Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Early Disclosure of Invention and Reduced Duplication: An Empirical Test

Sonja Lück, Benjamin Balsmeier, Florian Seliger, Lee Fleming

informs
http://pubsonline.informs.org/journal/mnsc
MANAGEMENT SCIENCE
Articles in Advance, pp. 1–9
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Early Disclosure of Invention and Reduced Duplication: An Empirical Test

Sonja Lück,[a] Benjamin Balsmeier,[b,c] Florian Seliger,[c] Lee Fleming[d]

[a] University of Paderborn, 33098 Paderborn, Germany; [b] University of Luxembourg, 4365 Esch-sur-Alzette, Luxembourg;
[c] ETH Zurich, 8092 Zurich, Switzerland; [d] University of California, Berkeley, Berkeley, California 94720

Contact: sonja.lueck@uni-paderborn.de, http://orcid.org/0000-0003-0380-1965 (SL); benjamin.balsmeier@uni.lu, http://orcid.org/0000-0002-8806-1427 (BB); seliger@kof.ethz.ch, http://orcid.org/0000-0002-6277-8235 (FS); lfleming@berkeley.edu, http://orcid.org/0000-0003-1363-4478 (LF)

Received: July 25, 2019
Revised: September 30, 2019
Accepted: September 30, 2019
Published Online in Articles in Advance: March 13, 2020

https://doi.org/10.1287/mnsc.2019.3521

Abstract. Much work on innovation strategy assumes or theorizes that competition in innovation elicits duplication of research and that disclosure decreases such duplication. We validate this empirically using the American Inventors Protection Act (AIPA), three complementary identification strategies, and a new measure of blocked future patent applications. We show that AIPA—intended to reduce duplication, through default disclosure of patent applications 18 months after filing—reduced duplication in the U.S. and European patent systems. The blocking measure provides a clear and micro measure of technological competition that can be aggregated to facilitate the empirical investigation of innovation, firm strategy, and the positive and negative externalities of patenting.

History: Accepted by Joshua Gans, business strategy.
Funding: This work was supported by the National Science Foundation [Grants 1735650 and 1360228] and the Coleman Fung Institute for Engineering Leadership.
Supplemental Material: The online appendix is available at https://doi.org/10.1287/mnsc.2019.3521.

Keywords: AIPA • duplicative research • knowledge disclosure • patents

Information asymmetries cause inventors to face strategic tradeoffs between protecting and commercializing their ideas (Anton and Yao 1994, Gans et al. 2008). Disclosure facilitates commercialization, yet at the same time exposes ideas to expropriation. Policy makers prefer disclosure, in order to decrease duplicated research across an economy, and offer the inventor the opportunity to patent; in return for a temporary monopoly, the inventor makes enough details public, such that other inventors can either avoid duplicating the effort and/or build upon the idea more easily. The inventor calculates that disclosure and hopefully protected access to the market for ideas will prove more valuable than a trade secret. Society benefits from less duplication of research and faster and more differentiated follow-on research and commercialization, following publication of the patent.

Following this logic, the American Inventors Protection Act (AIPA) of 2000 was intended to encourage faster disclosure of inventions; aligning the U.S. Patent and Trademark Office (USPTO) with the rest of the world, AIPA stipulated that, by default, all applications would be published 18 months after first filing, rather than at issuance. Here, we analyze newly available administrative data (both American and European) that enables a direct measure of duplicated research effort, by observing future applications that are rejected or "blocked" because a patent examiner declares the application to be obvious or not novel, relative to the original and explicitly identified blocking patent.

This new measure of blockings enables validation of an untested assumption in innovation strategy research and the patent-racing literature in particular. Many theoretical models proceed from the assumption that—due to informational asymmetries—competition in innovation inevitably elicits duplicated research and development investment (Loury 1979, Dasgupta and Stiglitz 1980, Scotchmer 1991, Roin 2005, Thompson and Kuhn 2017). Perhaps the most famous anecdote used to motivate this literature is Elisha Gray's loss, by two hours, of his patent race for the telephone to Alexander Graham Bell. Although competition may elicit greater effort by inventors, the lack of disclosure can waste societal resources in duplication; secrets are kept longer, rivals waste resources reinventing the metaphorical wheel, and the societal benefits of technological improvement are delayed.

Inventors in a race, like Bell and Gray, are often confronted with a prisoner's dilemma type of problem. They cannot coordinate and/or credibly commit themselves toward early disclosure of all their efforts, which could make both parties (and third parties) better off. If only one party discloses, the other will often see no benefit in disclosing anymore, so no disclosure happens in the first place. The market may thus fail to provide the welfare-maximizing disclosure system, giving reason for the policy maker to step in.

Disclosure regulation intends to solve this dilemma and balance the interest of inventors and society (Hall and Harhoff 2012, Williams 2017). Empirical work has validated many of the assumptions of this intent. Early disclosure increases (more citations by future patents), accelerates (faster appearance of future citations), and improves (higher difference relative to prior art) follow-on invention (e.g., Hegde et al. 2019, Baruffaldi and Simeth 2018, and Graham and Hegde 2015). In contrast, mandated secrecy delays citations, at least temporarily (Gross 2019). Patents disclosed for standards settings are more likely to be litigated, although causality remains unclear (Simcoe 2005). Cosine distance measures of technological class proximity illustrate how early disclosure makes future "close" patents less similar and "distant" patents more similar, relative to the disclosed patent (Hegde et al. 2019), illustrating how inventors move on more quickly from already-claimed territory and incorporate ideas faster into a new recombinant search. The opening of patent libraries increases geographically proximal patenting activity (Furman et al. 2018). Reducing informational asymmetries through patent disclosure should increase the efficiency of the market for ideas (Gans and Stern 2003), as evidenced by slower licensing agreements after patent grant delays (Gans et al. 2008), more and faster licensing after early disclosure of patent applications (Hegde and Luo 2018), increased venture capital interest (Mohammadi and Khashabi 2017), and positive reputation and network effects (Muller and Pénin 2006).

It remains to be established, however, whether earlier disclosure actually reduces duplicated research, as predicted by theory and hoped for by AIPA policy makers. The assumed mechanism is that at least some inventors and/or their lawyers will read or become aware of and use the disclosed applications to avoid competition and overlap with already claimed technologies (for an explicit model of patent racing, see Thompson and Kuhn 2017). To establish that early disclosure does indeed reduce duplication, we use newly available data and the American Inventors Protection Act as a quasinatural experiment. This enables measurement of how earlier patent disclosure through AIPA had the intended impact of reduced duplication, as measured by fewer blockings of post-AIPA patent-application claims.

The contribution of the paper is twofold. First, we add to the disclosure literature by establishing a first-order, but empirically unaddressed, assumption that earlier knowledge disclosure should lead to reduced duplication. We demonstrate this with three complementary estimations that return consistent results, including a regression discontinuity design, twins matching between U.S. and European patents, and a difference in differences estimation. Second, we apply the new measure of blockings and illustrate its usefulness for empirical research in innovation, strategy and competition, and knowledge diffusion. The count of blockings provides a more nuanced measure of the positive and negative externalities of invention, thus complementing more widely used measures such as future prior art citations (Trajtenberg 1991), knowledge spillovers to close or distant technologies (Jaffe 1986), or a focal firm's stock-price reaction to a patent publication (Kogan et. al. 2018).

Measuring Blocked USPTO Patent Applications

The newly available USPTO Office Action Dataset provides the measure of blocked future inventions by a U.S. patent, based on office actions issued by examiners to applicants during the patent-examination process.[1] An "office action" generally discloses the grounds for a rejection, the claims affected, and the pertinent prior art. The USPTO released 4.4 million office actions between March 19, 2008, and July 11, 2017, sent to the applicants of 2.2 million unique patent applications (Lu et al. 2017), including action types "102" and "103" as stipulated by section 35 of the U.S. Code. Essentially, the patent examiner expresses doubts on novelty (102 action) or obviousness (103 action) regarding the patent application and refers to explicitly identified patents as a basis for this decision.[2]

We defined the number of blockings by a focal patent three different ways: (1) the number of 102 and 103 office actions generated by subsequent patent applications that refer to the focal patent as the blocking patent; (2) the number of subsequent patent applications that generated at least one office action referring to the focal patent; and (3) the number of subsequent patent applications that generated at least one office action referring to the focal patent and that were not granted eventually. Presented results are based on (1) and are very similar if alternatives (2) or (3) are chosen (please see Online Appendix A, Tables A2, A3, A6, and A7).

The U.S. blocking data cover only office actions from March 2008 onwards, causing left-censorship, and hence provide very few blocked patent applications that were filed right before and after AIPA was enacted (all patents that were either granted or abandoned before March 2008 and have been blocked remain unobservable). To avoid bias, we use and find consistent results with European Patent Office (EPO) data—which are available continuously since 1977.

Measuring Blocked EPO Patent Applications

To model the number of blocked inventions at the EPO by a focal U.S. patent, we exploit two unique (and unavailable in the U.S. data) features of the uncensored European data. EPO standard examination practice is to determine all relevant prior art and explicitly classify each citation according to its type of relevance. The resulting examiner's search report contains on average fewer citations than a typical USPTO patent, reflecting precise rules of selectivity and justification in prior art citation. Two categories, labeled X and Y by the EPO, can be described as "blocking" citations (please see Online Appendix G for more details and an example of an EPO search report).[3]

Analogous to the U.S. blockings, we define the number of EPO blockings as the number of X and Y citations referring to the focal U.S. patent as the blocking patent (results are robust to counting X and Y citations separately). Presented results are based on the sum of X and Y citations. Analogous to the variety of U.S. patent measures, we also estimated the number of EPO patent applications, of which at least one claim was blocked, and the number of eventually abandoned patent applications, and found very similar results.

The European blocking data and examination conventions also allow calculation of the proportion of X and Y citations out of all citations to a given U.S. patent as a dependent variable. The proportional measure should remain unbiased, even if European examiners changed their citation behavior, assuming that they changed it consistently across types of citations—for example, if they redirected citations to applications, rather than granted patents, following the passage of AIPA. Any change at all is also less likely in the European data, as EPO examiners are highly constrained in their criteria for citation and must explicitly provide the justification for each citation.[4] We also increased comparability by calculating similar time windows and periods for the U.S. and European data.[5] The research required the integration of many data sources; please see Online Appendix F for details on data sources.

Three Complementary Identification Strategies

Aligning the United States with the rest of the world, AIPA stipulated that, by default, all patents would be published 18 months after first filing of the application instead of at issuance (inventors could still choose to keep their application unpublished until grant, if they filed only in the United States and did not seek foreign protection). This law went into effect on November 29, 2000. We pursue three complementary identification strategies that exploit this regulatory change; a (1) regression discontinuity design (RDD), (2) twins study (TW), and (3) difference in differences (DiD) estimation, each of which have caveats but also address different threats to a causal interpretation. With consistent findings across the three approaches, we present results and descriptive data of our preferred RDD method first and provide a somewhat shorter overview of the latter two approaches, plus a placebo test. Consistently, we find that earlier disclosure induced by AIPA reduces the number of duplicated claims in future applications, by 2.9%–14.3%.

Regression Discontinuity Design

We model a RDD (Davis 2008, Lee and Lemieux 2010) that exploits the discontinuous jump to disclosure for U.S. patents filed on or after November 29, 2000, which did not seek parallel foreign protection (hereafter "U.S. only"). Two arguments motivate the RDD. First, the hypothesized change in duplicated future patent claims is observable for patents filed immediately before and after AIPA.[6] Restricting the analysis to a small time window limits potentially confounding influences from law, policy, or economic changes concurrent to AIPA—for example, the bust of the dot-com bubble. Second, we find little evidence of avoidance strategies by inventors—for example, filing as many patents as possible before or after the regime shift (the number of patent applications in the weeks before and after AIPA remained relatively stable) or a shift in patenting with or without parallel foreign protection (the fraction of applications with parallel foreign protection in weeks before and after AIPA remained stable). Other patent characteristics also remained stable across windows of varying length, lessening the concern that inventors or their firms modified how they wrote patents. For nuances, descriptive statistics, and analysis of potential strategic behavior of few firms, please see below and Online Appendices A and E.

To further minimize potential confounds and avoid adding endogenous variables to the regressions, we estimated models with all available data as well as a matched data set. For the balanced sample, we matched each after-AIPA patent to a corresponding patent before AIPA, using coarsened exact matching (CEM), to minimize differences in observable patent characteristics. We match on backward cites to capture potential differences in novelty, six National Bureau of Economic Research (NBER) technology classes that capture differences in technological popularity, number of independent claims, dependent claims per independent claim, word count of the first claim, average number of words per claim, whether an attorney was involved, and patent pendency (time from application to grant date) to capture differences in patent scope and writing (please see Online Appendix A, Table A5 for no significant differences in observables after matching and no loss in common support). Although it is difficult to completely rule out concerns of strategic shifts in patenting behavior, these analyses should lessen these concerns (the twins design addresses such concerns of strategic maneuvering explicitly).[7]

Our baseline specification is

$$\log(\textit{blocked patents} + 1)_i = \beta_1 \cdot \textit{Post AIPA}_i + f(\textit{Days to AIPA})_i + \varepsilon_i, \quad (1)$$

where *blocked patents*$_i$ refers either to the number of 102 and 103 office actions referencing patent i, the number of times an EPO examiner references patent i with an X or Y in her search report, or the fraction of X and Y cites out of all references that patent i receives from EPO search reports, respectively. *Post AIPA*$_i$ is an indicator indicating whether patent i was filed on or after November 29, 2000, and *Days to AIPA*$_i$ is the assignment variable—that is, the difference in days between the filing date of patent i and November 29, 2000—and ε_i is the error term. Under the assumption that patents filed right before and after AIPA only differ in their filing date and disclosure, and f is correctly specified, β_1 will capture the causal influence of AIPA on the number of duplicated claims of future patent applications.

Alternatively, we estimate the same model as (1), but add technology fixed effects (six NBER classes), δ_i, technology class-specific trends, and firm fixed-effects (γ_i) that control for secular trends in technology popularity and unobserved time-invariant heterogeneity across firms. Thus

$$\log(\textit{blocked patents} + 1)_i = \textit{PostAIPA}_i + f(\textit{Days to AIPA})_i + \delta_i + \delta_i \cdot \textit{Days to AIPA}_i + \gamma_i + \varepsilon_i. \quad (2)$$

We allow for differing linear slopes f before and after AIPA (2), add NBER technology class fixed effects (3), technology class-specific trends (4), firm fixed effects (5), and again allow for nonlinear slopes of f before and after AIPA (6).[8] In Table 1, panel (1) estimates U.S.-only patents applied for within 1 month before and after AIPA, and panel (2) introduces a matched sample. Table 1 presents U.S. blockings, EPO blockings, and the fraction of X and Y cites from EPO search reports. Figure 1 illustrates the effects graphically. The year-window subgraphs (right-hand side) support the assumption of linear slopes. See Online Appendix A for corresponding results of yearly models, as well as estimations of Poisson, ordinary least squares (OLS) with levels, and linear probability models.

The RDD results suggest that AIPA reduced the number of blocked U.S. patent claims between 9.9% (Table 1, column (6), second model, matched data) and 13.3% (column (5), first model, raw data).[9] The effect appears consistent, but smaller, for European patents, with claims blocked by U.S. patents reduced by 5.4% (Table 1, column (6), fourth model, matched data) to 10.2% (Table 1, column (2), third model, raw data). The fractional measure of blocking cites decreases by between 2.7 percentage points (Table 1, column (6), sixth model, matched data) and 5.0 percentage *points* (Table 1, column (1), fifth model, raw data). Additional work in Online Appendix A illustrates how eventually abandoned patents also declined after AIPA, indicating that blocked inventors did more than simply drop some of their claims on eventually granted patents. This result is consistent with, but not captured by, data that consider reduced similarity to future granted patents within the same technology class (Hegde 2019).

Twins Design

Our second identification approach uses a matched twins design by focusing only on U.S. patent applications with a parallel EPO patent application of the same invention (Graham et al. 2003; Hegde et al. 2019). These "patent twins" are identified by a common patent family identifier available in the Patstat database. Identification comes from the difference between citations to the U.S. and EPO applications, with the assumption that they cover the same invention and should therefore disclose the same knowledge.

The major advantage of this approach is that the inclusion of patent family fixed effects should control for unobserved differences across inventions. It should thus be robust to any unobserved changes in patenting behavior, writing, and patent scope and, in particular, if firms and/or their lawyers reacted to early disclosure by writing patents to get around the previously disclosed patent. It relies on the assumption that EPO patent examiners remain unaffected by the publication of a corresponding U.S. patent—that is, EPO examiners keep on citing EPO patent applications and are not more likely to cite the parallel U.S. patent once it is published. These assumptions should not be critical. First, even if the assumption is not met, it should work against the hypothesized negative effect, because the bias would only increase citations to U.S. patents relative to the EPO counterpart. Second, we again rely on the estimation of the fraction of X and Y cites, which should remain unaffected by a potential shift in citation behavior, assuming other citations are similarly affected as the blocking cites.

If U.S. patents with parallel foreign protection are more valuable than U.S.-only patents—or are at least considered so by the applicant—there are likely even more filings in jurisdictions beyond the United States and Europe. This means that knowledge about an invention is more likely to leak out before publication, regardless of AIPA, such that any potential AIPA effect would be attenuated within this group of patents. Taken together, these arguments imply that the twins approach should provide a more conservative estimate.

Table 1. RDD Estimations

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: No. of 102 and 103 blocking actions (in logs)						
Panel 1: Raw data, 1 month						
Post-AIPA	−0.133***	−0.129***	−0.132***	−0.137***	−0.143***	−0.114**
	(0.023)	(0.023)	(0.023)	(0.023)	(0.033)	(0.046)
n	17,382	17,382	17,382	17,382	11,083	11,083
R^2	0.008	0.009	0.021	0.023	0.211	0.212
Panel 2: Matched data, 1 month						
Post-AIPA	−0.121***	−0.121***	−0.115***	−0.118***	−0.114***	−0.104**
	(0.026)	(0.026)	(0.026)	(0.026)	(0.037)	(0.052)
n	13,490	13,490	13,490	13,490	8,280	8,280
R^2	0.008	0.009	0.021	0.023	0.218	0.218
Dependent variable: No. of X and Y blocking actions (in logs)						
Panel 1: Raw data, 1 month						
Post-AIPA	−0.107***	−0.108***	−0.103***	−0.103***	−0.100***	−0.073***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.015)	(0.021)
n	17,382	17,382	17,382	17,382	11,080	11,080
R^2	0.023	0.023	0.030	0.031	0.236	0.237
Panel 2: Matched data, 1 month						
Post-AIPA	−0.100***	−0.100***	−0.098***	−0.098***	−0.081***	−0.055**
	(0.012)	(0.012)	(0.012)	(0.012)	(0.017)	(0.022)
n	13,481	13,481	13,481	13,481	8,277	8,277
R^2	0.020	0.020	0.027	0.027	0.249	0.249
Dependent variable: Fraction of X and Y blocking actions						
Panel 1: Raw data, 1 month						
Post-AIPA	−0.051***	−0.051***	−0.049***	−0.049***	−0.046***	−0.037***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.010)
n	17,382	17,382	17,382	17,382	11,080	11,080
R^2	0.020	0.020	0.025	0.025	0.210	0.210
Panel 2: Matched data, 1 month						
Post-AIPA	−0.045***	−0.045***	−0.045***	−0.046***	−0.035***	−0.027**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.012)
n	13,481	13,481	13,481	13,481	8,277	8,277
R^2	0.017	0.017	0.021	0.021	0.216	0.216
All panels						
RDD time controls	Linear (same slope)	Linear (different slopes)	Linear (different slopes)	Linear (different slopes)	Linear (different slopes)	Quadratic (different slopes)
Technology class fixed effects	No	No	Yes	Yes	Yes	Yes
Technology trends	No	No	No	Yes	Yes	Yes
Firm fixed effects	No	No	No	No	Yes	Yes

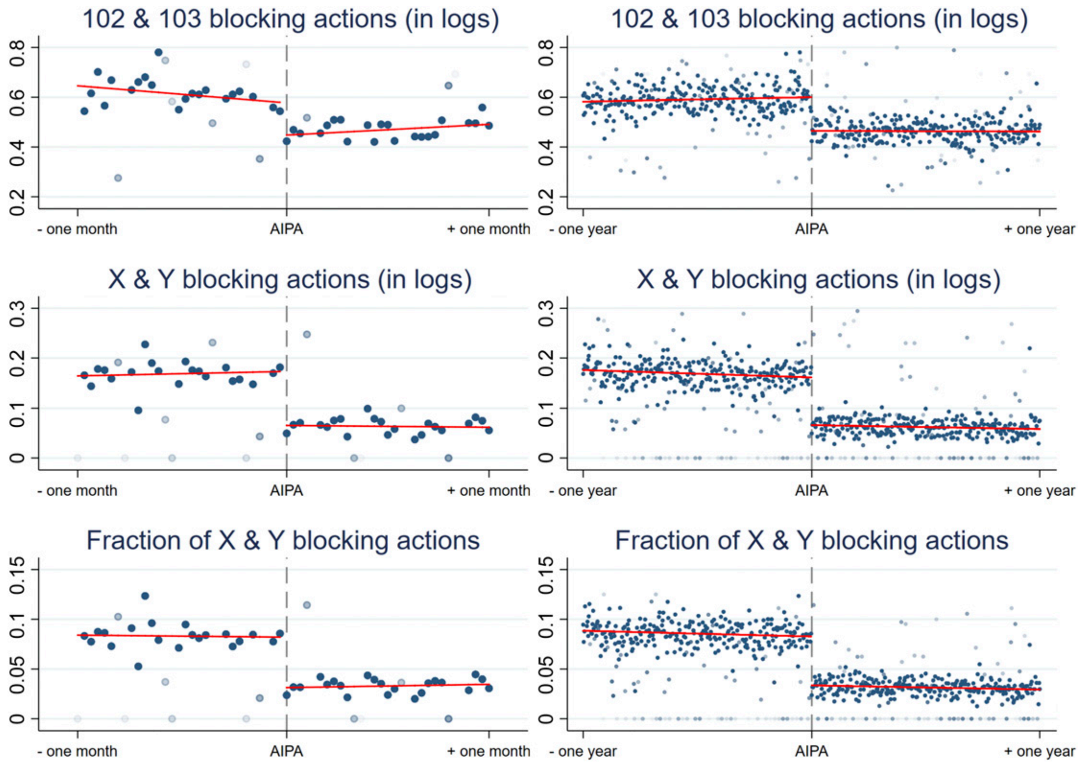
Notes. This table reports results of RDD models, Equations (1) and (2). Technology class fixed effects are from six NBER technology classes, technology trends are linear time trends per technology class using filing days. Raw data are all U.S. patent applications filed between 1 month before and 1 month after November 29, 2000. Post-AIPA is an indicator that equals 1 if the patent was filed on or after November 29, 2000. Matched data are a CEM balanced before and after AIPA sample, based on backward cites, six NBER technology classes, number of independent claims, dependent claims per independent claim, word count of the first claim, average number of words per claim, whether an attorney was involved, and patent pendency to capture differences in patent scope and writing. Online Appendix A, Table A5 shows the descriptive statistics of the matched data. Heteroscedasticity robust standard errors are presented in parentheses.

***, **, and *Statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively.

Note that these models are only possible for European blockings, since 102 and 103 references to non-U.S. patents are not available from the USPTO.

We estimated three different specifications using OLS, starting with patent family fixed effects, which control for unobserved differences across inventions and absorb technology and firm fixed effects. We then add year fixed effects (2) and technology class-specific trends. The number of X and Y cites, as well as the fraction of X and Y cites, serve as dependent variables. Online Appendix B shows how the number of X and Y blockings decreases by 7.2–7.3% and the fraction of

Figure 1. (Color online) RDD Graphs



Notes. These graphs illustrate the discontinuous difference in blocking actions 1 month (year) before and after AIPA became effective on November 29, 2000. Dots represent the average amount of 102 and 103 office actions (log of X and Y cites and fraction of X and Y cites, respectively) referring to patents filed at the USPTO on a given day. AIPA represents November 29, 2000. The lines represent fitted values, and the shading represents the number of patents used to calculate the shown mean. A very light shade therefore stands for 1 to 2 patents.

X and Y cites decreases by 4.9–5.0 percentage points, depending on specification.

Difference in Differences Design

The third identification approach uses a DiD estimation where U.S.-only patents serve as the treated group and all U.S. patents with parallel foreign applications filed before or at the time of the U.S. patent application serve as the control group (Graham and Hegde 2015, Hegde and Luo 2018). The latter group is supposedly unaffected, or at least less affected, by AIPA because the foreign invention is published by the foreign patenting entity 18 months after filing in its jurisdiction. The parallel trend assumption appears valid (see Online Appendix C, Figure C1). The prior analysis as well as graphical inspection suggest, however, that U.S. patents with parallel foreign protection are not unaffected by AIPA. Similar to their U.S.-only counterparts, the number of blocked patents by U.S. patents with foreign publication drops significantly after AIPA. This is most likely because patent publications in a foreign jurisdiction do not receive as much public attention as a USPTO publication, and patent examiners as well as applicants tend to search and cite mainly the U.S. database. The DiD is thus likely to underestimate the impact of AIPA. More precisely, it will estimate how much stronger the impact was for U.S.-only patent applications relative to U.S. patent applications with parallel foreign applications.

Online Appendix C estimates 5 OLS specifications: (1) basic DiD without controls, adding year fixed effects (2), NBER technology-class fixed effects (3), technology class-specific trends (4), and firm fixed effects (5). The impact of AIPA appears less than the estimates based on the RDD and the twins-study approach. We find a reduction in blocked U.S. patent claims (102 and 103) by U.S.-only patents between 4.5% and 3.7%, a reduction in EPO blockings by 2.9%–4.0%, and a reduction in the fraction of X and Y cites by 1.3–1.8 percentage points (all Online Appendix C, Table C1). Given that these estimates reflect the additional effect of AIPA on U.S.-only patents over patents with parallel foreign protection, which was estimated previously to drop by about 10%, the small numbers remain consistent with our estimates based on the RDD.

Placebo Test

Our placebo test draws on the opt-out option included in AIPA. Upon request and without additional costs, inventors could actively opt for pregnant secrecy if they did not seek parallel foreign protection. These patents—about 15% of all U.S.-only applications (~7% in total)—were thus treated by the USPTO as if they would have been submitted under the pre-AIPA regime; thus, we would expect AIPA to have had no effect on this subgroup of patents. Not surprisingly, however, these patents are a highly selected group. Therefore, we (CEM) matched each patent that requested secrecy after AIPA to a pre-AIPA patent that does not significantly differ in terms of backward cites, technology class, number of independent claims, average number of words per independent claim, number of words in first claim, number of dependent claims to number of independent claims, an attorney dummy, and pendency. Rerunning all previous regressions on this subset of patents reveals insignificant results throughout the specifications (see Online Appendix D).

We find positive effects on future blockings for the unmatched sample (Online Appendix D), most likely reflecting the selection of particularly important patents into secrecy. This may cause a slight upward bias in our previous RDD and DiD regressions (the twins study remains unaffected due to the restriction to patents with parallel foreign protection), where the secrecy patents were always included to get as close as possible to a treatment-effect estimate. Because the number of secrecy patents is rather low in total, however, we only find slightly more negative coefficients that do not significantly differ from previous regressions, which exclude the secrecy patents, or adding an indicator for applications that requested secrecy (results available upon request). As an additional placebo test, we tested whether inventors and firms block themselves and found the rates to be 0.2% and 0.49%, relative to 11% for all patents (Online Appendix D).

Informal Estimates of Upper and Lower Bounds of Impact

It remains difficult and speculative to estimate any efficiency gains that might have resulted from AIPA. However, even if a lower number of blocked patents does not reduce duplication in performing research, the results might imply a more efficient application process that avoids processing unsuccessful claims in the first place. In speculating about these bounds, we use RDD estimates of 9.9% and 5.4% for U.S. and European patents, rounded to 10% and 5%, respectively. The USPTO reports an absolute amount of 12.3 million (M) 102 and 103 references in 4.4M office actions between March 2008 and July 2017 (USPTO 2018), and the EPO reports 2.7M X and Y cites over the same period (EPO 2017; often, one cite corresponds to one office action in the EPO). This implies 1.37M avoided references and 0.49M avoided actions in the U.S. system and 0.15M avoided cites in the European system (10% and 5% less for the 9-year period, for the United States and EPO, respectively). Keep in mind that the EPO grants fewer patents each year than the USPTO.

To provide a lower bound on increased efficiency, one could assume that research and development (R&D) spending and allocation remained unaffected by AIPA (i.e., the prior amount of duplicated research continues unchanged). In that case, the gains would accrue only from decreased time and expenses for inventors, their lawyers, and examiners. Given the time to craft a claim and possibly respond to a rejection (or retain an expensive lawyer to do so) on the part of the inventor, and on the part of the examiner, to research, find a basis for rejection, and write up and send a rejection, each rejected claim probably takes multiple—at least tens—of professional hours. This implies millions of hours of saved effort over the time period for highly paid innovation professionals (e.g., at least 0.49M actions $\times$ 10 hours/action). At the high end, if patenting entities redirected the savings of 10% or 5% into new avenues of research, multiplied with an estimate of North American and European R&D of $668 billion in 2010,[10] this implies yearly savings in the billions of dollars. The actual savings is surely lower and probably incorporates saved time and nonduplicative R&D.

Although duplicated patent claims decreased after AIPA, we do not observe a drop in R&D investments or in patents. This may not be surprising, given prior results that future inventors seem to build more frequently and quickly on earlier disclosed knowledge and diverge more significantly from prior research (Hegde et al. 2019). Although we do not observe the direction of innovation, ideally, resources are not just saved, but spent differently and more efficiently.

Conclusion

Information asymmetries and the nature of competition for innovation confront inventors with a strategic dilemma (Anton and Yao 1994, Gans et al. 2008) and policy makers with an opportunity. Delaying disclosure decreases the inventor's exposure to expropriation, even though early disclosure could reduce wasteful duplication and facilitate commercialization. Our results suggest that AIPA helped to mitigate this conundrum. The results confirm a longstanding presumption in models that rely on and predict duplication of efforts in patent races.

The empirical quantification of a first-order effect adds to our understanding of the patent system's costs and benefits and its disclosure function in particular (e.g., Roin 2005, Thompson and Kuhn 2017, Williams 2017, and Sampat and Williams 2019).

The costs of creative destruction have also been difficult to measure completely, and, although much theory has dwelled on the negative externalities of innovation, less work has successfully modeled these ideas empirically (e.g., Bloom et al. 2013 and Kogan et al. 2018). For example, research suggests that positive spillovers from innovation dominate business-stealing effects, but where the latter actually happen and what drives them remain less clear.

In the course of investigating possible selection bias (see Online Appendix E), we also uncovered what appears to be strategic behavior on the part of a few firms. Nintendo, for example, submitted 29 patents the day before AIPA, out of 33 total for the entire 2 months before and after (these patents became salient when we found that they had an average of 201 backward citations). Looking more systematically, we found that 4.5% of firms that applied for patents over the month before and after AIPA made their maximum applications the day before AIPA. Removal of this 1 day of data did not change results, suggesting that this limited strategic behavior was not intended to influence blockings. The more interesting strategic question would consider a firm's choice of secrecy versus disclosure, but is beyond the scope of this work.

Although the measures differ, our results remain broadly consistent with prior work. Using a cosine similarity vector based on assignment across 7,154 seven-digit International Patent Classification (IPC) classes, for populations within four-digit IPC classes, Hegde et al. (2019) demonstrate that patents in the 5th percentile of the similarity distributions after AIPA are less similar and that patents in the 95th percentile are more similar. Put another way, more similar patents become less similar, and less similar patents become more similar. They interpret this as evidence that close technological competitors are using disclosure to move further away from the disclosed work and that distant competitors or noncompetitors are using disclosure to incorporate novel ideas. In contrast, our work relies on explicit identification of the link between individual blocked and blocking patents. This provides advantages, including easy identification of the inventor and assignee (this can be exploited in future research), taking account of all abandoned patents (the prior work used a private data set), and avoiding reliance upon technology classifications and similarity measures.

The possible applications of the blocking measure go beyond measuring negative externalities. Empirical strategy research on innovation currently lacks a direct measure of technological competition; citations measure precedence and remain a tenuous measure of (wanted or unwanted) knowledge diffusion; proximity (either technology class or lexically based) measures of similarity are often too coarse-grained to identify specific patent races; and financial measures are only now being developed that can estimate the negative impact of a patent on a competitor's stock prices (developing the idea published in Kogan et al. 2018). Blockings provide the best-to-date insight as to when inventors (and their firms) collide in technology space. They offer a clear picture of who innovated first and who tried to subsequently follow. This almost micro picture of competition between inventors at the level of individual patents can then be aggregated to illuminate competition between firms, regions, industries, and countries. The measure can be tabulated to illustrate temporal changes as a function of managerial strategy, policy change, court rulings, and technological breakthroughs. We hope that future strategy work can exploit this new empirical opportunity.

Acknowledgments

The authors thank Guan Cheng Li for invaluable research assistance; Stu Graham for detailed and very helpful advice on an earlier related research project; and Bronwyn Hall, Gaetan de Rassenfosse, Christian Helmers, Hong Lou, Deepak Hegde, Olav Sorenson, and Dennis Yao for very helpful comments and suggestions. Errors and omissions remain the authors'.

Endnotes

[1] USPTO Office Action Research Dataset: https://www.uspto.gov/learning-and-resources/electronic-data-products/office-action-research-dataset-patents.

[2] The USPTO provides the following definitions:- 102 "not novel": A claimed invention may be rejected under 35 U.S.C. 102 when the invention is anticipated (or is "not novel") over a disclosure that is available as prior art. Source: https://www.uspto.gov/web/offices/pac/mpep/s2131.html.- 103 "obviousness": A patent for a claimed invention may not be obtained, notwithstanding that the claimed invention is not identically disclosed as set forth in section 102, if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains. Patentability shall not be negated by the manner in which the invention was made. Source: https://www.uspto.gov/web/offices/pac/mpep/s2141.html.

[3] The EPO defines an X citation as: "where a document is such that when taken alone, a claimed invention cannot be considered novel or cannot be considered to involve an inventive step," and a Y citation as, "applicable where a document is such that a claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other documents of the same category, such combination being obvious to a person skilled in the art." Source: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/b_x_9_2_1.htm.

[4] For full list, see: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/b_x_9_2_1.htm.

[5] To maximize comparability over time, we count blocking cites to *granted* patents that could always, before and after AIPA, be observed. To the extent that the publication of pregrant documents changed the citations to granted patents, this change should be fully

absorbed by our fractional measure or cancelled out in the DiD estimation.

[6] Note that the immediate change in blockings estimated and illustrated does not reflect an immediate change in applications; rather, the applications filed after Nov. 29, 2000, become less likely to block *future* patent applications.

[7] One might argue that patent lawyers may have started to write patents in a different way after AIPA. Given that observable characteristics of patents did not change, this would seem less likely.

[8] Modelling f more flexibly with higher-order polynomials reveals similar results. The same is true if we disaggregate technology classes to NBER subclass levels (36) and allow technology class-specific trends to vary before and after AIPA.

[9] Because of the log specification, the economic magnitude of the effect is calculated as exp(β)-1.

[10] The 2010 spending in 2010 dollars; see http://uis.unesco.org/apps/visualisations/research-and-development-spending/, keeping in mind that not all of this investment aims for a patent or would be influenced by AIPA.

References

Anton J, Yao D (1994) Expropriation and inventions: Appropriable rents in the absence of property rights. *Amer. Econom. Rev.* 84(1): 190–209.

Baruffaldi S, Simeth M (2018) Patents and knowledge diffusion: The effect of early disclosure. Working paper, Max Planck Institute for Innovation and Competition, Munich, Germany.

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

Davis LW (2008) The effect of driving restrictions on air quality in Mexico City. *J. Political Econom* 116(1):38–81.

Dasgupta P, Stiglitz J (1980) Industrial structure and the speed of R&D. *Bell J. Econom.* 11(1):1–28.

EPO (2017) 9.2. Categories of documents (X, Y, P, A, D, etc.). *Guidelines for Examination.* Accessed November 5, 2019, https://www.epo.org/law-practice/legal-texts/html/guidelines/e/b_x_9_2.htm.

Furman J, Nagler M, Watzinger M (2018) Disclosure and subsequent innovation: Evidence from the patent depository library program. NBER Working Paper 24660, National Bureau of Economic Research, Cambridge, MA.

Gans J, Stern S (2003) The product market and the market for "ideas": Commercialization strategies for technology entrepreneurs. *Res. Policy*. 32(2):333–350.

Gans J, Hsu D, Stern S (2008) The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. *Management Sci.* 54(5):982–997.

Graham S, Hegde D (2015) Disclosing patents' secrets. *Science* 347(6219):236–237.

Graham S, Hall B, Harhoff D, Mowery D (2003) Patent quality control: A comparison of U.S. patent re-examinations and European Patent oppositions. *Patents in the Knowledge-Based Economy* (National Academies Press, Washington, DC), 74–119.

Gross D (2019) The consequences of invention secrecy: Evidence from the USPTO patent secrecy program in World War II. NBER Working Paper 25545, National Bureau of Economic Research, Cambridge, MA.

Hall BH, Harhoff D (2012) Recent research on the economics of patents. *Annual Rev. Econom.* 4:541–565.

Hegde D, Luo H (2018) Patent publication and the market for ideas. *Management Sci.* 64(2):652–672.

Hegde D, Herkenhoff K, Zhu C (2019) Patent publication and innovation. Working paper, New York University, New York.

Jaffe AB (1986) Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits and market value. *Amer. Econom. Rev.* 76(5):984–1001.

Kogan L, Papanikolaou D, Schmidt LDW, Song J (2018) Technological innovation and the distribution of labor income growth. Working paper, Massachusetts Institute of Technology, Cambridge.

Lee DS, Lemieux T (2010) Regression discontinuity designs in economics. *J. Econom. Literature* 48(2):281–355.

Loury GC (1979) Market structure and innovation. *Quart. J. Econom.* 93(3):395–410.

Lu Q, Myers A, Beliveau S (2017) USPTO patent prosecution research data: Unlocking office action traits. USPTO Economic Working Paper 2017-10, U.S. Patent and Trademark Office, Alexandria, VA.

Mohammadi A, Khashabi P (2017) Embracing the sharks: The impact of information exposure on the likelihood and quality of CVC investments. *Acad. Management Proc.*, ePub ahead of print October 30, https://doi.org/10.5465/AMBPP.2017.252.

Muller P, Pénin J (2006) Why do firms disclose knowledge and how does it matter? *J. Evolutionary Econom.* 16(1-2):85–108.

Roin BN (2005) The disclosure function of the patent system (or lack thereof). *Harvard Law Rev.* 118:2007–2028.

Sampat B, Williams HL (2019) How do patents affect follow-on innovation? Evidence from the human genome. *Amer. Econom. Rev.* 109(1):203–236.

Scotchmer S (1991) Standing on the shoulders of giants: Cumulative research and the patent law. *J. Econom. Perspect.* 5(1):29–41.

Simcoe T (2005) Explaining the increase in intellectual property disclosure. Working paper, Boston University, Boston.

Trajtenberg M (1991) A penny for your quotes: Patent citations and the value of innovations. *RAND J. Econom.* 21(1): 172–187.

Thompson N, Kuhn J (2017) Does winning a patent race lead to more follow-on innovation? Working paper, Massachusetts Institute of Technology, Cambridge.

USPTO (2018) USPTO Office Action Research Dataset (U.S. Patent and Trademark Office, Alexandria, VA). Accessed November 5, 2019, https://www.uspto.gov/learning-and-resources/electronic-data-products/office-action-research-dataset-patents.

Williams H (2017) How do patents affect research investments? *Annu. Rev. Econom.* 9:441–469.