OPEN DISCLOSURE USING INVENTION PLEDGES: A CASE STUDY OF IBM

Abstract

We study the nature and consequences of IBM's invention pledge program, the largest of its kind in US history. Invention pledges involve the open disclosure of patentable inventions without seeking patent protection, unlike patent pledges for which firms waive patent rights after patenting. Competing theoretical views characterize invention pledges either as low-quality inventions with poor organizational fit or as patentable inventions pledged openly by firms with weak firm-specific patent rights and lead-time advantages. We test these theoretical claims using a novel method that compares the quality of subsequent patents citing IBM invention pledges with similar others. Citing IBM invention pledges is positively correlated with patent quality, inconsistent with theories of adverse selection. IBM patents cite its invention pledges on average two years before others, consistent with private benefits of open disclosure. Cited invention pledges spur follow-on innovation, further highlighting their social benefits. We discuss the generalizability of our results obtained using the sample of cited invention pledges to uncited invention pledges. Overall, our results highlight the tradeoffs firms face between knowledge protection and disclosure and generalize to a small number of other disclosing firms.

Keywords: invention pledge, lead time advantage, open disclosure, follow-on innovation
JEL Codes: O36, O32, O31

1. INTRODUCTION

To protect their intellectual property, firms rely on patents, which are used both historically and internationally (see Williams (2017) for a recent review). Yet studies show that patents are not always the preferred mechanism for intellectual property protection and that firms use alternative mechanisms such as secrecy, lead-time advantage, and complementary capabilities to protect and profit from inventions (Cohen *et al*., 2000; Sakakibara and Branstetter, 2001; Lerner, 2009). These empirical findings raise three important questions: Why do firms not seek patent protection for some potentially patentable inventions? How do inventions that receive patent protection differ from those associated with more open disclosure regimes? What are the consequences of these open disclosure strategies for the firm and the society?

To address these questions, we examine the characteristics of openly disclosed inventions, via *invention pledges*, and how open disclosure of inventions affects cumulative innovation. Invention pledges are the open disclosure—through corporate invention disclosure journals—of seemingly patentable inventions without ever seeking patent protection.[1] In the 20th century, several exceptionally large firms facing antitrust enforcement began publishing invention pledges through corporate journals to balance the tradeoffs among patenting, secrecy, and open disclosure. On the one hand, patents grant the right to exclude others from selling in theory, but exercising those rights in practice can be costly for such exceptionally large firms. This is so because obtaining patents involves considerable costs, but it is difficult for large firms to enforce their patents and, if needed, extract licensing revenues from smaller rivals when they are facing antitrust scrutiny. On the other hand, secrecy is also an attractive option in theory, but it is ineffective in practice because the threat of patenting by rivals limits large firms' freedom of action. Secrecy also limits incentives for inventors who value open recognition and a public record of their achievements due to career concerns. Despite the significant historical use of invention pledges, few studies examine empirically the motivation behind their open disclosure and the consequences.

We address this gap in the literature by studying, from its beginning to its end, the largest corporate open disclosure program of its kind in US history: IBM's invention pledge program, from 1958 to 1998. To assess the characteristics of invention pledges, such as quality and relatedness to the firm's core technological area, we would ideally compare IBM's invention pledges to patents that are identical in terms of the underlying invention and differ only in the level of intellectual property protection. But IBM invention pledges provide relatively short descriptions and are not assigned to patent classes, making it difficult to match them directly to patents and to assess the quality of inventions contained in them using comparable measures of quality. Therefore, we propose a novel method to infer the quality of invention pledges by examining their impact on cumulative innovation—that is, examining the quality of subsequent patents that do or do not cite them.

[1] Invention pledges are distinct from patent pledges in important ways. While patent pledges involve pledging non-assertion of patent rights after obtaining patents, invention pledges involve the disclosure of patentable inventions without ever applying for a patent. Invention pledges, therefore, constitute an irreversible and indivisible commitment of voluntary restraint whereas patent pledges can be applied selectively and sometimes firms renege on their pledges (Contreras 2015). Moreover, invention pledges incur lower costs compared to patent pledges due to the addition cost of patenting and do not raise a firm's ability to exercise market power through the acquisition of intellectual property rights. Therefore, invention pledges provide a useful yet underexplored setting to study the implications of open disclosure of patentable inventions.

To do this, we construct a dataset beginning with the list of inventions described in the *IBM Technical Disclosure Bulletin* between 1958 and 1998. Then we extend the dataset of IBM inventions by incorporating IBM patent data from 1976 to 2006. Using these original inventions as the root data, we construct a core dataset consisting of the USPTO patents that directly cite either an IBM invention pledge or an IBM patent (or both), which, together, comprise the "IBM invention pool." Using the dataset of USPTO patents that cite the IBM invention pool, we examine the quality and distance of patents that cite IBM invention pledges to the IBM patent portfolio, relative to a comparison sample of patents that do not cite IBM invention pledges, controlling for patents' reliance on other non-patent literature.

This method has the obvious disadvantage that it cannot be applied to invention pledges that are never cited. In our estimation sample, more than 17,225 IBM invention pledges—about 21 percent of the 83,902 IBM pledges during 1958-1998—are cited at least once, allowing us to study the nature of knowledge openly disclosed in a significant share of the invention pledges.[2] Our results are, therefore, specific to cited invention pledges. We discuss in Section 7 how the sample of *cited* invention pledges differs from the *uncited* and whether our results generalize to all invention pledges by IBM during this period.

Our results indicate that patents citing the above-mentioned cited invention pledges are better in quality than comparison patents, as measured by patents' forward citations. Moreover, the higher the number of citations to an IBM invention pledge, the higher is the number of forward citations received by the patents that cite this invention pledge. These results, in turn, indicate that the quality of knowledge contained in cited invention pledges is significant and persistent in cumulative innovation. Using a measure of the distance between follow-on inventions and IBM's patent portfolio (Jaffe 1986), we also examine whether the knowledge disclosed in cited invention pledges is central or peripheral to the portfolio. Our results indicate that patents citing IBM invention pledges are closer to IBM's patent portfolio than are

[2] Focusing on this sample is reasonable given its scope: invention pledges with at least one citation represent a large sample, particularly when compared to the number of patents IBM pledged in 2005—namely, 500. Prior studies have documented the consequences of patent pledges using this sample of 500 invention pledges (e.g., Wen et al. 2015; Sundaresan et al., 2017; and Ayvazyan and Matr, 2019).

patents that do not cite invention pledges. These results suggest it is less likely that IBM openly disclosed cited invention pledges due to an absence of organizational fit.

Our results are robust to controlling for the differences in industries and commonly employed firm-level determinants of patenting such as R&D intensity and capital intensity in subsamples of Compustat firms (Hall and Ziedonis 2001). Our results can be driven by the potential correlation between follow-on patent quality and heterogeneity in the quality of search for prior art. If inventors of higher-quality patents are more likely to search for prior art causing them to cite IBM invention pledges, then we misattribute differences in the quality of search for prior art to differences in invention quality. Therefore, we construct a comparison sample of patents that do not cite IBM invention pledges but have at least one inventor in common with follow-on patents that do cite IBM invention pledges. This ensures that both follow-on patents' and comparison patents' inventors are aware of IBM invention pledges and mitigates the differences in their search strategies for prior art.[3] It is reasonable to question whether our results are sensitive to the definition of comparison sample. Our results are similar when follow-on and comparison patents are matched on additional dimensions, such as patent assignee, patent year and patent class.

If cited invention pledges are related to IBM's core businesses, how did IBM capture value from invention pledges? We address this question by examining the relative recency of citations to invention pledges in IBM patents, compared to other patents citing IBM invention pledges. We find that IBM's own patents cited more invention pledges and, on average, did so two years before rivals cited them, indicating significant lead-time advantages in follow-on invention. Overall, these results indicate private benefits of open disclosure.

Next, we examine the social benefits of open disclosure. If cited invention pledges involve valuable inventions but accessible at no cost, we expect these pledges to spur follow-on invention by firms other than IBM. Building on the identification strategy of Moser and Voena (2012), we compare within the same

[3] The disadvantage with this latter approach, for our purposes, is that matching on patent class eliminates any technological distance between follow-on and comparison patents, an aspect we want to investigate. Therefore, in the Appendix, we provide results using alternative comparison samples that show the robustness of our results across different comparison samples.

patent class, the quantity and quality of patents filed in subclasses in which patents cite IBM pledges, to the quantity and quality of patents filed in subclasses in which patents do not cite IBM pledges. Our results indicate that invention is higher in subclasses influenced by IBM invention pledge activity, consistent with underlying cited invention pledges being valuable sources of knowledge that spur follow-on invention. Our results generalize to a small number of other disclosing firms, such as Xerox.

Our primary contribution is to the literature on the role of patents as incentives for innovation. The strength of patent rights and its impact on innovation have generated significant debate (see, for a review, Williams (2017)). The strength of patent rights is generally associated with the nature of the industry and the national patent system (Cohen *et al*., 2000; Lerner, 2009). However, even within a national patent system and an industry, firms can experience varying levels of patent protection for disparate reasons, resulting in firm-specific appropriation strategies. Therefore, we add the insight that the patent system plays a heterogeneous role as an incentive for innovation for different firms, even within an industry. The result is that some firms adopt various non-patent strategies and partial patent strategies (Cohen *et al*., 2000; Anton and Yao, 2004; Contreras, 2015; Parchomovsky and Mattioli, 2011).

Second, we contribute to understanding how organizations cope with varying levels of intellectual property protection. Specifically, we examine how exceptionally large firms cope with a weak firm-specific patent regime created by strong antitrust enforcement. We find that invention pledges through corporate journals represent a two-tiered system of invention disclosure practiced by exceptionally large firms facing antitrust enforcement—one at the organizational level and the other through the national patent system. Invention pledges are, therefore, the organizational analogue to the patent system that: (i) preserves inventors' incentives within the organization through a system of allocating credit and rewarding its inventors; (ii) avoids the cost of obtaining unassertible patents while preserving freedom of action by openly disclosing inventions, thus making them a prior art that is easily accessible to patent examiners; and (iii) exploits other mechanisms, such as lead-time advantage, to appropriate returns from invention pledges.

Third, analyzing an historically significant episode of open disclosure by an exceptionally large firm (IBM), we contribute to the ongoing debate about the appropriate role of industrial R&D in the US.

As firms moved away from research to development (and papers to patents) in the 1990s (Arora *et al*., 2019; Arora *et al*., 2018; Rosenbloom and Spencer, 1996), IBM moved away from open disclosure to become the largest patenting organization at the USPTO. As IBM's ability to assert its patent rights grew in the early 1990s, its R&D activities shifted from basic to applied (Buderi, 2000), indicating how stronger patent rights differentially influence incentives for basic and applied innovation.

Finally, there is renewed interest in understanding disclosure effects of the patent system, particularly on follow-on innovation (Hegde *et al*., 2018; Furman *et al*., 2018). In our empirical setting, the exclusionary effects of open disclosure through invention pledges are relatively lower compared to disclosure through patents, allowing us to study the enabling effects of disclosure. The use of invention pledges as a substitute for patenting offers a useful insight: those with lead-time advantages can adopt open disclosure over secrecy under antitrust enforcement, and such disclosure can inspire follow-on innovation (see, also, Watzinger *et al*. (2017)).

The paper is organized as follows. We review the literature in Section 2. In Section 3, we introduce the research setting. We present the data and methodology in Section 4 and the results in Section 5. In Section 6, we discuss several robustness tests and the generalizability of our results—which are based on *cited* invention pledges—to *uncited* invention pledges, discussed in Section 7. We conclude in Section 8.

2. LITERATURE REVIEW AND RESEARCH QUESTIONS

New knowledge is a key source of competitive advantage for firms (Grant, 1996). Investing to create new knowledge, managing knowledge diffusion and absorption, and profiting from this knowledge involve considerable challenges (Nelson, 1992; Rosenberg, 1990; Cohen and Levinthal, 1990; Teece, 1986). Knowledge spillovers can undermine the competitive advantage of the creator of knowledge by stimulating follow-on innovations and strategic entrepreneurship (Agarwal *et al*., 2010). Understanding how firms create incentives for the generation of new knowledge and how they protect and benefit from it remain important questions in the literature on incentives for innovation.

Open disclosure of patentable knowledge generates spillovers that can be beneficial to rivals and undermine appropriability, leading to questions about why firms pledge inventions. Several theoretical studies model open disclosures without patent protection as an optimal disclosure strategy to either deter or delay rival innovation. For example, suppose that two firms are engaged in a patent race that spans an intermediate stage producing an intermediate result and a final stage in which a licensable innovation (a patent) may be developed. In the intermediate stage, one of the two firms, the leader, discovers an intermediate result, whereas the follower has yet to discover it. Gill (2008) notes that the leader's disclosure of the intermediate result has two effects: (i) a spillover effect that helps the follower; and (ii) a signaling effect that can induce the follower to exit the race. Gill (2008) shows that the costs of development in the intermediate and final stages determine the leader's incentive to disclose intermediate research results. Accordingly, on the one hand, if the development cost of the final stage is low, the follower is less likely to be deterred by the leader's disclosure; thus, the leader has little incentive to disclose openly. On the other hand, if the development cost is high, then open disclosure can serve as a signal to deter followers.

Others model how open disclosure can slow down, if not deter, rivals in a winner-take-all race to reach a significant technological milestone (Bar, 2006; Baker and Mazzetti, 2005; Parchomovsky, 2000; Lichtman *et al.*, 2000). In these types of models, if the firm lagging behind in a patent race can disclose its intermediate results openly and establish them as "prior art"—and, thus, advance the prior art marginally—the laggard makes it difficult for the leading firm to reach a level of innovation that is patentable (relative to the new prior art containing the intermediate result) and, hence, extends the patent race.[4]

In the above theoretical literature, the intermediate results are not patentable, as they do not reach the level of invention necessary to receive a patent; however, the disclosure has intentional or inadvertent spillover benefits (e.g., De Fraja, 1993). Pacheco-de Almeida and Zemsky (2012) arrive at similar conclusions, although they assume no spillover benefits that compensate innovators for their open disclosure. In their setting, two competing firms in a patent race experience time-compression diseconomies. The two firms can develop the new technology concurrently or sequentially. In the latter case, one firm

[4] The broader literature on patent races is extensive (for a review, see Thompson and Kuhn (2020)). In early random-discovery models, firms compete until a patent is obtained (summarized also in Reinganum (1989)), while later models allowed for leapfrogging (Fudenberg *et al.*, 1983). We focus on a subset of this literature that uses disclosure as a strategy in the patent race (e.g., Baker and Mazzetti, 2005).

waits until the other has introduced the new technology in the market. Pacheco-de Almeida and Zemsky (2012) find that the follower prefers to wait rather than develop concurrently if the leader discloses openly and spillovers are sufficiently high. The leader finds open disclosure optimal because it softens competition.

The invention pledges that we study are made by exceptionally large, pioneering firms (albeit under the shadow of regulatory action) rather than by laggards in a patent race, thus undermining the empirical validity of some of the theoretical models. To reconcile this observation, Johnson (2014) examines motivations for the open disclosure of patentable inventions by a leading firm (as in our empirical setting). He finds that inventions that are not too technically challenging and are easy to invent around are openly disclosed. A general conclusion of the above studies is that less-valuable inventions—either intermediate or easy-to-imitate results—are openly disclosed for strategic long-term advantage (e.g., Eaton and Eswaran, 2001). This has contributed to the idea that invention pledges such as those contained in the *IBM Technical Disclosure Bulletin* reveal knowledge that is not comparable to the knowledge disclosed through patents. Others argue that openly pledged inventions may also lack complementarity and fit (*á la* Cassiman and Ueda, 2006; Hellmann and Perotti, 2011), which firms find too distant from their core business and, hence, not optimal to commit their limited resources to secure patent rights.

An alternative set of studies explains open disclosure of patentable inventions as an organizational response to prevailing institutional pressures that weaken firm-specific patent rights and promote open disclosure norms. In explaining such disclosure practices, scholars highlight the important role that antitrust policy played in the US innovation system, particularly during the post-war period (Arora *et al*., 2019; Hounshell, 1996, pp. 41-48; Mowery, 1992; Hart, 2001). Open disclosure is a relatively more attractive option for industry giants facing antitrust enforcement, as it avoids the cost of securing patents that cannot be asserted against smaller rivals, while also preserving the freedom of action for industry giants and their inventors' incentives through open recognition and rewards. Open disclosure through corporate invention disclosure journals is, therefore, an organizational analogue developed by exceptionally large firms as a complement to the patent system. Open disclosure norms became prevalent among firms in the US in the 1900s, as they began establishing industrial R&D labs and sought to attract and retain scientific personnel

(Hounshell, 1996). In these R&D labs, firms did not always tightly control newly generated knowledge, especially in industries in which patents were not effective (Nelson, 1992). Even in industries in which patents were effective, firms still divulged information without losing their proprietary edge because such open disclosure by for-profit firms could be complementary to public investment and the broader development of technology. Prior theoretical studies suggest that when innovations are sequential and complementary (Bessen and Maskin, 2009) and when the time horizon for commercialization is too long, firms can adopt a pro-publication environment and promote open disclosure (Aghion *et al*., 2008). These competing perspectives give rise to our first research question.

Research Question 1: *How valuable is the knowledge contained in IBM invention pledges? How proximate or distant are IBM invention pledges to IBM's own patent portfolio?*

The prior literature documents the implications of open disclosure for patenting and value capture. Scholars of innovation have long wrestled with an apparent contradiction between firms' profit motive and the Mertonian norms of open science (Merton, 1973). Rosenberg (1990) challenges the notion that published research is easily appropriated by free-riding firms. Implementing an innovation openly disclosed in a scientific publication or as an invention pledge requires experts that understand them. Scholars have argued that participating in the open disclosure of scientific results is essential to absorbing external knowledge (Cohen and Levinthal, 1989). Unless firms maintain a reputation for open disclosure, they may find themselves excluded from formal conferences, as well as from informal interactions at the frontier of science. Gambardella (1992) reviews case studies of pharmaceutical companies and shows a correlation between the extent of publishing and patenting. Cockburn and Henderson (1998) likewise combine qualitative and quantitative data to show that in the pharmaceutical field, firms achieve "connectedness" to the public sector not only by employing scientists, but also by explicitly considering their publications for purposes of internal promotion. Publishing openly may also serve to burnish a firm's reputation with external stakeholders (Dasgupta and David, 1994; Hicks, 1995; Audretsch and Stephan, 1996; Stern, 2004; Hicks, 1995; Sauermann and Cohen, 2010). Yang *et al*. (2010) argue that innovators who disclose knowledge benefit from learning vicariously from recipients who create complementary knowledge. In a seminal study, Cohen *et al*., (2000) identify lead-time advantages and complementary marketing and manufacturing capabilities as important mechanisms for capturing value from inventions, aside from patenting and secrecy. If a firm adopts open disclosure and freedom of action as an intellectual property strategy under antitrust enforcement, then exploiting these mechanisms is vital to profiting from invention pledges. This leads to our next research question.

Research Question 2: *How did IBM benefit from open disclosure of invention pledges?*

Disclosure is an important part of the grand bargain struck in the design of the patent system, as firms are required to disclose new knowledge to society in exchange for receiving monopoly rights through a patent.[5] Studying the motivation and consequences of disclosure using patent data, however, is fraught with challenges. Since disclosure is a precondition to granting a patent, firms' motivation to voluntarily disclose patentable innovations cannot be easily examined using patent data. In addition, it is difficult to isolate the exclusionary and enabling effects of disclosure through patents on follow-on innovation (Hegde and Luo, 2018; Hegde *et al*., 2018; Furman *et al*., 2018).

Given the growing popularity of patent donations, scholars have used patent pledges as an avenue to study the social benefits of open disclosure. Pledged patents are those for which firms waive patent rights. Pledging not to assert patent rights after making the costly investment to secure patents appears counterintuitive, but several firms, such as Tesla, Toyota, IBM, and Monsanto, have pledged some of their patents (e.g., Contreras, 2015), indicating strategic benefits from such openness (Alexy *et al*., 2018). The prior literature on patent pledges focuses largely on the consequences of such pledges. For example, Wen *et al*. (2015) find that IBM's 2005 patent pledges stimulated new open-source software product introductions. Sundaresan *et al*. (2017) find that IBM's patent pledges lowered citations by external inventors, and Ayvazyan and Matr (2019) show that pledges led to more breakthroughs and trade in patents.

---

[5] A growing literature examines how disclosure is related to competitive advantage (Pacheco-de Almeida and Zemsky, 2012; Alexy *et al*., 2013); why firms choose to disclose inventions during the patent application process (Graham and Hegde, 2015); how disclosure through patenting affects the sale of patents in the market for ideas (Hegde and Luo, 2018); follow-on innovation (Sampat and Williams, 2019; Hegde *et al*., 2018; Furman *et al*., 2018; Gross, 2019; Barrufaldi and Simeth, 2018; Thompson and Kuhn, 2020); exploration of new research directions (Murray *et al*., 2016; Bessen and Maskin, 2009; Aghion *et al*., 2008); and knowledge reabsorption (Belenzon, 2012; Yang *et al*., 2010).

Alexy and Reitzig (2013) view IBM patent pledges as a coordination strategy to jointly produce a common open resource (such as open-source software) with other firms in a private collective and to derive benefit from producing proprietary complements.

Despite their merits, patent pledges have some limitations. First, patent pledges are voluntary commitments of restraint on which firms can and do renege, suspending the pledge defensively when it is convenient for them to do so. Second, the legal enforceability of patent pledges remains an open question (e.g., Contreras, 2015). Third, patent pledges are sometimes selective and incomplete, as firms declare their intention not to seek royalties from small firms. Finally, patent pledges are often confined to technologies for which the benefits of strategic openness are generally obvious and relate to setting industry standards or limiting holdup due to patent litigation (Ziedonis, 2004).

Studying the effect of disclosure on follow-on innovation using invention pledges is useful because the exclusionary effects of disclosure created by the patent system are mitigated. Since follow-on inventors of IBM invention pledges do not have to compensate IBM through licensing fee, the enabling effects of the disclosure can be observed. Prior studies show, however, that disclosure can have entry-deterring effects, as rivals are discouraged by potentially aggressive competitors (e.g., Gill, 2008). Nevertheless, compared to patents, we expect invention pledges to have lower exclusionary effects. Therefore, if the knowledge disclosed through invention pledges is valuable, it is more likely to spur follow-on innovation. These insights lead to our final research question.

Research Question 3: *Do we see instances or (evidence of) follow-on innovation associated with IBM invention pledges?*

3. RESEARCH SETTING

The historical beginnings, rise, and the eventual fall of the open disclosure program at IBM can be understood in the context of antitrust enforcement and the strength of the prevailing patenting regime. Sustained antitrust scrutiny from the Department of Justice and IBM's competitors, collectively known as the "seven dwarves" (Hart, 2001, 2007), coupled with the Department of Justice's restrictions on patent licensing until the 1990s (Gilbert *et al*., 1997), significantly limited IBM's ability to exercise monopoly rights through patents (Grindley and Teece, 1997; Hounshell, 1996; Usselman, 1993; 2009). In response, IBM established its *IBM Technical Disclosure Bulletin* and pledged inventions in response to antitrust enforcement that ended in a consent decree in 1956, rendering IBM patents not assertible in practice due to compulsory licensing restrictions (see, for an analysis of the AT&T case, Watzinger *et al*. (2017) and for the DuPont case, Mullin and Snyder (Forthcoming)).

The program of invention pledges through the *IBM Technical Disclosure Bulletin* (TDB) lasted for nearly four decades from 1958 to 1998, when patent rights became significantly strong in the US leading to patent explosion (Hall, 2004) and transforming IBM into the top patenting organization at the USPTO.[6] Figure 1 plots the number of invention pledges published in IBM TDB's monthly volumes. The number of invention pledges increased to, on average, 200 per month in the 1970s but crossed 400 in the later decades. Xerox's disclosure program had similar beginnings in 1976, after signing the consent decree in 1975, limiting its patent rights; the Xerox program ended in 2001. Following the example set by AT&T's Bell Labs (Hounshell, 1996, pp. 41-48), IBM promoted an open research environment after it signed the consent decree, launching the *IBM Journal of Research and Development* in 1957 and IBM Systems Journal in 1962, increasing the number of PhDs employed from 105 in 1956 to 898 in 1960, rewarding its most exceptional R&D employees with the title of "IBM Fellow" and unfettered freedom and resources to pursue their ideas (Bhaskarabhatla and Pennings, 2014).

-- Place Figure 1 here --

While IBM maintained multiple platforms for open disclosure during this period, the invention pledges are distinct from scientific publications, for which IBM maintained a separate journals, namely IBM Journal of Research and Development and IBM Systems Journal. The invention pledges in IBM TDB were more similar to patents than publications; they contained patent drawings illustrating the invention

[6] Lopatka (2000) notes: "Over the years, some of the provisions of the decree became obsolete, and in 1994 IBM sought termination of its remaining portions, a request that the government ultimately supported. In 1995 the Second Circuit removed the judge overseeing the case, Judge Edelstein, for apparent lack of impartiality. See Second Circuit Orders Judge's Recusal from Termination of IBM Consent Decree, 68 Antitrust & Trade Reg. Rep. (BNA) 87 (Jan. 26, 1995). The new judge terminated the decree, and in 1998 the Second Circuit affirmed. United States v. IBM Corp., 163 F.3d 737 (2d Cir.1998)."

and its embodiments. Bhaskarabhatla and Pennings (2014) note the striking similarity between invention pledges and patents and note that some disclosing firms such as Xerox even categorized their invention pledges based on patent classes. According to Bhaskarabhatla and Hegde (2014), until 1989 IBM inventors submitted their invention disclosures to internal review committees comprising research managers and attorneys from the company's legal division. The review committees then elected to either pursue a patent or publish the invention in IBM's *Technical Disclosure Bulletin*. Jim McGroddy, the newly appointed R&D Director at IBM, ushered in a pro-patent regime within IBM and shifted incentives away from invention pledges to patents in 1989 (Bhaskarabhatla and Hegde, 2014), eventually shutting the invention pledge program down in 1998.

Since the inception of the invention pledge program, IBM implemented a points-based promotion system for its scientists that awarded three points for each patent and one point for each invention pledge, indicating that IBM created incentives for its employees to contribute to invention pledges. Since invention pledges became an integral part of evaluating R&D personnel at IBM during this period, inventors had less incentive to keep valuable inventions secret. The directors of R&D at IBM during 1970-1996, Ralph Gomory, John Armstrong, and Jim McGroddy, as well as business historians specializing in IBM, document an open environment at IBM's R&D division for decades and the growing need for less open disclosure as IBM's financial performance suffered in the early 1990s (McGroddy, 1998; 2001, Armstrong, 1996; Gomory, 1989; Buderi, 2000; Pugh, 1995).

Since the turnaround, IBM has received more patents than any other organization in the last 25 years, securing 9,043 in 2017 alone (IBM, 2018). Yet, at the peak of the invention pledge program in 1990, four in five inventions were disclosed in the *IBM Technical Disclosure Bulletin*, while only 20 percent of the inventions were patented. Reflecting the value of the invention pledges, the *IBM Technical Disclosure Bulletin* was cited more than 80,000 times by subsequent US patents. In an early analysis of nonpatent literature citing patents, Narin and Noma (1985) note that *IBM Technical Disclosure Bulletin* "receives as many references from patents as all SCl-covered physics journals combined." Similarly, the *Xerox*

Disclosure Journal received more than 2,000 citations from US patents. *Research Disclosure*, a journal used for open disclosure by several firms, received more than 10,000 citations from US patents.

IBM invention pledges are an historically significant and well-suited setting to examine the impact of open disclosure on cumulative innovation. The invention pledge program institutionalized open disclosure for inventors working at IBM with significant implications for IBM and for the industry. For IBM, an increasing volume of invention pledges combined with a constant number of patents until the early 1990s resulted in the diffusion of knowledge and little licensing revenues from outside IBM (Bhaskarabhatla and Hegde 2014). For the industry, IBM's open disclosure led to entry and competition in product markets which it pioneered (e.g., disk drives, personal computers, memory, databases, among others). We deepen the analysis of the IBM invention pledge program by examining the nature of cited invention pledges and their impact of cumulative innovation.

4. DATA AND METHODOLOGY

4.1 Data Construction

Our IBM invention pledge data extend from 1958, when the *Technical Disclosure Bulletin* was established, to the program's closure in 1998. We complement these data with NBER patent dataset. Since NBER data begin from 1976, we are able to examine patents that cite invention pledges from this period onwards to 2006, when the dataset ends. As noted earlier, Figure 1 depicts the evolution of the invention pledge program at IBM. Figure 2 shows the diffusion of knowledge originated in IBM inventions, as captured by citations. The root of the diffusion tree consists of all IBM inventions, including IBM patents and IBM pledges. The first generation of the citation tree includes patents that cite either IBM patents or IBM pledges or both. Since IBM pledges do not list references, we do not know whether the pledges are built upon other IBM pledges or patents. For this reason, IBM pledges exist at the root of the tree. Patents in the first generation include both IBM patents and patents by others. Some pledges are cited by patents, allowing us to examine their quality indirectly by comparing the characteristics of patents that either cite or do not cite the pledges. In the empirical strategy to be described in Section 4.2, we denote similar patents that do not cite IBM inventions as *comparison patents*.

-- Place Figure 2 here --

We follow several steps to construct the tree-structured data (for details on data construction, please see Appendix A). Our final sample of the *root invention pool* contains 25,274 IBM patents and 17,225 cited IBM invention pledges. We also compile financial information for a subset of firms listed in the Compustat dataset (for a list of variables, see Table 1, and for descriptive statistics, see Table 2).

-- Place Tables 1 and 2 here --

Assessing the quality of disclosures is vital to understanding the rationale for open disclosure of knowledge. Accordingly, we collect data on forward citations to follow-on patents that cite IBM patents and/or invention pledges. We also construct a comparison sample of patents, which is composed of patents with at least one inventor in common with the follow-on patent, but not citing IBM invention pledges or patents. In doing so, we exploit data on non-patent literature citations of all US patents.

To examine lead-time advantages and follow-on innovation, we collect data on the number of times patents citing IBM invention pledges are themselves cited, and the number of years between the patent application year and the publication year of the cited invention pledge. To accurately compute this latter measure, we extract the publication year for each cited IBM invention pledge (Table B1 in Appendix contains a list of variations of *IBM Technical Disclosure Bulletin* we found in the nonpatent references). When the publication year is missing, we exploit volume numbers of the invention pledges in nonpatent literature references, with volume 1 corresponding to 1958, volume 2 to 1959, and so on. Using this method, we identify publication years for more than 99 percent of the cited invention pledges. Some patents cite multiple IBM invention pledges from different publication years. To address this, we construct three alternative measures of the gap between citing patent's application year and the publication years of the IBM invention pledges, one based on the oldest, the other based on the most recent, and the third based on the average publication year of the invention pledge.

4.2 Methodology

Our methodological approach involves comparing the characteristics of patents that cite an IBM invention pledge and/or an IBM patent with the characteristics of patents that cite neither IBM invention pledges nor

IBM patents. To this end, we compare a patent citing IBM inventions (pledges, patents, or both) with similar patents that do not cite IBM inventions. Patents in the comparison sample contain at least one inventor in common with the patent citing IBM inventions, but they neither cite IBM invention pledges nor IBM patents. As noted earlier, in robustness checks we use an alternative comparison sample by matching on patent assignee, patent year and patent class.

We estimate the following equation:

$$y_i = \alpha_0 + \alpha_1 P_i + \alpha_2 D_i + \alpha_3 PD_i + X_i\alpha + \delta_{ci} + \mu_{ti} + \epsilon_i, \qquad (1)$$

where i denotes patent, and y_i refers to patent i's measure for quality or distance. We measure quality using forward citations and measure distance to IBM's patent portfolio using the measure described in Jaffe (1986). Since measures of quality are heavily right-skewed and have fat left tails, the logarithm is applied to alleviate the skewness. P_i is a binary variable indicating whether patent i cites IBM patents but not IBM invention pledges. D_i is a binary variable indicating whether patent i cites IBM invention pledges but not IBM patents. PD_i is a binary variable indicating whether patent i cites both IBM pledges and IBM patents. X_i is a vector of control variables, including the following variables: 1) the logarithm of the total number of backward citations of patent i, which measures the technical nature of the invention contained in the patent; 2) the logarithm of the number of patent classes in patent i's backward citations, which measures the breadth of the patent's knowledge base; 3) the logarithm of patent i's total number of non-patent citations, which measures the patent's reliance on non-patent literature; and 4) the logarithm of patent i's number of inventors. When estimating Equation (1) with Compustat firms, we also add Compustat controls, including R&D; log(firm age); log(number of employees); three-year moving averages of log sales; change in operating income; stock of firm patents; R&D intensity; and capital intensity. δ_{ci} represents assignee-patent class fixed effects and μ_{ti} application year fixed effect. ϵ_i is the idiosyncratic error term. Compustat controls allow us to better control for time-variant unobservables and mitigate concerns about endogeneity. In the regression on forward citations, a positive value of α_2 would imply that patents citing IBM invention pledges are more valuable than patents citing neither IBM patents nor pledges, all else equal. In the regression on distance, a positive value of α_2 would imply that patents citing IBM pledges are less central to IBM's patent portfolio, compared with patents citing neither IBM patents nor pledges, all else equal. We estimate an Ordinary Least Squares (OLS) regression for the full sample and a subsample excluding IBM patents, as IBM's own first-generation patents can drive our results.

The next set of regressions concerns lead-time advantages in follow-on invention. The regression equation is specified as follows:

$$y_i = \alpha_0 + \alpha_1 IBM_i + X_i\alpha + \delta_{ci} + \mu_{ti} + \epsilon_i \, , \qquad (2)$$

where *i* denotes follow-on patent and *t* denotes year. y_i is the number of forward citations, or the number or recency of IBM invention pledges cited. IBM_i is a binary variable indicating whether patent *i* is an IBM patent. X_i is a vector of control variables. When estimating Equation (2) with Compustat firms, we also add Compustat controls as before. We also include additional firm-specific controls, such as whether the patent assignee is a US firm.

To examine whether IBM invention pledge activity is correlated with the quantity and quality of follow-on innovations, we analyze the patenting activity at the level of patent classes and subclasses. We follow an estimation strategy similar to that of Moser and Voena (2012) and estimate the following equation:

$$y_{s,t} = \alpha_0 + \alpha_1 D_{s,t-1} + \alpha_2 P_{s,t-1} + \alpha_3 K_{s,t-1} + \delta_{ct} + \epsilon_{s,t} \, , \qquad (3)$$

where *s* denotes a patent subclass within a three-digit patent class *c*. $y_{s,t}$ refers to the number of patents, as well as to the number of citations received by patents, in subclass *s* of a patent class in year *t*. The main explanatory variable $D_{s,t-1}$ measures the number of patents in subclass *s* that cite an IBM invention pledge in year *t-1*. We also control for the number of patents *P* in subclass *s* that cite IBM patents in year *t-1*, $P_{s,t-1}$. We control for patent stock *K* at the subclass level in year *t-1*, $K_{s,t-1}$. Alternatively, we control for average patent stock over the last 3 years. We also control for patent class-year fixed effects, δ_{ct} . We expect that the subclasses of a three-digit patent class that receive patents citing IBM invention pledges in year *t-1* will experience more and better innovations in year *t*.

4.3 Descriptive Statistics

To understand and interpret the regression results that follow, we provide a series of figures that describe the data. Figure 3 shows by patent application year, the average forward citations received by IBM and non-IBM patents, as well as the average backward citations made by IBM and non-IBM patents. The citations show a similar pattern, particularly so for backward citations. Figure 4 shows by patent application year, the average forward citations received by patents that cite IBM Invention Pledges, IBM Patents, and non-IBM patents, as well as the average backward citations made by each of these groups of patents. Patents citing an invention pledge have more forward citations compared to patents that do not cite any IBM inventions. However, patents citing IBM patents or both patents and pledges have, on average, even more forward citations. Among cited inventions, only a small proportion (7.5%) of IBM patents are cited exclusively by IBM itself, whereas 23.8% of IBM invention pledges are cited exclusively by IBM itself, indicating that cited invention pledges may be of more particular use to IBM than they are to others. The patterns are qualitatively similar for the subsample of firms in Compustat. Nearly 63% of the cited invention pledges are cited exclusively by external firms, indicating their potential role in spurring follow-on innovations. We report the descriptive statistics for our full sample and the Compustat sample in Table 2.

-- Place Figures 3 and 4 here --

5. RESULTS

5.1 Quality and Distance

Table 3 shows estimation results for Equation (1) with the entire sample of first-generation patents and comparison patents. Panel A of Table 3 reports results for the overall sample and Panel B for the Compustat subsample. In column 1, the coefficient estimate of *Citing IBM Patent* is positive and significant ($b=0.065$, $p<0.01$), indicating that building on IBM knowledge disclosed in patents led to 6.5 percent more citations compared to a comparison patent that does not cite IBM innovation. The coefficient estimate of *Citing IBM Pledge* is also positive and significant, indicating that building on openly disclosed IBM knowledge was also valuable and associated with 5.8 percent additional citations compared to a comparison patent that does not cite an IBM innovation. The results in columns 3 and 4 indicate that patents citing IBM pledges are closer to the IBM patent portfolio although patents citing IBM patents or both patents and invention pledges are further closer to the IBM patent portfolio. We obtain qualitatively similar results in Panel B when focusing on Compustat firms and controlling for a comprehensive set of Compustat controls. These results are consistent with the view that *cited* IBM invention pledges were central to IBM's patent portfolio and are positively associated with the quality of cumulative innovation.

-- Place Table 3 here --

Next, we examine the moderating effect of invention pledge quality on first-generation citing patents' quality and distance to IBM's patent portfolio. To this end, we interact Invention Pledge Quality, which measures the number of times an invention pledge is cited by patents. We expect that a first-generation patent citing an invention pledge that is highly cited is likely to receive more citations. The key independent variable of interest is the interaction term *Citing IBM Pledge (& not patent)* × *Pledge Quality*, whose coefficient estimate in column 1 of Panel A in Table 4 is positive and significant (b=0.068, p<0.1). The coefficient estimate in columns 1 and 2 of Panel B is positive and statistically significant. The results on distance reported in columns 3 and 4 of Table 4 are positive in Panel A, and insignificant in Panel B, indicating that higher quality invention pledges are no closer than others to IBM's portfolio. These results are broadly consistent with cited invention pledges containing valuable knowledge that is close to IBM's patent portfolio.

-- Place Table 4 here --

5.2 Lead-time Advantage

To examine lead-time advantage in follow-on invention, we analyze a dataset of all first-generation patents citing IBM invention pledges. The dependent variables in these regressions are the log number of forward citations, the log number of IBM invention pledges cited, and the number of years between the patent application year and the year in which the cited invention pledge is published. Since some first-generation patents cite more than one invention pledge, we use alternative dependent variables that measure the gap in years between the application year and the oldest invention pledge (instead of the most recently published invention pledge), as well as the average gap across all cited invention pledges.

The results of the estimation are shown in Table 5. In Panel A, the coefficient estimates of IBM in columns 1 and 2 are positive and significant (b=0.198, p<0.01; b=0.176, p<0.01), indicating that IBM's own first-generation patents citing IBM invention pledges were higher in both quality and quantity. The coefficient estimate of IBM in column 3 is negative and significant (b=-2.010, p<0.01), indicating that IBM had nearly two years of lead time relative to others on recently disclosed invention pledges. The results are qualitatively similar across different dependent variables in columns 4 and 5. Overall, IBM was, on average, 19 months ahead of others in patenting inventions based on its invention pledges, as shown by the results in column 5. The results are similar for Compustat firms in Panel B.

-- Place Table 5 here --

Bhaskarabhatla and Hegde (2014) exploit the shift from invention pledges to patents within IBM. To account for these changes, we add an additional variable, *IBM × After 1989*, to examine the relative change in lead time advantage after 1989. We expect IBM to have been more proactive in exploiting its lead time advantage after 1989 than before. Consistent with this intuition, we find that first-generation IBM patents cited more-recent invention pledges after 1989 relative to before (columns 3, 4, and 5). We do not find any difference in the quality of first-generation patents produced by IBM after 1989 relative to before (see column 1). The results remain broadly similar in a subsample of Compustat firms, as shown in Panel B of Table 5, and robust to using models suited for analyzing count dependent variables.

5.3 Follow-on Innovation

To examine the impact on follow-on innovation, we begin with a sample containing 12 million non-patent literature citations by US patents and identify citations to IBM invention pledges. We count the number of patents citing IBM invention pledges by year during 1976-2011 but limiting the data until 2006 does not change our results. The data contain 428 three-digit patent classes and 130,396 patent class-subclass combinations. Among them, 316 (74 percent) three-digit patent classes and 13,019 patent class-subclass combinations (10 percent) received patents that cited IBM invention pledges during 1976-2011. We consider them treated subclasses.

-- Place Table 6 here --

We estimate Equation (3) and report the results in Table 6. The coefficient estimate of the variable measuring the number of patents citing an IBM invention pledge in year *t-1* in column 1 is positive and significant, indicating that more patents are filed in year *t* when more patents building on IBM pledges are filed in subclass *s* in year *t-1*. The magnitude of the coefficient estimate is small, indicating that IBM pledge activity has a modest effect in promoting entry into innovation. By contrast, the coefficient estimate of the same variable in column 2 is large in magnitude (two percent), positive, and significant, indicating that the patents filed in these subclasses receive more forward citations compared to patents in other subclasses of the same three-digit patent class. We control for the stock of patents over the last three years can control for any temporal trend at the subclass level. The coefficient estimates for the number of patents citing IBM patents in columns 1 and 2 are small in magnitude, indicating that patenting by IBM has a modest effect on the quantity and quality of subsequent innovation within the patent class. Our results remain robust to using alternative measures for citing IBM pledges and patents in columns 3 and 4. Our results are consistent with prior research that shows that open disclosure of knowledge spurs follow-on innovation, in that it suggests that cited IBM pledges are as effective as patents in supporting innovation. The difference between the effect of citing IBM pledges and the effect of citing IBM patents is larger in column 1 than in column 2, suggesting that IBM pledge activities can lead to better rather than more innovations.

Extending the analysis on follow-on innovation to other firms would involve considerable effort to collect data on individual invention pledges and match the names of inventors for each of the invention pledges to their patenting histories. While we do not have data at that level of granularity for Xerox, we can extend subclass-level analyses concerning the quantity and quality of follow-on innovations to Xerox. We follow a procedure very similar to that for IBM and identify at the subclass level, the number of patents citing *Xerox Disclosure Journal* and Xerox patents in the previous year(s). We estimate Equation

(4) and report the results in the second panel of Table 6. We find that Xerox's disclosure activity is not positively associated with the rate of innovation but with the quality of innovations (8 and 14 percent in columns 2 and 4, respectively). By contrast, the coefficient estimates associated with Xerox patenting activity are close to zero, indicating their limited role in stimulating follow-on innovations within a subclass.

Some firms have historically outsourced the open disclosure of their patentable inventions instead of maintaining an internal journal. A major outlet for such disclosures is a publication titled *Research Disclosure* (Baker and Mazzetti 2005). Although the identity of the firm disclosing the invention can also be disclosed, doing so is not necessary for establishing prior art. Consequently, many publications in *Research Disclosure* are anonymously disclosed inventions. We compile nearly 10,000 citations to such invention pledges, and it forms the second largest set of invention pledges after IBM's *Technical Disclosure Bulletin*. We estimate Equation (4) using citations to *Research Disclosure* in the non-patent literature of the US patents and report the results in the bottom panel of Table 6. Our results are consistent with the view that open disclosure of innovations spurs the rate and quality of follow-on innovations. Since we do not know the identity of the firm that discloses using Research Disclosure, we cannot control for patenting activity in the bottom panel of Table 6.

6. ROBUSTNESS CHECKS AND ADDITIONAL RESULTS

Our results on quality and distance above (see Table 3) are based on a comparison sample that contains at least one common inventor with the sample of follow-on patents. We do so for two reasons: (i) to ensure that both the follow-on patents and comparison patents' inventor teams are aware of the *IBM Technical Disclosure Bulletin*, mitigating concerns about their prior art search and citation strategies; and (ii) so that we do not mechanically match follow-on patents and comparison patents on patent class, resulting in zero distance between them. Nevertheless, one may wonder whether our results on quality are driven by the definition of comparison sample. To address this, we construct an alternative comparison group of patents by matching on patent assignee, patent subclass, and patent application year for each follow-on patent. We present results for both quality and distance in Table A1. Consistent with previous results, these results indicate that invention pledges have a significant and persistent effect on the quality of cumulative innovation. Since we match patent classes, we expect there to be no significant differences in the distance measure between follow-on and comparison patents. Not surprisingly, the coefficient estimates are all small in magnitude (approximately 0.002 to 0.007), indicating that no significant differences in distance could be inferred from this analysis. These analyses also validate our choice of using a comparison sample defined by the existence of common inventors for our primary analyses. We report results of additional analyses examining the role of pledge quality in Table A2. These results, based on the alternative comparison group, are similar to those obtained in Table 4, indicating that the choice of comparison sample does not drive our results.

Besides forward citations, patent quality can also be measured by originality and generality. We estimate Equation (1) with originality and generality of first-generation patents by Compustat firms as dependent variables. The measures of originality and generality are based on the diversity of backward and forward citations respectively (Trajtenberg *et al*., 1997). The results, shown in Table A3, indicate that, compared with patents not citing IBM innovations, patents citing IBM invention pledges are more original and no less general in both estimation samples that use alternative definitions of the comparison group.

In additional analyses, we examine the role of citing IBM invention pledges on the scope of patents. If IBM invention pledges disclose valuable inventions, they are expected to limit the scope of follow-on innovations. The results, shown in Table A4, indicate that patents citing IBM invention pledges have fewer claims than patents not citing IBM invention pledges, implying that these patents are narrower in scope. This, in turn, implies that the cited invention pledge is valuable, as it limits the citing patent's ability to make marginal contributions. The results are broadly similar in both estimation samples that use alternative definitions of the comparison group.

We also examine how IBM's lead-time advantage in follow-on innovation are associated with scope, originality, and generality. The results, shown in Table A5, indicate that first generation IBM patents citing IBM invention pledges are narrower in scope and have lower levels of originality and generality in the overall sample, particularly before 1989. This suggests that compared with non-IBM patents citing IBM invention pledges, IBM patents citing IBM invention pledges rely on a less diverse knowledge source and generate more concentrated citations. In the Compustat subsample, IBM patents citing IBM invention pledges are less original but do not differ significantly in terms of generality with non-IBM patents citing IBM invention pledges.

IBM's innovation management strategy underwent a significant change in 1989, as its new director of R&D, James McGroddy, prioritized patents over invention pledges (Bhaskarabhatla and Hegde 2014). To see if this change affected the quality of patents and invention pledges, we divide the sample into pre-1989 and post-1989. The average five-year citation count for IBM invention pledges published after 1989 was higher than that for pledges published before 1989 (1.2 vs. 0.5). Hence, we find no evidence that IBM invention pledges disclosed after 1989 were less valuable than before. This implies that despite shifting focus towards patenting after 1989, IBM continued to disclose valuable knowledge openly until the avenue to do so was foreclosed in 1998. One alternative motivation for disclosure activity was IBM's collaboration with academic institutions and universities. However, unlike scientific publications in academic journals, IBM's *Technical Disclosure Bulletin* did not involve collaborative work between IBM employees and university-based researchers, foreclosing this explanation.

Finally, we examine to what extent follow-on patents citing IBM invention pledges build on the knowledge disclosed in them. We do so by studying the first-hand descriptions of patent authors of follow-on patents citing IBM invention pledges. We searched the text of more than 48,000 follow-on patents that cite IBM invention pledge and find that about 5,595 of these patents cite the *Bulletin* in the "Description" section of the patent (although they all cite the *Bulletin* in "Other References"). In the Description section, the invention and its relationship to the prior art are summarized using sentences such as "Bulletin teaches," "Bulletin discloses," "Bulletin shows," or "Bulletin illustrates." These patents and the descriptions of invention pledges contained in them, shown in Table B2 in the Appendix, indicate that the cited invention pledges and the follow-on patents are technologically related.

7. GENERALIZABILITY OF RESULTS TO UNCITED INVENTION PLEDGES

The methodology we develop and employ to examine the nature of the invention pledge program and its consequences for cumulative innovation and the above empirical results are limited to IBM's cited invention pledges. Indeed, it is reasonable to question whether the results generalize to uncited invention pledges, as it is possible that the motivation and consequences of uncited invention pledges are systematically different from those of cited invention pledges. Uncited invention pledges may involve low-quality inventions and ones that are outside IBM's focus in terms of the technology.

Our methodology, which is based on characteristics of patents citing invention pledges, cannot be used to directly to address the issue of generalizability of our results to *uncited* invention pledges. Therefore, we examine whether *cited* and *uncited* invention pledges differ systematically and how this influences our interpretation of the results. We begin by examining the similarities between cited and uncited invention pledges. We first pre-process words in the titles and abstracts of cited and uncited invention pledges by removing stopwords and stemming. Based on the vector of pre-processed word frequency, we compute the cosine similarity between cited and uncited invention pledges. To account for possible changes over time, we do this for each of the four decades of the IBM invention pledge program. If the cited and uncited invention pledges significantly differed in their technological areas, they would show a reasonable separation in the vocabulary. Yet, the results, shown in Table 7, indicate a high degree of similarity (ranging from 0.87 to 0.97 for titles and 0.93 to 0.98 for abstracts) between cited and uncited invention pledges. Higher-quality invention pledges may involve a more elaborate description, as opposed to the simpler descriptions of potentially trivial inventions disclosed in uncited invention pledges. To investigate this, we compute the lengths of titles and abstracts, as measured by word counts, and find little difference between cited and uncited invention pledges on these dimensions.

-- Place Table 7 here --

Next, we examine how the propensity to patent differs at the inventor level based on their *cited* and *uncited* invention pledges. For each inventor at IBM, we examine the correlation between the total number of invention pledges and patents filed in a given year. The results shown in Table 8 indicate that invention pledges and patents are positively and significantly correlated, reflecting that these activities are complementary rather than conflicting. We further divide the total number of invention pledges into counts of *cited* and *uncited* invention pledges. The results, shown in Column 2 of Table 8, indicate that both cited and uncited invention pledges are separately positively and significantly correlated with patenting, although the magnitude of the coefficient estimate for cited is three times larger than that for uncited. We repeat the estimation with originality and generality as alternative measures of patent quality and obtain similar results, as shown in the top and bottom panels of Table A6 in the Appendix.

-- Place Table 8 here --

We also examine whether there are systematic differences between when cited and uncited invention pledges are published over the years. It is possible that invention pledges are concentrated in a small cohort of publication years and significantly differ in terms of the overall distribution of invention pledges. However, Figure 5 shows that the distributions of cited and uncited inventions are broadly similar, particularly before 1989.

-- Place Figure 5 here --

We also examined whether cited invention pledges are more likely to be published in some months of the year, as inventors may hurry to publish their valuable inventions as invention pledges, which requires less time than patenting, before their annual performance reviews. However, we found no such concentration of cited invention pledges in any given month.

Finally, we examine how prior levels of patenting and invention pledging affect current invention pledging level. We reason that the limited number of patents that IBM filed during this period likely reflected a quota of around 500 patents at the organizational level and a limited organizational capacity to draft patent applications and file them with the USPTO (Bhaskarabhatla and Hegde, 2014). This likely, in turn, imposed inventor-level quotas such that no inventor at IBM received a significantly large number of patents in one year even when the inventor generated more patentable inventions than the organization patented. Since there was no such limit on invention pledges published through the *IBM Technical*

Disclosure Bulletin, as IBM invested more and its inventors generated more inventions, the *Bulletin* grew rapidly over the years while patenting stagnated.

The results, shown in the first column of Table A7, indicate a positive correlation between past patenting and current month's invention pledges, as well as between past invention pledges and current month's invention pledges controlling for the number of patents in the current month. When we split the analysis in the first half of the calendar year (January to June) and the second half (July to December) as in the second column, we find intriguing patterns. First, we find a negative correlation between current month's patenting and invention pledges during the second half of the year (but not the first), indicating substitution between patenting and invention pledging later in the year. We also find the number of patents earlier in the year is positively correlated with the number of invention pledges during the second half of the year but not the first half of the year. On the contrary, the number of pledges earlier in the year is positively correlated with the current month's invention pledges throughout the year. These patterns indicate that as inventors received more patents in a year, they were more likely to disclose invention pledges.

Notwithstanding the above analyses, we acknowledge that the generalizability of our results to uncited invention pledges remains a topic for future study.

8. CONCLUSION

Using novel data on IBM's patents and openly pledged inventions, we examine two competing rationales to explain IBM's invention pledge program between 1958 and 1998. We interpret the pieces of evidence we gather as supporting the view that IBM disclosed valuable knowledge as invention pledges, which stimulated innovation in patent subclasses in which it was cited, indicating positive knowledge spillovers from cited invention pledges. IBM subsequently reabsorbed the knowledge rooted in IBM's invention pledges by exploiting its lead-time advantages. Our evidence is less consistent with other rationales for *cited* invention pledges rooted in adverse selection.

Beyond our empirical evidence, several qualitative aspects of the invention pledge program are also inconsistent with the theoretical models discussed above. First, IBM openly pledged more innovations during this period across all its core businesses, whereas the extent of patenting remained nearly constant, at around 500 per year for much of this period, indicating that disclosure was not an intermediate step towards patents but, rather, an expanding program. Indeed, empirical analyses show that as the intensity of antitrust cases increased, IBM increased its level of invention pledges (Bhaskarabhatla and Pennings, 2014). Second, if invention pledges help win patent races, it is unclear why they were prominent in an era when patenting itself was not popular and were shut down when patenting became important. Also, IBM and others are still engaged in technological competition, although they rarely make invention pledges nowadays. Thompson and Kuhn (2020) find that about ten to 11 percent of patents are likely part of a patent race and that the winners conduct more follow-on innovation. However, they find no evidence of the use of open disclosure outside of patents to prolong patent races. Also, IBM's invention pledges, coupled with antitrust enforcement against its dominance—IBM accounted for 85 percent of the domestic market in the mid-1950s (see Usselman (1993))—led to significant entry into disk drives, personal computers and components, contrary to the prediction that open disclosure is expected to soften competition or deter entry.

Our study is not without limitations. First, a major limitation is that our method can be applied only to the study of *cited* invention pledges. However, nearly 80 percent of the invention pledges are not cited and may differ significantly from cited invention pledges. Therefore, our empirical results cannot rule out the possibility that some of these uncited invention pledges relate to less-valuable inventions or inventions on the periphery of IBM's patent portfolio. Similarly, if uncited invention pledges hinder follow-on innovation, then our results on follow-on innovation cannot be generalized to uncited invention pledges. To address this limitation, we conduct additional analyses, but a future study may arrive at more decisive conclusions.

Second, we do not have data on the internal decision process that assigned inventions to patenting, disclosure, and secrecy, and our estimates can be biased due to the non-random nature of IBM's decision process. This limitation gives rise to endogeneity concerns regarding our estimates, given that we do not have the ideal experiment to establish causality. Without a clear understanding of how IBM chose between invention pledges and patenting, our results may be interpreted as correlations, and we should be extremely careful about drawing conclusions and making prescriptive recommendations from the statistical associations in the analysis.

Third, as in the majority of studies set in the period before 2001, we assume that patent applicants added citations to IBM invention pledges. However, in cases in which patent examiners added citations to IBM invention pledges, our interpretation of the results can be incorrect. We expect this to be a less severe limitation given that applicants' share of citations to non-patent literature is generally high (94 percent in a one-percent sample of patents issued in 2007) (Cotropia *et al*., 2013).

Fourth, firms can avoid disclosure by not applying for patent protection, keeping their knowledge secret, and protecting it using laws and institutions governing trade secrecy, but empirically measuring secrecy remains a challenge (Gross, 2019; Ganglmair and Reimers, 2019). We do not measure what IBM kept secret during this period, which can affect our estimates and their interpretation. However, keeping secrets is not optimal for firms such as IBM, which have faced antitrust enforcement rendering their patent rights less valuable and their smaller rivals' patent rights more valuable. This asymmetry in patent rights implies that secrecy can lead to holdup if one of the rival firms subsequently patents an invention that is essential for commercializing IBM's products. Therefore, as part of a strategy to protect its freedom of action, IBM disclosed more and kept few secrets, as evidenced by its large-scale invention pledge activity.

Finally, many of our analyses are based on IBM alone and may not generalize to other innovators, especially those with characteristics distinct from IBM's. The focus on IBM is reasonable, as the firm was a clear leader in voluntary disclosure and had a structured process for distributing its invention pledges. The focus on IBM, however, does raise concerns about external validity, given that IBM was an exceptionally successful and dominant firm in a rapidly expanding industry. IBM created its Invention Pledge program not out of an inherent commitment to open innovation, but as a response to pressure from the Department of Justice. While some of our findings may generalize to Xerox and other firms that employed *Research Disclosure*, our findings may not generalize to other firms that are not in IBM's situation. Future work on other disclosing firms can be fruitful in examining diverse motivations for open disclosure. Moreover, future research can exploit invention pledges to better identify the disclosure effect of invention.

References

Agarwal R, Audretsch D, Sarkar MB. 2010. Knowledge spillovers and strategic entrepreneurship. *Strategic Entrepreneurship Journal* 4(4):271-283.

Aghion P, Dewatripont M, Stein J. 2008. Academic freedom, private-sector focus, and the process of innovation. *RAND Journal of Economics* 39(3):617-635.

Alexy O, George G, Salter A. 2013. Cui Bono? The selective revealing of knowledge and its implications for innovative activity. *Academy of Management Review* 38(2):270-291.

Alexy O, Reitzig M. 2013. Private–collective innovation, competition, and firms' counterintuitive appropriation strategies. *Research Policy* 42(4):895–913.

Alexy O, West J, Klapper H, Reitzig M. 2018. Surrendering control to gain advantage: Reconciling openness and the resource-based view of the firm. *Strategic Management Journal* 39(6):1704-1727.

Anton J, Yao D. 2004. Little patents and big secrets: managing intellectual property. *RAND Journal of Economics* 35(1):1-22.

Armstrong J. 1996. Reinventing research at IBM. Engines of Innovation, US Industrial Research at the End of an Era. Harvard Business School Press.

Arora A, Belenzon S, Patacconi A, Suh J. 2019. The Changing Structure of American Innovation: Some Cautionary Remarks for Economic Growth (No. w25893). *National Bureau of Economic Research*.

Arora A, Belenzon S, Patacconi A. 2018. The decline of science in corporate R&D. *Strategic Management Journal* 39(1):3-32.

Ayvazyan A, Matr S. 2019. Sharing is Caring. Outbound Open Innovation and the Subsequent Innovation Process. Presented at the Academy of Management Conference in Boston.

Baker S, Mezzetti C. 2005. Disclosure as a strategy in the patent race. *Journal of Law and Economics* 48(1):173-194.

Bar T. 2006. Defensive publications in a patent race. *Journal of Economics & Management Strategy* 15(1):229-254.

Bessen J, Maskin E. 2009. Sequential innovation, patents, and imitation. *RAND Journal of Economics* 40(4):611-635.

Bhaskarabhatla A, Hegde D. 2014. An organizational perspective on patenting and open innovation. *Organization Science* 25(6):1744-1763.

Bhaskarabhatla A, Pennings E. 2014. Defensive disclosure under antitrust enforcement. *Industry and Innovation* 21(7):533-552.

Bhattacharya S, Ritter J. 1983. Innovation and communication: Signaling with partial disclosure. *Review of Economic Studies* 50(2):331-346.

Buderi R. 2000. Funding central research. *Research Technology Management* 43(4):18-25.

Cassiman B, Ueda M. 2006. Optimal project rejection and new firm start-ups. *Management Science* 52(2):262-275.

Cockburn I, Henderson R. 1998. Absorptive capacity, coauthoring behavior, and the organisation of research in drug discovery. *Journal of Industrial Economics* 46(2):157-182.

Cohen W, Levinthal D. 1990. Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* 35(1):128-152.

Cohen W, Nelson R, Walsh J. 2000. Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not). NBER Working paper no. 7552.

Contreras J. 2015. Patent pledges. *Arizona State Law Review* 47(3),543-608.

Cotropia C, Lemley M, Sampat B. 2013. Do applicant patent citations matter? *Research Policy* 42(4):844-854.
Dasgupta P, David PA. 1994. Towards a new economics of science. *Research Policy* 23(5): 487-522.
De Fraja G. 1993. Strategic spillovers in patent races. *International Journal of Industrial Organization* 11(1):139-146.
Eaton C, Eswaran M. 2001. Know-how sharing with stochastic innovations. *Canadian Journal of Economics* 34(2):525-548.
Fudenberg, D., Gilbert, R., Stiglitz, J., & Tirole, J. (1983). Preemption, leapfrogging and competition in patent races. European Economic Review, 22(1), 3-31.
Furman JL, Nagler M, Watzinger M. 2018. Disclosure and subsequent innovation: evidence from the patent depository library program. NBER Working Paper No. 24660.
Gambardella A. 1992. Competitive advantages from in-house scientific research—The U.S. pharmaceutical industry in the 1980s. *Research Policy* 21(5):391-407.
Ganglmair, Bernhard and Reimers, Imke. 2022. Visibility of Technology and Cumulative Innovation: Evidence from Trade Secrets Laws. Working Paper.
Gilbert R, Shapiro C, Kaplow L, Gertner R. 1997. Antitrust issues in the licensing of intellectual property: The nine no-no's meet the nineties. Brookings Papers on Economic Activity. Microeconomics 283-349.
Gill D. 2008. Strategic disclosure of intermediate research results. *Journal of Economics & Management Strategy* 17(3):733-758.
Gomory R. 1989. Moving IBM's technology from research to development. *Research Technology Management* 32(6):27-32.
Graham S, Hegde D. 2015. Disclosing patents' secrets. *Science* 347(6219):236-237.
Grant R. 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal* 17(S2):109-122.
Grindley PC, Teece DJ. 1997. Managing intellectual capital: licensing and cross-licensing in semiconductors and electronics. *California Management Review* 39(2):8-41.
Gross DP. 2019. The consequences of invention secrecy: evidence from the USPTO patent secrecy program in World War II (May 12, 2019). Harvard Business School Strategy Unit Working Paper No. 19-090.
Hall B. 2004. Exploring patent explosion. Journal of Technology Transfer 30(2):35-48.
Hall B, Ziedonis R. 2001. The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry, 1979-1995. *RAND Journal of Economics* 32(1):101-128.
Hart DM. 2001. Antitrust and technological innovation in the US: ideas, institutions, decisions, and impacts, 1890–2000. *Research Policy* 30(6):923-936.
Hart DM. 2007. Red, White, and "Big Blue": IBM and the business-government interface in the United States, 1956–2000. *Enterprise & Society* 8(1):1-34.
Hegde D, Luo H. 2018. Patent publication and the market for ideas. *Management Science* 64(2):652-672.
Hegde D, Herkenhoff K, Zhu C. 2018. Patent disclosure and innovation. Available at SSRN: https://ssrn.com/abstract=3158031
Hellmann T, Perotti E. 2011. The circulation of ideas in firms and markets. *Management Science* 57(10):1813-1826.
Hicks, D. 1995. Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and Corporate Change*, 4(2), 401-424.
Hounshell D. 1996. The evolution of industrial research in the United States. Engines of Innovation, US Industrial Research at the End of an Era. Harvard Business School Press.
IBM 2018. IBM breaks records to top U.S. patent list for 25th consecutive year. *IBM Press Release*.
Jaffe A. 1986. Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review* 76(5):984-1001.
Jaffe, A. B., Trajtenberg, M., & Henderson, R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, 108(3), 577-598.

Johnson J. 2014. Defensive publishing by a leading firm. *Information Economics and Policy* 28:15-27.
Lerner, J. (2009). The empirical impact of intellectual property rights on innovation: Puzzles and clues. American Economic Review, 99(2), 343-48.
Li G, Lai R, D'Amour A, Doolin D, Sun Y, Torvik V, Yu A, Fleming L. 2014. Disambiguation and co-authorship networks of the U.S. patent inventor database (1975–2010). *Research Policy* 43(6):941-55.
Lichtman D, Baker S, Kraus K. 2000. Strategic disclosure in the patent system. *Vanderbilt Law Review* 53(6):2175–2217.
Lopatka, J. E. (2000). United States v. IBM: A Monument of Arrogance. Antitrust Law Journal, 68, 145.
McGroddy J. 1998. Visiting committee on advanced technology congressional testimony. Committee on Science, U.S. House of Representatives. Retrieved from https://www.nist.gov/director/vcat/vcat-1998-congressional-testimony
McGroddy, J. 2001. Raising mice in the elephants' cage. Lewis M. Branscomb and Philip Auerswald, Taking Technical Risks: How Innovators, Executives, and Investors Manage High Tech Risks.
Merton RK. 1973. The sociology of science: theoretical and empirical investigations. Univ. of Chicago Press.
Moser P, Voena A. 2012. Compulsory licensing: evidence from the Trading with the Enemy Act. *American Economic Review* 102(1):396-427.
Mowery DC. 1992. The US national innovation system: origins and prospects for change. *Research Policy* 21, 125–144.
Mullin, W.P. Snyder, C.M., Forthcoming. A Simple Method for Bounding the Elasticity of Growing Demand with Applications to the Analysis of Historic Antitrust Cases. *American Economic Journal: Microeconomics*.
Murray F, Aghion P, Dewatripont M, Kolev J, Stern, S. 2016. Of mice and academics: examining the effect of openness on innovation. *American Economic Journal: Economic Policy* 8(1):212-252.
Nelson RR. 1992. What is commercial and what is public about technology and what should be? In N. Rosenberg, R. Landau, and D. Mowery (eds.), Technology and the Wealth of Nations. Stanford Press.
Narin, F., & Noma, E. (1985). Is technology becoming science?. Scientometrics, 7(3-6), 369-381.
Pacheco-de Almeida G, Zemsky PB. 2012. Some like it free: Innovators' strategic use of disclosure to slow down competition. *Strategic Management Journal* 33(7):773-793.
Parchomovsky G. 2000. Publish or perish. *Michigan Law Review* 98(4):926-952.
Parchomovsky G, Mattioli M. 2011. Partial patents. *Columbia Law Review* 111,207-254.
Pugh EW. 1995. Building IBM: Shaping an Industry and Its Technology. MIT Press.
Reinganum, J. F. 1989. "The Timing of Innovation: Research, Development, and Diffusion." In Handbook of Industrial Organization, edited by Richard Schmalensee and Robert Willig, pp. 1:849–908. New York: Elsevier Science Publishing Co.
Rosenberg N. 1990. Why do firms do basic research (with their own money)? *Research Policy* 19(2):165–174.
Rosenbloom RS, Spencer WJ. 1996. Engines of innovation: US industrial research at the end of an era. Harvard Business School Press.
Sakakibara, M., & Branstetter, L. (2001). Do stronger patents induce more innovation? Evidence from the 1988 Japanese patent law reforms. RAND Journal of Economics, 77-100.
Sarasohn H. 1973. The Technical Journal – who needs it? *IEEE Transactions on Professional Communication* 16(3):129-130.
Stern S. 2004. Do scientists pay to be scientists? *Management Science* 50(6):835-853.
Teece DJ. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* 15(6):285-305.
Thompson, N. C., & Kuhn, J. M. (2020). Does Winning a Patent Race lead to more follow-on Innovation?. Journal of Legal Analysis, 12, 183-220.

Usselman, S. 1993. IBM and its imitators: Organizational capabilities and the emergence of the international computer industry. Business and Economic History, 1-35.
Usselman S. 2009. Unbundling IBM: antitrust and the incentives to innovation in American computing, in Clarke *et al*., eds., The Challenge of Remaining Innovative, Stanford University Press, pp. 249-279.
Watzinger M, Fackler TA, Nagler M, Schnitzer M. 2017. How antitrust enforcement can spur innovation: Bell labs and the 1956 consent decree. CEPR Discussion Paper 11793.
Wen W, Ceccagnoli M, Forman C. 2015. Opening up intellectual property strategy: Implications for open source software entry by start-up firms. *Management Science* 62(9):2668-2691.
Williams H. 2017. How do patents affect research investments? *Annual Review of Economics* 9(1):441-469.
Yang HY, Phelps C, Steensma HK. 2010. Learning from what others have learned from you: the effects of knowledge spillovers on originating firms. *Academy of Management Journal* 53(2):371-389.
Ziedonis RH. 2004. Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms. *Management Science* 50(6):804-820.

Figures and Tables

Figure 1. Distribution of invention pledges by IBM over time

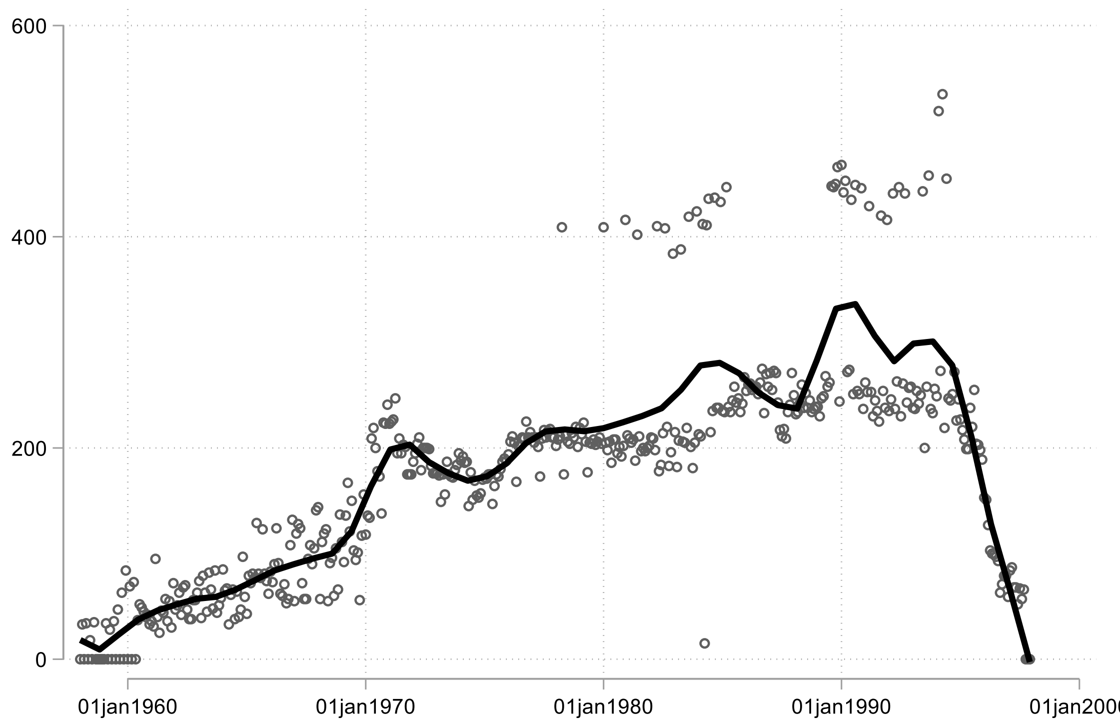


Figure 1 Notes—The scatter plot shows the number of IBM invention pledges published in its Technical Disclosure Bulletin during 1958-1998. The count is monthly and contains the number of invention pledges in each volume of the bulletin. A smoothed line is also plotted.

Figure 2. Diffusion of knowledge originated in IBM patents and invention pledges

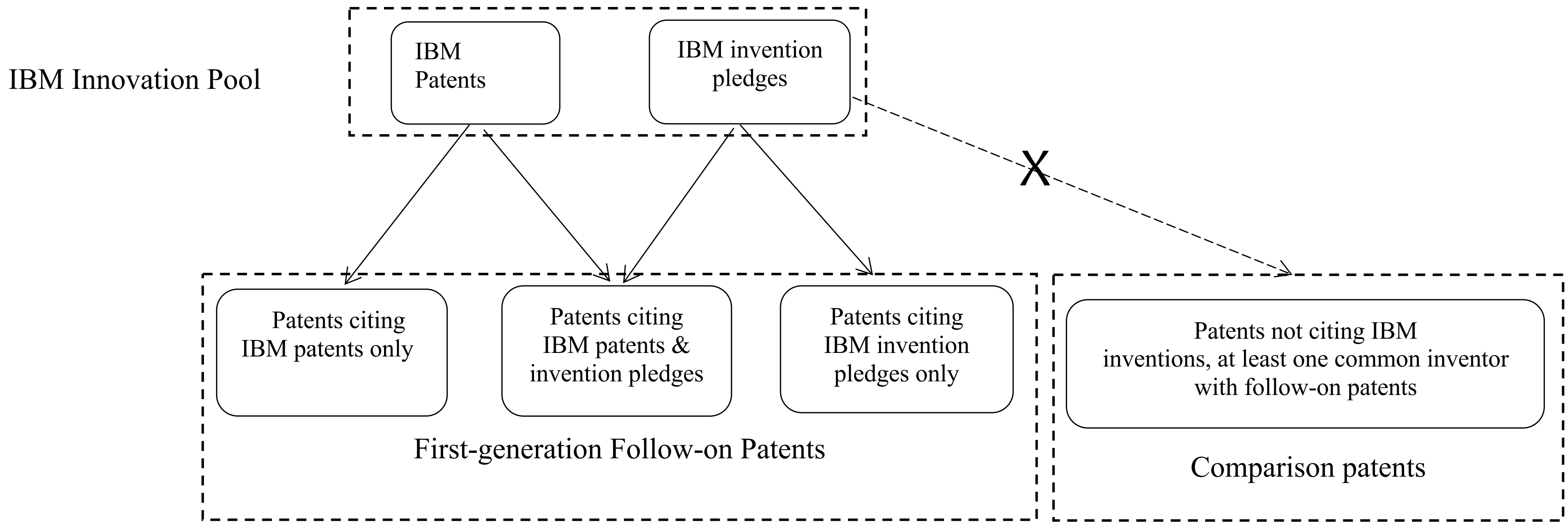
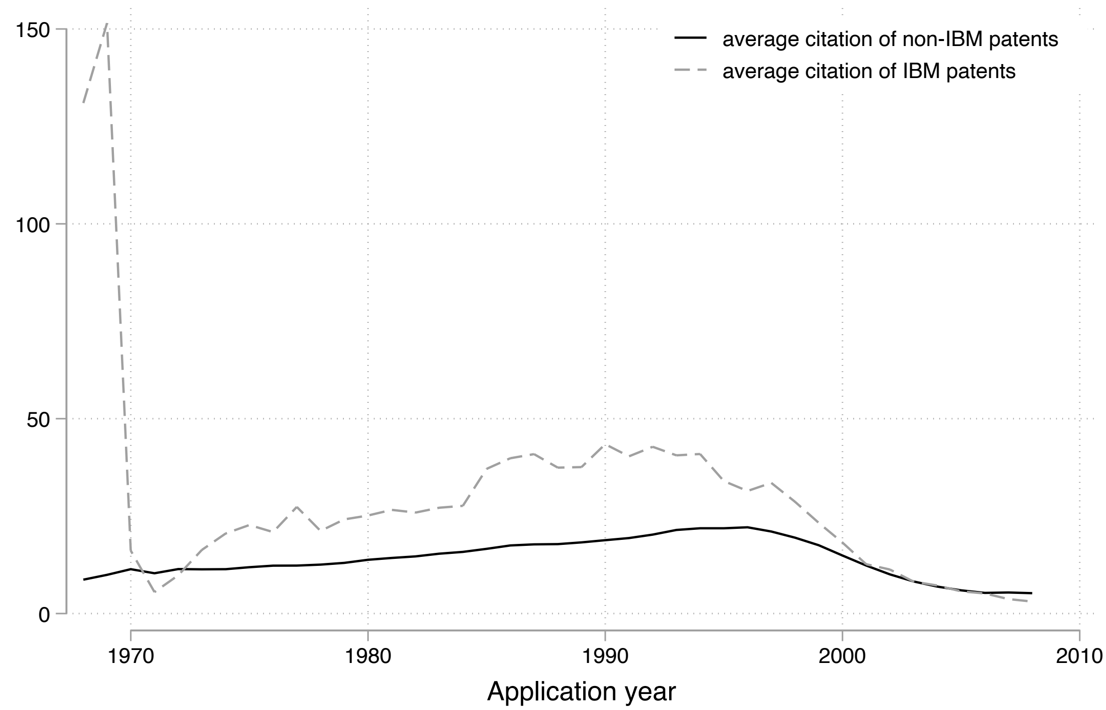


Figure 2 Notes—The figure shows the diffusion of knowledge originated in IBM innovations, as captured by citations. The root of the diffusion tree consists of all IBM innovations, including IBM patents and IBM invention pledges. The first generation of the citation tree includes patents that cite either IBM patents or IBM invention pledges or both, and include both IBM patents and patents by other assignees. The comparison sample consists of patents with at least one inventor in common with the first-generation follow-on patents, but not citing IBM invention pledges or patents.

Figure 3. Forward and backward citations to IBM and non-IBM Patents

Panel A: Forward Citations



Panel B: Backward Citations

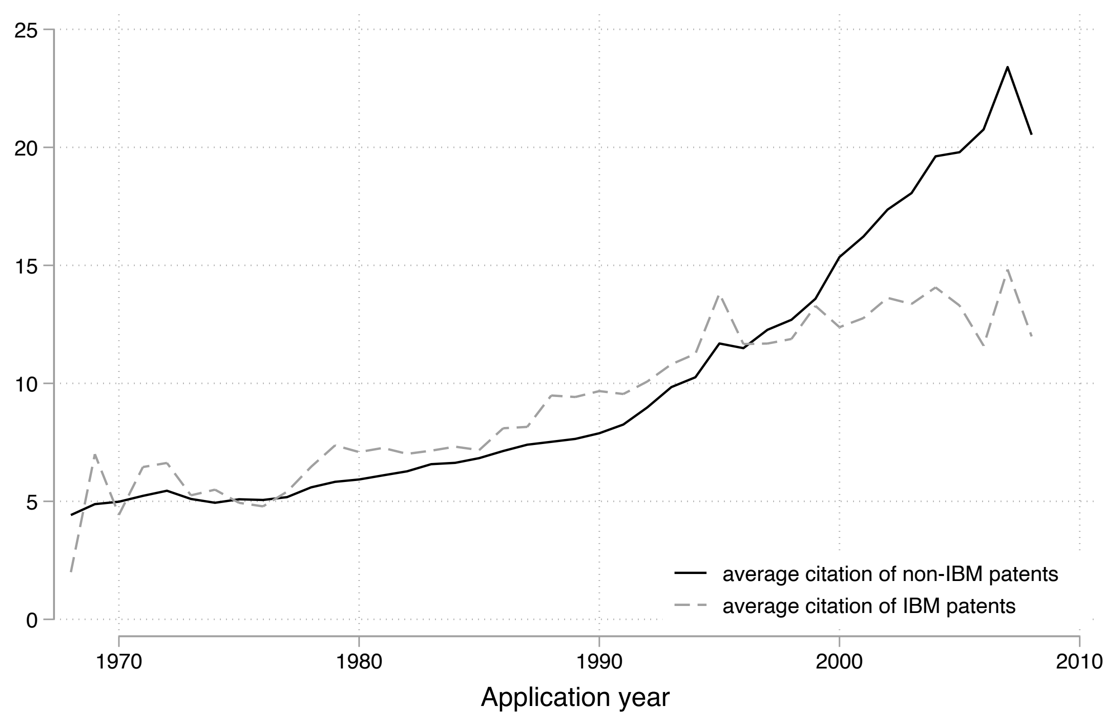
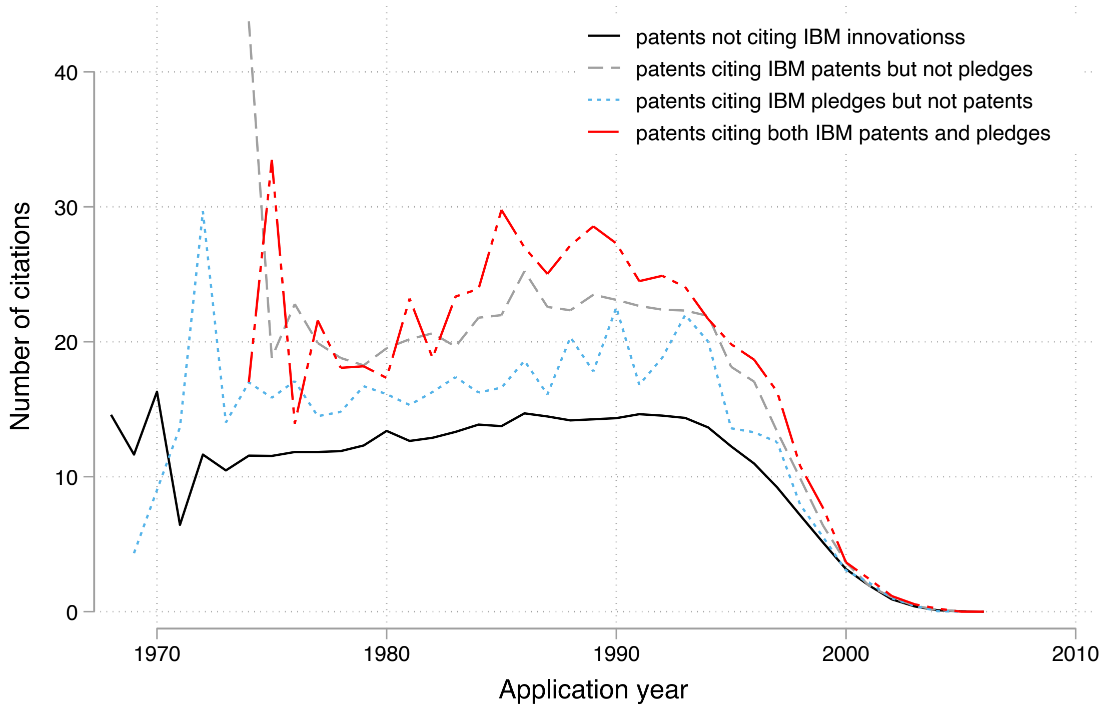


Figure 4. Forward and backward citations received by patents that cite IBM Invention Pledges, IBM Patents, and non-IBM Inventions (grouped by patent application year)

Panel A: Forward Citations



Panel B: Backward Citations

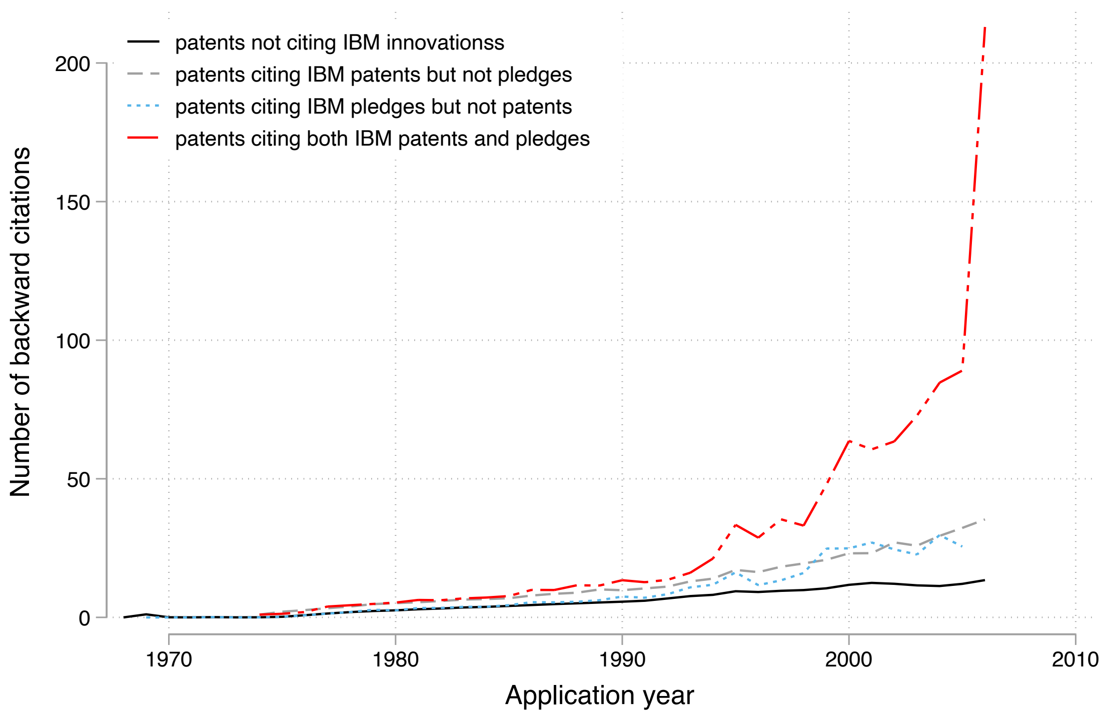


Figure 5. Distribution of cited and uncited invention pledges by IBM over time

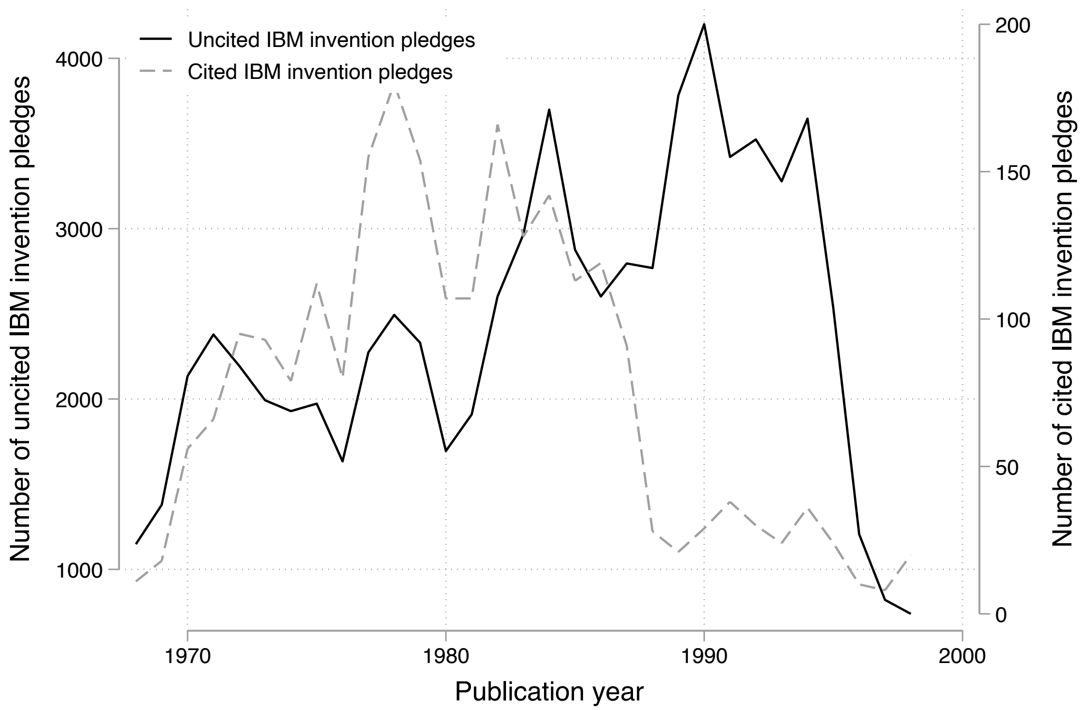


Table 1. Description of variables

Variable	Description
Innovation and output measures	
Forward citations	Total number of forward citations to the patent excluding self-citations
Distance to IBM Patents	A patent's technical distance to IBM's patent portfolio by Jaffe (1986)'s distance measure
Generality	One minus the sum of squared shares of forward citations in each patent class
Originality	One minus the sum of squared shares of backward citations in each patent class
Patent Scope	Number of claims contained in the patent
Lead time advantage measures	
Pledge citation lag (measure 1)	Patent Application Year – Most Recent Cited Invention Pledge's Publication Year
Pledge citation lag (measure 2)	Patent Application Year – Oldest Cited Invention Pledge's Publication Year
Pledge citation lag (measure 3)	Patent Application Year – Average Cited Invention Pledge's Publication Year
Baseline Controls	
Backward citations	Total number of citations made by the patent
Number of patent classes in backward citations	Total number of patent class in the patents cited by the focal patent
Number of non-patent citations	Total number of non-patent citations made by the patent
Number of inventors	Total number of inventors on the focal patent
Patent stock	Sum of patents over the past three years
Additional Controls for Compustat firms	
Firm age	Firm's age in year when a patent is filed
R&D Dummy	Dummy of whether the firm reports R&D expenditure in the year
R&D intensity	Three R&D Expenditures/Sales averaged over the past three years
Capital intensity	PP&E/Sales averaged over the past three years, where PP&E is Property, Plant and Equipment expenditure
Sales	Firm's averaged sales over the past three years
Operating income change	Change in operating income of firm's averaged over the past three years
Employees	Number of employees in the year

Table 2. Summary statistics for first generation follow-on patents

Variable	Table 3, Panel A, including IBM	Table 3, Panel A, excluding IBM	Table 3, Panel B, including IBM	Table 3, Panel B, excluding IBM
Observations	1,593,251	1,510,291	674,001	592,925
Innovation and output measures				
If citing IBM Patent (& not Pledge)	0.19 [0.39]	0.17 [0.38]	0.25 [0.43]	0.20 [0.40]
If citing IBM Pledge (& not patent)	0.01 [0.11]	0.01 [0.07]	0.01 [0.11]	0.01 [0.11]
If citing IBM Patent & Pledge	0.01 [0.12]	0.01 [0.09]	0.03 [0.16]	0.01 [0.11]
Forward citations	10.78 [16.93]	10.64 [16.78]	11.20 [16.66]	10.92 [16.24]
Distance to IBM Patents	0.76 [0.26]	0.77 [0.25]	0.71 [0.28]	0.74 [0.27]
Backward citations	9.12 [15.64]	9.06 [15.87]	9.71 [15.29]	9.60 [15.82]
Number of patent classes in backward citations	2.73 [2.36]	2.73 [2.37]	2.81 [2.31]	2.80 [2.34]
Number of non-patent citations	2.35 [9.50]	2.33 [9.5684]	2.22 [8.76]	2.14 [8.80]
Number of inventors	3.85 [2.54]	3.86 [2.54]	3.93 [2.71]	3.91 [2.75]
Firm characteristics				
Firm age	--	--	2.69 [0.51]	2.66 [0.52]
Dummy R&D	--	--	0.99 [0.11]	0.99 [0.11]
R&D intensity	--	--	0.07 [0.09]	0.07 [0.09]
Capital intensity	--	--	0.27 [0.14]	0.27 [0.15]
Sales	--	--	9.89 [1.62]	9.67 [1.60]
Operating income change	--	--	0.14 [0.55]	0.10 [0.50]
Employees	--	--	4.18 [1.50]	3.97 [1.48]
Patent stock	--	--	7.93 [1.63]	7.71 [1.60]

Table 2. Notes—The table presents mean and standard deviation (in brackets) of variables on first-generation patents for the whole sample (Panel A) and for the subsample of firms listed in the Compustat database (Panel B).

Table 3. Estimates for the differences among follow-on patents and comparison patents

Column	1	2	3	4
	Panel A: All Firms			
Dependent Variable	Log Forward Citations		Distance to IBM Patents	
Comparison sample	At Least One Common Inventor Among Follow-on and Comparison Patents			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (& not pledge)	0.065***	0.069***	-0.053***	-0.053***
	[0.007]	[0.006]	[0.002]	[0.002]
Citing IBM Pledge (& not patent)	**0.058******	**0.061******	**-0.026******	**-0.028******
	[0.013]	[0.014]	[0.004]	[0.003]
Citing IBM Patent & Pledge	0.036	-0.002	-0.043***	-0.043***
	[0.026]	[0.025]	[0.004]	[0.006]
Controls	Yes	Yes	Yes	Yes
Inventor Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,593,251	1,510,291	1,593,251	1,510,291
R-squared	0.4754	0.4763	0.6704	0.6572
	Panel B: Compustat Firms			
Dependent Variable	Log Forward Citations		Distance to IBM Patents	
Comparison sample	At Least One Common Inventor Among Follow-on and Comparison Patents			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (& not pledge)	0.056***	0.063***	-0.043***	-0.043***
	[0.010]	[0.009]	[0.003]	[0.004]
Citing IBM Pledge (& not patent)	**0.037***	**0.042**	**-0.012****	**-0.017******
	[0.021]	[0.028]	[0.006]	[0.004]
Citing IBM Patent & Pledge	0.043*	-0.005	-0.028***	-0.023***
	[0.023]	[0.039]	[0.004]	[0.008]
Controls	Yes	Yes	Yes	Yes
Additional Compustat Controls	Yes	Yes	Yes	Yes
Inventor Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	674,001	592,925	674,001	592,925
R-squared	0.4946	0.4974	0.7117	0.6993

Table 3 Notes— The table presents OLS estimates of the forward citation (Columns 1 and 2) and distance to IBM patent portfolio (Columns 3 and 4) for patents applied between 1976 and 2006. Controls included but not reported in Panel A are: log(number of backward citations); log(number of patent classes in backward citations); log(number of non-patent citations); log(number of non-patent citations in the patent's backward citations); log(number of inventors). Additional Compustat controls included in Panel B are: whether the firm reports R&D; log(firm age); log(number of employees); three year moving average of log sales; three year moving average of change in operating income; stock of firm patents; three year moving average of R&D intensity; three year moving average of capital intensity. The constant term is included in all estimations but not reported in the table. Standard errors are clustered at firm level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Estimates for the differences among follow-on patents and comparison patents (including moderating effect of pledge quality)

Column	1	2	3	4
Panel A: All Firms				
Dependent Variable	Log Forward Citations		Distance to IBM Patents	
Comparison sample	At Least One Common Inventor Among Follow-on and Comparison Patents			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (& not pledge)	0.065***	0.069***	-0.053***	-0.053***
	[0.007]	[0.006]	[0.002]	[0.002]
Citing IBM Pledge (& not patent)	**-0.038**	**-0.005**	**-0.038*****	**-0.042*****
	[0.049]	[0.041]	[0.006]	[0.005]
Citing IBM Patent & Pledge	0.069	0.136***	-0.057***	-0.068***
	[0.062]	[0.049]	[0.010]	[0.008]
Citing IBM Pledge (& not patent) × Pledge Quality	**0.068***	**0.046**	**0.009*****	**0.010*****
	[0.036]	[0.032]	[0.003]	[0.004]
Citing IBM Patent & Pledge × Pledge Quality	-0.019	-0.076**	0.009	0.014***
	[0.048]	[0.031]	[0.005]	[0.005]
Controls	Yes	Yes	Yes	Yes
Inventor Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,593,251	1,510,291	1,593,251	1,510,291
R-squared	0.4755	0.4763	0.6704	0.6572
Panel B: Compustat Firms				
Dependent Variable	Log Forward Citations		Distance to IBM Patents	
Comparison sample	At Least One Common Inventor Among Follow-on and Comparison Patents by Compustat Firms			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (& not pledge)	0.056***	0.063***	-0.043***	-0.043***
	[0.009]	[0.008]	[0.003]	[0.004]
Citing IBM Pledge (& not patent)	**-0.132****	**-0.076**	**-0.021*****	**-0.027*****
	[0.062]	[0.046]	[0.007]	[0.008]
Citing IBM Patent & Pledge	0.005	0.078	-0.045***	-0.061***
	[0.069]	[0.084]	[0.010]	[0.009]
Citing IBM Pledge (& not patent) × Pledge Quality	**0.119****	**0.081****	**0.006**	**0.007**
	[0.048]	[0.039]	[0.005]	[0.006]
Citing IBM Patent & Pledge × Pledge Quality	0.024	-0.045	0.011	0.021***
	[0.052]	[0.050]	[0.007]	[0.005]
Controls	Yes	Yes	Yes	Yes
Additional Compustat Controls	Yes	Yes	Yes	Yes
Inventor Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	674,001	592,925	674,001	592,925
R-squared	0.4947	0.4974	0.7117	0.6994

Table 4 Notes— See detailed notes for Table 3. The constant term is included in all estimations but not reported in the table. Pledge Quality is measured as the number of times an invention pledge is cited. Standard errors are clustered at firm level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Estimates of IBM's lead time advantage in follow-on invention

Column	1	2	3	4	5
Panel A: All Firms					
Dependent Variable	Log Forward Citations	Log Number of IBM Invention Pledges Cited	Patent Application Year – Most Recent Cited Invention Pledge Year	Patent Application Year – Oldest Cited Invention Pledge Year	Patent Application Year – Average Cited Invention Pledge Year
Sample	All First-Generation Patents Citing IBM Invention Pledges				
IBM (=1 if yes, 0 otherwise)	**0.198*****	**0.176*****	**-2.010*****	**-1.139*****	**-1.613*****
	[0.024]	[0.015]	[0.187]	[0.204]	[0.180]
IBM × After 1989	0.002	-0.028	-0.673***	-0.604**	-0.614***
	[0.032]	[0.018]	[0.214]	[0.265]	[0.226]
Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Patent Class Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	48,153	48,153	48,153	48,153	48,153
R-squared	0.082	0.088	0.332	0.326	0.339
Panel B: Compustat Firms					
Sample	All First-Generation Patents by Compustat firms Citing IBM Invention Pledges				
IBM (=1 if yes, 0 otherwise)	**0.179*****	**0.216*****	**-2.058*****	**-1.621*****	**-1.812*****
	[0.035]	[0.018]	[0.292]	[0.313]	[0.291]
IBM × After 1989	-0.019	-0.021	-0.894***	-0.850**	-0.812***
	[0.040]	[0.022]	[0.268]	[0.333]	[0.281]
Controls	Yes	Yes	Yes	Yes	Yes
Additional Compustat Controls	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Patent Class Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	16,017	16,017	16,017	16,017	16,017
R-squared	0.112	0.115	0.348	0.314	0.327

Table 5 Notes—This table presents OLS estimates on first-generation patents citing IBM invention pledges for patents applied between 1976 and 2006. To examine lead time advantage in follow-on invention, we analyze a dataset of all patents citing IBM invention pledges. The dependent variables in these regressions are respectively: log number of forward citations in column (1), log number of IBM invention pledges cited in the first-generation patent in column (2), the number of years between the patent application year and the year in which the most recent cited invention pledge is published (column 3), the number of years between the application year and the oldest cited invention pledge (column 4), and the average gap across all cited invention pledges (column 5). To accurately compute citation gap measures, we extract the publication year for each cited IBM invention pledge. When publication year is missing, we exploit volume numbers of the invention pledges, with volume 1 corresponding to 1958, 2 to 1959 and so on, and we identify publication years for more than 99 percent of the cited invention pledges. Controls include a dummy for US firms (if patent assignee type is US firm) and number of invention pledges cited in columns 3, 4, and 5. Compustat controls include R&D; log(firm age); log(number of employees); three year moving average of log sales; three year moving average of change in operating income; stock of firm patents; three year moving average of R&D intensity; three year moving average of capital intensity. Constant is included but not reported in the table in all specifications. Standard errors clustered at the patent class level are reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Estimates on the quantity and quality of patents filed in subclasses citing invention pledge in each patent class

Column	1	2	3	4
Dependent Variable	$Log(Patents)_{st}$	$Log(Cites)_{st}$	$Log(Patents)_{st}$	$Log(Cites)_{st}$
Sample	All Patent Classes, 1976-2011			
N of Patents Citing IBM Pledges in subclass *s* in year *t*-1	**0.004****	**0.019****		
	[0.002]	[0.007]		
N of Patents Citing IBM Patents in subclass *s* in year *t*-1	0.003***	0.002*		
	[0.001]	[0.001]		
Average N of Patents Citing IBM Pledges in subclass *s* during *t*-1, *t*-2, *t*-3			**0.007****	**0.038*****
			[0.003]	[0.013]
Average N of Patents Citing IBM Patents in subclass *s* during *t*-1, *t*-2, *t*-3			0.003***	0.000
			[0.001]	[0.001]
Observations	469,722	469,722	469,722	469,722
R-squared	0.607	0.456	0.607	0.456
Dependent Variable	$Log(Patents)_{st}$	$Log(Cites)_{st}$	$Log(Patents)_{st}$	$Log(Cites)_{st}$
Sample	All Patent Classes, 1976-2011			
N of Patents Citing Xerox Pledges in subclass *s* in year *t*-1	**0.005**	**0.078*****		
	[0.019]	[0.029]		
N of Patents Citing Xerox Patents in subclass *s* in year *t*-1	0.005***	0.004***		
	[0.001]	[0.001]		
Average N of Patents Citing Xerox Pledges in subclass *s* during *t*-1, *t*-2, *t*-3			**0.006**	**0.142****
			[0.036]	[0.060]
Average N of Patents Citing Xerox Patents in subclass *s* during *t*-1, *t*-2, *t*-3			0.005***	0.003*
			[0.001]	[0.002]
Observations	469,722	469,722	469,722	469,722
R-squared	0.606	0.456	0.606	0.456
Dependent Variable	$Log(Patents)_{st}$	$Log(Cites)_{st}$	$Log(Patents)_{st}$	$Log(Cites)_{st}$
Sample	All Patent Classes, 1976-2011			
N of Patents Citing Research Disclosure in subclass *s* in year *t*-1	**0.050*****	**0.080*****		
	[0.010]	[0.025]		
Average N of Patents Citing Research Disclosure in subclass *s* during *t*-1, *t*-2, *t*-3			**0.078*****	**0.091***
			[0.024]	[0.049]
Observations	469,722	469,722	469,722	469,722
R-squared	0.605	0.456	0.605	0.456

Table 6 Notes—The table presents OLS estimates on the quantity and quality of patents filed in subclasses citing IBM invention pledges, Xerox invention pledges, or Research Disclosure in each patent class. The dependent variable is the log number of patents filed in each patent subclass *s* in year *t* in columns 1 and 3 and log citations of patents filed in subclass *s* in year *t* in columns 2 and 4. The sample for analysis is 1976-2011. Limiting the sample period to 1976-2006 does not change the results qualitatively. The sample contains all patent classes. Constant and log(Stock of patents) are included but not reported in the table. We include three-digit patent class-year fixed effects. Standard errors clustered at the patent subclass level are reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Comparison of cited and uncited invention pledges

Years	Average Abstract Word Counts (Std. Dev)		Average Title Word Counts (Std. Dev)		Abstract Similarity between cited and uncited	Title Similarity between cited and uncited
	Cited	Uncited	Cited	Uncited		
1958 to 1968	22.33(5.47)	22.12(5.63)	3.84(1.45)	3.87(1.50)	0.93	0.87
1969 to 1978	21.92(5.01)	21.99(4.97)	4.69(1.82)	4.50(1.78)	0.96	0.93
1979 to 1988	23.66(2.37)	23.14(3.27)	5.46(2.22)	5.17(2.17)	0.96	0.94
1989 to 1998	21.41(4.48)	20.90(4.88)	6.30(2.37)	5.83(2.25)	0.97	0.95
Total	22.01(4.61)	21.74(4.81)	5.29(2.23)	5.10(2.16)	0.98	0.97

Table 7 Notes—The table presents the length of title and abstract (measured in words) of cited and uncited invention pledges, as well as the similarity in frequently occurring words in title and abstract between cited and uncited invention pledges, for each of the four decades of the IBM invention pledge program.

Table 8. Estimates of correlation between invention pledges and the average citation of patents in the year

Column	1	2
Variables	Average citation of patents published in the year	
Total number of invention pledges published up to the focal year	0.208***	
	[0.010]	
Number of cited invention pledges published up to the focal year		0.556***
		[0.115]
Number of uncited invention pledges published up to the focal year		0.176***
		[0.0123]
Number of inventors	36,021	36,021
Inventor FE	Yes	Yes
Year FE	Yes	Yes
Observations	1,837,071	1,837,071
R-squared	0.067	0.067

Table 8 Notes— The table presents OLS estimates of the average quality (measured with citations) of patents at inventor-year level as the dependent variable, based on a balanced yearly panel of inventors for patents applied between 1976 and 2006. In years when the inventor does not have any patents, the dependent variable is set as zero. The constant term is included in all estimations but not reported in the table. Standard errors are clustered at inventor level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Appendix

Appendix A. Data Construction

1 IBM innovation pool (the root innovations)

The root IBM innovation contains all IBM patents and IBM invention pledges. The purpose of this step is to identify the root innovation and obtain their characteristics.

IBM patents in the root

We search NBER US patent database and identify all patents whose assignee is IBM. Root patents in the IBM innovation pool should not cite other IBM patents directly, so we search NBER US patent citation database to identify IBM patents that cite other IBM patents and exclude these patents from the set of root IBM patents.

Patent characteristics such as application year, patent subclass, and patent assignee are obtained from NBER US patent database. We add inventor ID from Li *et al*. (2014) to the root IBM patents based on patent numbers. For assignees in the Compustat database, we obtain financial information from Compustat database.

IBM invention pledges in the root

We obtain the complete list of all IBM invention pledges by manually scanning and coding the "Table of Contents" section of the IBM TDB and from IP.com. This resulted in 83,902 IBM pledges during 1958-1998. For each IBM invention pledge, we extract pledge ID, title, abstract, each inventor's full last name and first name initial, publication year, volume and issue number, and pages. Unlike patents, IBM invention pledges do not come with classifications. Our data do not include the body of IBM invention pledges.

In order to study the diffusion of knowledge originated in IBM innovations, we need to identify IBM invention pledges cited by any US patents. However, non-patent citations could be in various formats. In many cases, the complete information of the cited invention pledge is not included in citations. To identify the accurate citation relationship between IBM invention pledges and patents, we first identify non-patent citations that are likely related to IBM invention pledges. We then match the potential non-patent citations with details of IBM disclosures to confirm the citation relationship between patents and IBM invention pledges. Below we detail the steps.

Step 1. We conduct a thorough search on a series of key words in all US patents' non-patent citations, such as "IBM", "International Business Machines", "Disclosure", "Technical Disclosure Bulletin", "TDB" and "IBM TDB". We retain all non-patent citations that include any of the key words, and denote them as potential matches.

Step 2. For all potential matches, we matched invention pledges and patents' non-patent citations selected from Step 1. We compute the Jaro-Winkler distance and Jaccard distance between the body of non-patent citations and a string that includes invention pledge's title, inventor name(s) publication year, issue and volume number, and pages. First, among all IBM invention pledges matched with one non-patent citation, we keep the IBM invention pledges with the largest Jaro-Winkler distance. Second, if the Jaccard distance is no less than 0.5, the pair is identified as match. If the Jaccard distance is less than 0.4, the pair is identified as unmatch. If the Jaccard distance is less than 0.5, but greater than 0.4, the pair is manually checked to mark the matches. In the end, among 83,902 IBM invention pledges, we identify 17,225 invention pledges cited by 28,548 patents. IBM disclosures that are not cited by any patents are excluded from root innovations.

Next, we add inventor ID from Li *et al*. (2014) to the 17,225 root invention pledges. We do so based on inventor name, pledge year, and title, with the following steps.

Step 1. We match all IBM pledge inventors with IBM patent inventors on full last name and first name initial. For one-to-one matches, we assign the IBM patent inventor ID to the matched pledge inventor.

Step 2. For IBM pledge inventors who fail to match with any IBM patent inventors, we match the pledge inventor with all US patent inventors based on full last name and first name initial. For one-to-one matches, we assign the patent inventor ID to the matched IBM pledge inventor.

Step 3. For non-unique matches in Step 1 and Step 2, we need to identify the correct match. Because invention pledges and patents by the same inventor are likely in the similar technical field, we expect the titles and abstracts of invention pledges and patents to share similar terms. Hence, we compare the similarity of the titles/abstracts of the inventor's pledge and patent matches. First, for each pair of pledge and patent, we compute the Jaccard distance and Jaro-Winkler distance between the titles and between the abstracts. Second, we keep the pairs for which the title Jaccard distance is smaller than 0.7 or the title Jaro-Winkler distance is smaller than 0.2. Third, for the remaining pairs, we construct a score based on the four distances (two types of distances for both abstract and title). If a pair's distance is the smallest among all possible matches, we add 1 to the score, so the highest possible score is 4 and the lowest possible score is 0. We keep the match with the highest total score and assign the corresponding inventor ID to the pledge inventor.

If no potential matches are found, then we consider this inventor to be a pledge-only inventor, whose mobility cannot be tracked.

In total, there are 39,938 invention pledges inventors, of which 20,788 are matched with patent inventor ID either uniquely or non-uniquely. After refinement, 16,539 pledge inventors are assigned a unique inventor ID and there are 23,399 pledge-only inventors. After removing citing patents whose assignee ID is missing, our final sample of root innovation pool contains 15,744 IBM invention pledges and 45,146 IBM patents.

2 First-generation patents

The first-generation patents are those that cite either IBM patents or IBM invention pledges or both. NBER patent citation database enables us to identity the first-generation patents citing root IBM patents. Meanwhile, in the previous step, we have identified patents citing IBM invention pledges. The combination of these two groups of patents is the first-generation patents. We obtain patent characteristics similarly as for the root patents.

3. Inventor name matching

Because IBM invention pledges only provide the last name and first name initials (instead of full first names) of the inventors, matching the names of such inventors across the patent and invention pledge databases is more challenging and difficult to disambiguate. We follow several steps to match the names. First, for each IBM inventor with a patent, we assign a unique inventor identifier provided by Li *et al*. (2014) by matching on inventor name and patent number. Second, for each IBM inventor with an invention pledge, we search among IBM patents for potential inventor name matches. If there is a unique match, then the identifier of the patent's inventor is assigned to the inventor of the pledge. If there are multiple matches, we compute a text-based similarity measure comparing the titles of this inventor's pledges and potential patent matches, and match the inventor of the pledge to the inventor of the patent with the highest similarity. We also manually verified the quality of the match. Third, for inventors with an IBM invention pledge but no potential match among inventors of IBM patents, we look for matches outside IBM patents. We do so because it is possible that although such inventors did not file for patents while working at IBM, they could have filed for patents either before joining or after leaving IBM. Hence, we search in the entire patent database for potential matches. Again, we compute a text-based similarity measure comparing the titles of such inventors' disclosures and potential patent matches outside of IBM and assign matches using similarity scores. If no potential matches are found, then we consider this inventor to be an invention-pledge-only inventor. After identifying unique inventors, we are able to analyze the correlation between the propensity to patent and openly disclose invention pledges.

Table A1. Robustness check using alternative comparison sample: Estimates for the differences among follow-on patents and comparison patents

Column	1	2	3	4
Panel A: All Firms				
Dependent Variable	Log Forward Citations		Distance to IBM Patents	
Comparison sample	Common IPC Class, Assignee, Application Year Among Follow-on and Comparison Patents			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (and not disclosure)	0.064***	0.068***	-0.006***	-0.006***
	[0.0050]	[0.0037]	[0.0003]	[0.0009]
Citing IBM Disclosure (and not patent)	**0.037*** **	**0.040*** **	**0.004* **	**0.002**
	[0.0082]	[0.0085]	[0.0021]	[0.0019]
Citing IBM Patent & Disclosure	0.041***	0.051***	0.003	-0.000
	[0.0108]	[0.0118]	[0.0024]	[0.0026]
Controls, Assignee-IPC Fixed Effects, Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	566,127	535,232	566,127	535,232
R-squared	0.2194	0.2179	0.9048	0.9026
Panel B: Compustat Firms				
Dependent Variable	Log Forward Citations		Distance to IBM Patents	
Comparison sample	Common IPC Class, Assignee, Application Year Among Follow-on and Comparison Patents			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (and not disclosure)	0.048***	0.052***	-0.005***	-0.005***
	[0.0055]	[0.0049]	[0.0009]	[0.0011]
Citing IBM Disclosure (and not patent)	**0.041*** **	**0.054*** **	**0.007*** **	**0.005* **
	[0.0148]	[0.0122]	[0.0026]	[0.0028]
Citing IBM Patent & Disclosure	0.014	0.025	0.003*	0.002
	[0.0110]	[0.0181]	[0.0019]	[0.0035]
Controls, Compustat Controls, Assignee-IPC Fixed Effects, Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	284,162	254,219	284,162	254,219
R-squared	0.2450	0.2451	0.9177	0.9148

Table A1 Notes— The table presents OLS estimates of the forward citation (Columns 1 and 2) and distance to IBM patent portfolio (Columns 3 and 4) for patents applied between 1976 and 2006. Controls included but not reported in Panel A are: log(number of backward citations); log(number of patent classes in backward citations); log(number of non-patent citations); log(number of non-patent citations in the patent's backward citations); log(number of inventors). Additional Compustat controls included in Panel B are: whether the firm reports R&D; log (firm age); log (number of employees); three year moving average of log sales; three year moving average of change in operating income; stock of firm patents; three year moving average of R&D intensity; three year moving average of capital intensity. The constant term is included in all estimations but not reported in the table. Standard errors are clustered at firm level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A2. Robustness check using alternative comparison group: Estimates for the differences among follow-on patents and comparison patents (including moderating effect of pledge quality)

Column	1	2	3	4
Panel A: All Firms				
Dependent Variable	Log Forward Citations		Distance to IBM Patents	
Sample	Common Patent Class, Assignee, Application Year Among Follow-on and Comparison Patents			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (& not pledge)	0.152***	0.160***	-0.006***	-0.006***
	[0.011]	[0.008]	[0.001]	[0.001]
Citing IBM Pledge (& not patent)	**-0.147*****	**-0.132*****	**-0.001**	**-0.003**
	[0.034]	[0.033]	[0.003]	[0.003]
Citing IBM Patent & Pledge	0.042	0.100*	0.007**	0.003
	[0.056]	[0.057]	[0.003]	[0.004]
Citing IBM Pledge (& not patent) × Pledge Quality	**0.155*****	**0.147*****	**0.003***	**0.003****
	[0.025]	[0.026]	[0.002]	[0.002]
Citing IBM Patent & Pledge × Pledge Quality	0.020	-0.007	-0.005***	-0.004**
	[0.032]	[0.036]	[0.002]	[0.002]
Controls	Yes	Yes	Yes	Yes
Assignee-Patent Class Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	575,120	544,199	575,120	544,199
R-squared	0.2169	0.2152	0.9045	0.9023
Panel C: Compustat Firms				
Dependent Variable	Log Forward Citations		Distance to IBM Patents	
Sample	Common Patent Class, Assignee, Application Year Among Follow-on and Comparison Patents			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (& not pledge)	0.116***	0.127***	-0.005***	-0.004***
	[0.013]	[0.011]	[0.001]	[0.001]
Citing IBM Pledge (& not patent)	**-0.196*****	**-0.161*****	**0.003**	**0.001**
	[0.045]	[0.045]	[0.004]	[0.005]
Citing IBM Patent & Pledge	-0.005	0.088	0.007**	0.001
	[0.071]	[0.089]	[0.003]	[0.005]
Citing IBM Pledge (& not patent) × Pledge Quality	**0.182*****	**0.165*****	**0.001**	**0.002**
	[0.030]	[0.033]	[0.002]	[0.003]
Citing IBM Patent & Pledge × Pledge Quality	0.027	-0.029	-0.005***	-0.003
	[0.045]	[0.054]	[0.002]	[0.003]
Controls	Yes	Yes	Yes	Yes
Additional Compustat Controls	Yes	Yes	Yes	Yes
Assignee-Patent Class Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	286,814	256,834	286,814	256,834
R-squared	0.2441	0.2442	0.9184	0.9156

Table A2 Notes— See detailed notes for Table 3. The constant term is included in all estimations but not reported in the table. Pledge Quality is measured as the number of times an invention pledge is cited. Standard errors are clustered at firm level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A3. Additional results using originality and generality as the dependent variable: Estimates for the differences among follow-on patents and comparison patents

Column	1	2	3	4
Sample	At Least One Common Inventor Among Follow-on and Comparison Patents			
Dependent Variable	Originality		Generality	
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (and not disclosure)	0.017***	0.017***	0.012***	0.013***
	[0.0030]	[0.0036]	[0.0028]	[0.0031]
Citing IBM Disclosure (and not patent)	**0.032*** **	**0.033*** **	**-0.004**	**-0.007**
	[0.0090]	[0.0122]	[0.0107]	[0.0143]
Citing IBM Patent & Disclosure	-0.015***	-0.015**	0.002	-0.006
	[0.0061]	[0.0141]	[0.0067]	[0.0147]
Controls, Compustat Controls, Assignee-IPC Fixed Effects, Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	479,485	415,110	479,485	415,110
R-squared	0.6476	0.6532	0.4157	0.4225
Dependent Variable	Originality		Generality	
Sample	Common IPC Class, Assignee, Application Year Among Follow-on and Comparison Patents			
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (and not disclosure)	0.025***	0.025***	0.027***	0.029***
	[0.0029]	[0.0032]	[0.0033]	[0.0032]
Citing IBM Disclosure (and not patent)	**0.034*** **	**0.033*** **	**-0.006**	**-0.006**
	[0.0078]	[0.0098]	[0.0098]	[0.0124]
Citing IBM Patent & Disclosure	-0.005	-0.009	0.005	-0.006
	[0.0057]	[0.0105]	[0.0069]	[0.0110]
Controls, Compustat Controls, Assignee-IPC Fixed Effects, Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	195,345	172,765	195,345	172,765
R-squared	0.4518	0.4562	0.1355	0.1386

Table A3 Notes— The table presents OLS estimates of originality (Columns 1 and 2) and generality (Columns 3 and 4) for patents applied between 1976 and 2006. Controls included but not reported in Panel A are: log(number of backward citations); log(number of patent classes in backward citations); log(number of non-patent citations); log(number of non-patent citations in the patent's backward citations); log(number of inventors). Additional Compustat controls included in Panel B are: whether the firm reports R&D; log(firm age); log(number of employees); three year moving average of log sales; three year moving average of change in operating income; stock of firm patents; three year moving average of R&D intensity; three year moving average of capital intensity. The constant term is included in all estimations but not reported in the table. Standard errors are clustered at firm level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Additional results using number of claims as the dependent variable: Estimates for the differences among follow-on patents and comparison patents

Column	1	2	3	4
Panel A: All Firms				
Dependent Variable	Log Claims			
Comparison sample	At Least One Common Inventor Among Follow-on and Comparison Patents		Common IPC Class, Assignee, Application Year Among Follow-on and Comparison Patents	
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (and not disclosure)	0.001	0.001	0.008***	0.008***
	[0.0014]	[0.0015]	[0.0020]	[0.0021]
Citing IBM Disclosure (and not patent)	**-0.012***	**-0.014****	**-0.022*****	**-0.023*****
	[0.0064]	[0.0068]	[0.0051]	[0.0055]
Citing IBM Patent & Disclosure	-0.006	-0.002	-0.022***	-0.018**
	[0.0067]	[0.0105]	[0.0057]	[0.0075]
Controls, Assignee-IPC Fixed Effects, Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,574,846	1,492,098	566,127	535,232
R-squared	0.3976	0.4009	0.1506	0.1518
Panel B: Compustat Firms				
Dependent Variable	Log Claims			
Comparison sample	At Least One Common Inventor Follow-on and Comparison Patents		Common IPC Class, Assignee, Application Year Among Follow-on and Comparison Patents	
IBM Patents Included?	Yes	No	Yes	No
Citing IBM Patent (and not disclosure)	-0.004**	-0.004***	0.003	0.001
	[0.0017]	[0.0019]	[0.0025]	[0.0022]
Citing IBM Disclosure (and not patent)	**-0.011**	**-0.016**	**-0.024*****	**-0.024*****
	[0.0103]	[0.0118]	[0.0075]	[0.0090]
Citing IBM Patent & Disclosure	-0.013*	-0.010	-0.033***	-0.038***
	[0.0073]	[0.0157]	[0.0076]	[0.0109]
Controls, Compustat Controls, Assignee-IPC Fixed Effects, Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	669,137	588,234	284,162	254,219
R-squared	0.3924	0.3986	0.1593	0.1640

Table A4 Notes— The table presents OLS estimates of number of claims as the dependent variable for patents applied between 1976 and 2006. Controls included but not reported in Panel A are: log(number of backward citations); log(number of patent classes in backward citations); log(number of non-patent citations); log(number of non-patent citations in the patent's backward citations); log(number of inventors). Additional Compustat controls included in Panel B are: whether the firm reports R&D; log(firm age); log(number of employees); three year moving average of log sales; three year moving average of change in operating income; stock of firm patents; three year moving average of R&D intensity; three year moving average of capital intensity. The constant term is included in all estimations but not reported in the table. Standard errors are clustered at firm level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A5. Additional results using number of claims, originality, and generality: Estimates of IBM's lead time advantage in follow-on invention

Column	1	2	3	4	5	6
Panel A: All Firms						
Dependent Variable	Log Claims		Originality		Generality	
Sample	All First-Generation Patents Citing IBM Invention Pledges					
IBM (=1 if yes, 0 otherwise)	-0.147***	-0.064**	-0.027***	-0.030***	-0.019***	-0.031***
	[0.020]	[0.028]	[0.006]	[0.004]	[0.006]	[0.004]
IBM × After 1989		-0.113***		0.005		0.018***
		[0.035]		[0.008]		[0.006]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Patent Class Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	46,895	46,895	48,153	48,153	48,153	48,153
R-squared	0.163	0.164	0.702	0.702	0.473	0.473
Panel B: Compustat Firms						
Sample	All First-Generation Patents by Compustat firms Citing IBM Invention Pledges					
IBM (=1 if yes, 0 otherwise)	-0.014	0.033	-0.014**	-0.022***	0.003	-0.004
	[0.028]	[0.033]	[0.007]	[0.007]	[0.007]	[0.007]
IBM × After 1989		-0.075**		0.012		0.012
		[0.032]		[0.009]		[0.009]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Additional Compustat Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Patent Class Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	15,707	15,707	16,017	16,017	16,017	16,017
R-squared	0.19	0.19	0.685	0.685	0.503	0.503

Table A5 Notes—This table presents OLS estimates on first-generation patents citing IBM invention pledges for patents applied between 1976 and 2006. To examine lead time advantage in follow-on invention, we analyze a dataset of all patents citing IBM invention pledges. The dependent variables in these regressions are respectively: log number of claims in columns (1) and (2), originality in columns (3) and (4), and generality in columns (5) and (6). Controls include a dummy for US firms (if patent assignee type is US firm) and number of invention pledges cited in columns 3, 4, and 5. Compustat controls include R&D; log(firm age); log(number of employees); three year moving average of log sales; three year moving average of change in operating income; stock of firm patents; three year moving average of R&D intensity; three year moving average of capital intensity. Constant is included but not reported in the table in all specifications. Standard errors clustered at the patent class level are reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1

Table A6. Additional results on the correlation between invention pledges and the average quality of patents in the year

Column	1	2
Variables	Originality	
Total number of invention pledges published up to the focal year	0.004***	
	[0.000]	
Number of cited invention pledges published up to the focal year		0.002*
		[0.001]
Number of uncited invention pledges published up to the focal year		0.004***
		[0.000]
Number of inventors	36,021	36,021
Inventor FE	Yes	Yes
Year FE	Yes	Yes
Observations	1,837,071	1,837,071
R-squared	0.104	0.104
Variables	Generality	
Total number of invention pledges published up to the focal year	0.004***	
	[0.000]	
Number of cited invention pledges published up to the focal year		0.002*
		[0.001]
Number of uncited invention pledges published up to the focal year		0.004***
		[0.000]
Number of inventors	36,021	36,021
Inventor FE	Yes	Yes
Year FE	Yes	Yes
Observations	1,837,071	1,837,071
R-squared	0.090	0.090

Table A6 Notes— The table presents OLS estimates of the average quality (measured with originality and generality) of patents at inventor-year level as the dependent variable, based on a balanced yearly panel of inventors for patents applied between 1976 and 2006. In years when the inventor does not have any patents, the dependent variable is set as zero. The constant term is included in all estimations but not reported in the table. Standard errors are clustered at inventor level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Table A7. Association between monthly invention pledge activity and cumulative patent filings and invention pledges published within a year

Column	1	2
Dependent Variable	Number of IBM Invention Pledges Published by an Inventor in the Current Month	
Number of IBM Patents Filed in Current Month	-0.000	
	[0.005]	
Number of IBM Patents Filed in Current Month × Jan-June		0.005
		[0.007]
Number of IBM Patents Filed in Current Month × July-Dec		-0.006**
		[0.003]
Cumulative Number of IBM Patents Filed Until Previous Month in Current Year	0.002*	
	[0.001]	
Cumulative Number of IBM Patents Filed Until Previous Month in Current Year × Jan-June		-0.000
		[0.002]
Cumulative Number of IBM Patents Filed Until Previous Month in Current Year × July-Dec		0.003***
		[0.001]
Cumulative Number of IBM Invention Pledges Published Until Previous Month in Current Year	0.040***	
	[0.005]	
Cumulative Number of IBM Invention Pledges Published Until Previous Month in Current Year × Jan-June		0.064***
		[0.010]
Cumulative Number of IBM Invention Pledges Published Until Previous Month in Current Year × July-Dec		0.028***
		[0.004]
Inventor Fixed Effects	Yes	Yes
Year-Month Fixed Effects	Yes	Yes
Observations	2,186,893	2,186,893
R-Squared	0.093	0.096

Table A7 Notes— The table presents OLS estimates of the number of IBM invention pledges published by an inventor in the current month as the dependent variable, based on a balanced monthly panel of inventors for patents applied between 1976 and 2006. In months when the inventor does not have any patents, the dependent variable is set as zero. All patent- and pledge-related variables are measured at the inventor level. Jan-June and July-Dec are dummy variables, set to one for the corresponding month during the year and zero otherwise. The constant term is included in all estimations but not reported in the table. Standard errors are clustered at inventor level and reported in parentheses; significance: *** p<0.01, ** p<0.05, * p<0.1.

Appendix B. Additional Material

Table B1. Variations in citing the IBM Technical Disclosure Bulletin in the nonpatent references

IBM Tech. Disclosure Bulletin
IBM Technical Disclosure Bulletin
IBM Tech. Disc. Bulletin
IBM Tech. Disclosure Bull.
IBM Tech. Discl. Bull.
IBM Tech. Disc. Bull.
IBM Technical Disclosure
IBM Tech. Discl. Bul.
IBM Tech. Dis. Bull.
IBM Tech. Discl. Bulletin
IBM Technical disclosure Bulletin
IBM Tech. Bull.
IBM Technical Disclosure Bulletin
IBM Tech. Discl.
IBM TECH. DISC. BULL.
IBM *Disc. Bull.
IBM Tech. Disclosure
IBM Technical Bulletin
IBM Techn. Discl. Bulletin
IBM, Tech. Dis. Bul.
IBM Tech. Disc. Bul.
IBM Tech., Disclosure Bulletin
IBM Tech. Dis. Bul.
IBM Tec. Disc. Bulletin
IBM Tech. Discl Bulletin
IBM Tech. Bul.
IBM Tec. Disclosure Bulletin
IBM Tec. Bull.
IBM Technical Disclosure Bulletin
IBM Tech. Disc. Bull.
IBM ch. Disc. Bull.
IBM Bulletin
IBM Technical Disc. Bulletin
IBM Tech Discl. Bull.
IBM Tech Discl. Bull.
IBM, Technical Disclosure Bulletin
IBM Tech. Bulletin
IBM Tech. Journal
IBM Tech. Disc. B.
IBM, Tech. Bul.
IBM Tech. Disc. Bull
IBM Disclosure Bulletin
IBM Tech. Dis.
IBM, Tech. Disc.
IBM (Tech Bul.)
IBM (Tech. Bul.)
IBM. Tech. Discl. Bull.
IBM Tech. Disclos. Bull.
IBM Tech. Disc.
IBM Technical Discl. Bulletin
IBM, Tech. Discl. Bull.
IBM Tech. Dis. Bull.
IBM TDB
IBM Technical Discosure Bulletin
IBM Tech Disclosure Bulletin
IBM, Tech. Disclosure Bulletin
IBM Tech. Disclosure
IBM Tech. Dicl. Bull.
IBM Tech. Disel. Bull.
IBM Tech. Discl. Bull
IBM - Tech. Bul.
IBM Tech. Disclre. Bulletin
IBM Technical Disc. Bull.
IBM T.D.B.
IBMTDB
IBM-TDB
IBM Tec* Dis* Bul
IBM, TDB
IBM Te* Di* Bul
IBM??TDB
IBM/TDB
IBM Technical * Bulletin
IBM - Tech. Dis. Bul.
IBM TECHNICAL DISCLOSURE BULLETIN
IBM Technical Bulletin
IBM disclosure Bulletin
IBM--Tech. Bul.--
IBM - TDB
IBM-Tech. Discl. Bull.
IBM Discl Bulletin
IBM TDM
IBM technical disclosure
IBM Disc
IBM Research Disclosure
IBM TBS
Disc Bull
IBM Tech D B
IBM Technical Bu*
IBM T D B
IBM Technical Bull
IBM Tech Bul
IBM Disclosure Bull
IBM Tech Sisc Bullet
Technical Disclosure Bulletin
Disclosure Bull
IBM Tech Disclre Blln
IBM Tech disclosure bulletin
Tech Discl Bull
IBM Technical disclosure bulletin
IBM Tech Dislre Blltn
IBM Bulletin Technical Disclosure
IBM Tech Disclosure
IBM Tech Dsclre Bulltn
IBM Tech Disclre Buttn
IBM Technical Dis
IBM Tech Dscl
IBM Tech Dis

Table B2. First-hand (verbatim) descriptions of IBM invention pledges in first-generation patents citing them

We compiled a list of nearly 5,595 citations to IBM invention pledges in granted US patents in their description sections. The file contains patent number and three sentences from patent description: one sentence preceding the sentence in which IBM Technical Disclosure Bulletin is mentioned, the matching sentence itself, and the sentence following the matching sentence. The sentences often depict the nature of invention and its relation to the patent. We make the file containing these descriptions available anonymously using the link below:

https://figshare.com/s/a1691b25bebb0688fd8a

We also provide a motivating example. In February 1979, IBM's 4,199,767 patent disclosed an improved nozzle valve for IBM 6640 ink jet printers, one of the first ink jet printers used in offices. Since black ink was sprayed out of a single nozzle, broken into drops, and each drop was deflected by electromagnetic fields to form individual letters or symbols, the functioning of nozzle was pivotal to print quality. The invention in the '767 patent prevents ink drops from forming on the machine (the problem is described as dribbling) when a printer starts or shuts down. The '767 patent cites four references to the IBM Technical Disclosure Bulletin inventions. All the four invention pledges, credited to different IBM inventors, offer different solutions to the same problem: the 1973 pledge involves an electromagnetic valve; the May 1976 pledge involves a cover for multiple nozzles while the December one relates to a quick cutoff of the ink stream internally; and the 1977 pledge involves a cap. The '767 patent notes "None of the references cited uses a valve member which presses against the outlet orifice proper and has, by it pressing engagement against the orifice, a seal effected regardless of pump pressure against the member. In view of the above, it is a principle object of the present invention to provide novel nozzle valve apparatus which may be positioned adjacent the outlet of the nozzle of an inkjet printer and may function with the nozzle orifice for either inhibiting the emission of the stream of ink drops from the nozzle or allowing the stream of ink drops to proceed uninhibited." The above example shows that the nature of inventions pledged openly is similar to those patented, consistent with our results.