captures the tendency of new words in a firm's product description to also appear in competitors' descriptions in following years, with higher values corresponding to a greater intensity of product market rivalry, and was downloaded from the Hoberg-Phillips data library. For details on its construction, see Hoberg, Phillips, and Prabhala (2014). All other variables are as defined previously. The regressions are estimated by OLS with firm- and contract-level controls (specified in Appendix A.3), as well as technology-industry (SIC3) and year-fixed effects, with standard errors clustered at the technology-industry (SIC3) level.

Panel A: Redaction and Capital Market Outcomes

	$\overline{PrcImpact}$	$\log(PrcImpact)$	$Turnover_{t+6}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+6}$	$\log(InstOwn)_{t+12}$
$\mathbf{1}\{REDACT\}$			1.779*	–0.158**	–0.597***	0.172**
			[1.021]	[0.077]	[0.112]	[0.079]
$\mathbf{1}\{REDACT\} \times Post$	–25.212***	–24.277***				
	[7.026]	[8.678]				
$\mathbf{1}\{REDACT\} \times Post \times ProdMktFluid$	1.749***	1.429**				
	[0.608]	[0.717]				
$\mathbf{1}\{REDACT\} \times ProdMktFluid$			–0.122	0.007	0.038***	–0.003
			[0.077]	[0.006]	[0.009]	[0.007]
Analysis	Event-Time	Event-Time	OLS	OLS	OLS	OLS
Event FE	Y	Y	N	N	N	N
Industry + Year FE	N	N	Y	Y	Y	Y
Contract + Firm Controls	N	N	Y	Y	Y	Y
N	70,209	70,234	3,224	3,192	3,224	3,026
R^2	0.012	0.255	0.348	0.851	0.785	0.717

Panel B: Redaction and Innovation Outcomes

	$\log(FNPATS)_{t+1}$	$\log(FNPATS)_{t+2}$	$\log(TSM)_{t+1}$	$\log(TSM)_{t+2}$
$\mathbf{1}\{REDACT\}$	0.135	0.196	–0.038	0.042
	[0.160]	[0.199]	[0.246]	[0.291]
$\mathbf{1}\{REDACT\} \times ProdMktFluid$	0.009	–0.002	0.029*	0.013
	[0.011]	[0.014]	[0.016]	[0.018]
Analysis	OLS	OLS	OLS	OLS
Industry + Year FE	Y	Y	Y	Y
Contract + Firm Controls	Y	Y	Y	Y
N	2,562	2,253	2,562	2,253
R^2	0.592	0.593	0.604	0.604

Table A.5.3: Robustness With Industry-by-Year-Fixed Effects

This table displays results of specifications with industry-by-year-fixed effects, as well as additional controls including lagged market capitalization, market capitalization squared and turnover. The regression takes the form $Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \sum_t \alpha_l \times \alpha_t + \epsilon_{i,j,l,t}$. All variables are as previously defined. Standard errors for the OLS analysis (in brackets) are clustered at the technology industry level.

	$Turnover_{t+6}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+6}$	Ret_{t+6}	$\log(InstOwn)_{t+12}$	$\log(FNPATS)_{t+1}$	$\log(FNPATS)_{t+3}$	$\log(TSM)_{t+1}$	$\log(TSM)_{t+3}$
$\mathbf{1}\{REDACT\}$	2.148***	–0.089***	–0.116***	0.047	0.104*	0.266***	0.306***	0.347***	0.447***
	[0.586]	[0.019]	[0.035]	[0.050]	[0.062]	[0.051]	[0.087]	[0.079]	[0.119]
$Level_{t-1}$	0.140	0.479***	0.417***	–0.140	0.567***	0.012	–0.229	–0.175	–0.589
	[0.094]	[0.034]	[0.052]	[0.102]	[0.023]	[0.161]	[0.258]	[0.182]	[0.422]
Analysis	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Industry-by-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Contract + Firm Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	3,610	3,610	3,610	3,091	3,342	2,564	2,049	2,564	2,049
R^2	0.274	0.883	0.868	0.657	0.773	0.636	0.629	0.710	0.672

Table A.5.4: Robustness With Change Specification

This table displays results of specifications where instead of using the level of the outcome variable Y as the dependent variable, we use its first difference, ΔY. We also include industry-by-year-fixed effects, as well as additional controls including lagged market capitalization, market capitalization squared and turnover. The regression takes the form $\Delta Y_{i,j,l,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,l,t} + \beta_2 Y_{i,j,l,t-1} + \mathbf{X}\gamma + \sum_l \sum_t \alpha_l \times \alpha_t + \epsilon_{i,j,l,t}$. All variables are as previously defined. Standard errors for the OLS analysis (in brackets) are clustered at the technology industry level.

	$\Delta Turnover_{t+6}$	$\Delta Turnover_{t+12}$	$\Delta \log(Spread)_{t+6}$	$\Delta \log(Spread)_{t+12}$	$\Delta \log(Illiq)_{t+6}$	$\Delta \log(Illiq)_{t+12}$	$\log(InstOwn)_{t+12}$
$\mathbf{1}\{REDACT\}$	1.560***	2.148***	–0.054***	–0.046**	—0.132***	–0.112***	0.105*
	[0.396]	[0.586]	[0.019]	[0.019]	[0.051]	[0.036]	[0.059]
$Level_{t-1}$	–0.896***	–0.861***	–0.394***	–0.306***	–0.600***	–0.532***	–0.354***
	[0.074]	[0.094]	[0.026]	[0.020]	[0.052]	[0.047]	[0.019]
$MktCap_{t-1}$	–3.637**	–3.667***	–0.192***	–0.116**	–1.232***	–1.251***	0.950***
	[1.436]	[1.002]	[0.071]	[0.051]	[0.108]	[0.081]	[0.117]
$MktCap^2_{t-1}$	0.194***	0.202***	–0.000	–0.002	0.025***	0.028***	–0.033***
	[0.066]	[0.050]	[0.003]	[0.002]	[0.007]	[0.005]	[0.005]
Analysis	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Industry-by-Year FE	Y	Y	Y	Y	Y	Y	Y
Contract + Firm Controls	Y	Y	Y	Y	Y	Y	Y
N	3,582	3,610	3,582	3,610	3,582	3,610	33,42
R^2	0.748	0.860	0.426	0.498	0.607	0.625	0.471

Table A.5.5: Robustness With Firm Fixed-Effects

This table displays results of specifications with firm-fixed effects. The regression takes the form $Y_{i,j,t+k} = \beta_1 \mathbf{1}\{REDACT\}_{i,j,t} + \beta_2 \mathbf{1}\{REDACT\}_{i,j,t} \times FilerLicensor_{i,j,t} + \sum_j \alpha_j + \sum_t \alpha_t + \epsilon_{i,j,t}$. All variables are as previously defined.

	$Turnover_{t+6}$	$\log(Spread)_{t+6}$	$\log(Illiq)_{t+6}$	$\log(InstOwn)_{t+12}$	$\log(FNPATS)_{t+1}$	$\log(FNPATS)_{t+3}$	$\log(TSM)_{t+1}$	$\log(TSM)_{t+3}$
$\mathbf{1}\{REDACT\}$	0.003	0.052	0.168*	0.137	–0.036	0.005	–0.103	–0.018
	[0.655]	[0.066]	[0.089]	[0.091]	[0.050]	[0.088]	[0.066]	[0.098]
$\mathbf{1}\{REDACT\} \times FilerLicensor$	0.321	–0.170**	–0.337***	0.065	0.147***	0.120	0.295***	0.220**
	[0.501]	[0.086]	[0.126]	[0.106]	[0.038]	[0.077]	[0.041]	[0.095]
Analysis	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Firm + Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Contract + Firm Controls	Y	Y	Y	Y	Y	Y	Y	Y
N	4,665	4,646	4,665	4,415	4,597	4,044	4,597	4,044
R^2	0.956	0.899	0.862	0.904	0.903	0.905	0.902	0.903