# Online Appendix

"Patent Publication and Innovation" by Deepak Hegde, Kyle Herkenhoff, and Chenqi Zhu

February 11, 2022

# A Proofs

## A.1 Pre-AIPA.

**Assumption 1:** Suppose $V^p(\cdot)$ is bounded above, strictly inreasing and concave in each of its arguments, and is such that$V^p(z,0,Z_0) = 0$, and $\lim_{\Delta\to 0} V_2^p(z,\Delta,Z_0) = \infty$. Suppose $c(\cdot)$ is strictly convex, $c(0) = 0$, and it has a finite first derivative at the origin, $c_1(0) < \infty$. We further assume $c(\cdot)$ and $V^p(\cdot)$ are twice continuously differentiable in all arguments.

**Lemma 1.** Under assumption 1, (i) the patent scope decision is finite and unique, (ii) if public knowledge and scope are complements, then patent scope is increasing in public knowledge, (iii) if public knowledge and scope are substitutes, then patent scope is a decreasing function of public knowledge.

*Proof.* Since $c(\cdot)$ is strictly convex and increasing, it is unbounded. Define $\Delta^H > 0$ as point where $c(\Delta^H) = qV^p(z,\Delta^H,Z_0)$. By assumption $qV^p(z,0,Z_0) = c(0)$ and $qV^p(z,\epsilon,Z_0) > c(\epsilon)$ for sufficiently small $\epsilon$ since $c(\cdot)$ has a finite first derivative. Thus $qV^p(z,\Delta,Z_0) - c(\Delta)$ attains its maximum in the interior of the compact interval $[0,\Delta^H]$. Since $qV^p(z,\Delta,Z_0) - c(\Delta)$ is strictly concave, any local maximum is a global maximum. The interior solution is thus defined by the first order condition, $qV_2^p(z,\Delta,Z_0) = c_1(\Delta)$. Define $\Delta(z,Z_0)$ as the solution to this problem.

The implicit function theorem provides hypotheses under which $\Delta(z,Z_0)$ is a locally differentiable function of $z$ and $Z_0$. We assume those hypotheses are met, and we characterize the properties of $\Delta(z,Z_0)$ below. Taking the derivative with respect to $Z_0$ yields

$$\begin{aligned}
qV_{22}^p(z,\Delta,Z_0)\Delta_2(z,Z_0) + qV_{23}^p(z,\Delta,Z_0) &= c_{11}(\Delta)\Delta_2(z,Z_0) \\
qV_{23}^p(z,\Delta,Z_0) &= \left[c_{11}(\Delta) - qV_{22}^p(z,\Delta,Z_0)\right]\Delta_2(z,Z_0).
\end{aligned}$$

Rearranging,

$$\Delta_2(z,Z_0) = \frac{qV_{23}^p(z,\Delta,Z_0)}{[c_{11}(\Delta) - qV_{22}^p(z,\Delta,Z_0)]}.$$

By assumption 1, $c_{11}(\Delta) > 0$ and $V_{22}^p(z,\Delta,Z_0) < 0$. Thus if scope and public knowledge are complements, $V_{23}^p(z,\Delta,Z_0) > 0$, then scope is an increasing function of public knowledge $\Delta_2(z,Z_0) > 0$.

Likewise, we can characterize how $\Delta(z,Z_0)$ varies with $z$. Taking the derivative with respect to $z$ yields:

$$qV_{22}^p(z,\Delta,Z_0)\Delta_1(z,Z_0) + qV_{12}^p(z,\Delta,Z_0) = c_{11}(\Delta)\Delta_1(z,Z_0)$$

$$\Delta_1(z,Z_0) = \frac{qV_{12}^p(z,\Delta,Z_0)}{[c_{11}(\Delta) - qV_{22}^p(z,\Delta,Z_0)]}.$$

Thus scope is an increasing function of idea quality if they are complements, $V_{12}^p(z,\Delta,Z_0) > 0$.

■

**Corollary 1**: Under assumption 1, the expected payoff to patenting is monotone increasing in $z$, and $Z_0$.

*Proof:* Let $\Delta(z, Z_0)$ denote the optimal investment in scope for an idea of quality $z$. We know $\Pi(z, \Delta(z, Z_0), Z_0) < \Pi(z', \Delta(z, Z_0), Z_0))$ for $z' > z$ since $V^p$ is strictly increasing in its inputs. Since the firm can do at least as well as maintaining a constant $\Delta(z, Z_0)$, it must be the case that reoptimization weakly improves payoffs, $\Pi(z, \Delta(z, Z_0), Z_0)) < \Pi(z', \Delta(z, Z_0), Z_0)) \leq \Pi(z', \Delta(z', Z_0), Z_0)$. Thus the expected payoff from patenting is monotone in $z$. A similar argument holds for $Z_0$. ■

**Assumption 2:** there exist $\underline{z}$ and $\overline{z}$ such that $V^c > \Pi(\underline{z}, \Delta(\underline{z}, Z_0), Z_0)$ and $\Pi(\overline{z}, \Delta(\overline{z}, Z_0), Z_0) > V^c$.

**Lemma 2:** Under Assumptions 1 and 2, a unique interior patenting threshold $z^p$ exists and is monotone decreasing in public knowledge $Z_0$.

**Proof:** Since $\Pi(\cdot)$ is strictly increasing in $z$, $\Pi(\cdot)$ must cross $V^c$ only once in the interval $[\underline{z}, \overline{z}]$. Taking the derivative with respect to $Z_0$, we have that monotonicity of the objective in its inputs implies $z_1^p$ is a decreasing function of public knowledge:

$$\frac{d}{dZ_0}\left[\Pi(z^p(Z_0), \Delta(z^p(Z_0), Z_0), Z_0)\right] = qV_1^p z_1^p + qV_2^p[\Delta_1 z_1^p + \Delta_2] + qV_3^p - c_1[\Delta_1 z_1^p + \Delta_2]$$

$$\frac{d}{dZ_0}\left[V^c\right] = 0.$$

Differentiating the threshold condition for patenting we have

$$\begin{aligned} 0 &= qV_1^p z_1^p + qV_2^p[\Delta_1 z_1^p + \Delta_2] + qV_3^p - c_1[\Delta_1 z_1^p + \Delta_2] \\ &= qV_1^p z_1^p + [\underbrace{qV_2^p - c_1}_{\text{0 from foc for }\Delta}]\Delta_1 z_1^p + qV_3^p + \underbrace{[qV_2^p - c_1]}_{\text{0 from foc for }\Delta}\Delta_2]. \end{aligned}$$

Using optimality of $\Delta$, and noting that the profitability from patenting is strictly increasing in its arguments $V_3^p > 0$ and $V_1^p > 0$, the patenting threshold is decreasing since

$$\begin{aligned} -qV_3^p &= qV_1^p z_1^p \\ \frac{-V_3^p}{V_1^p} &= z_1^p \\ z_1^p &= \frac{-V_3^p}{V_1^p} < 0. \qquad ■ \end{aligned}$$

## A.2 Post-AIPA.

For the remainder of the theoretic characterization, we maintain assumptions 1 and 2 as hypotheses.

**Lemma 4:** The expected payoffs of patenting are unambiguously greater post-AIPA.

*Proof.* Using the optimized value pre-AIPA and the pre-AIPA policy function for scope, it is clear that since $q < 1$, the post-AIPA value of patenting increases.

$$\begin{aligned} qV^p(z,\Delta(z,Z_0),Z_0)+(1-q)V^c - c(\Delta(z,Z_0)) &< qV^p(z,\Delta(z,Z_0),Z_1)+(1-q)V^c - qc(\Delta(z,Z_0)) \\ &= q\left[V^p(z,\Delta(z,Z_0),Z_1) - c(\Delta(z,Z_0))\right] + (1-q)V^c \\ &\leq \max_{\Delta} q\left[V^p(z,\Delta,Z_1) - c(\Delta)\right] + (1-q)V^c. \end{aligned}$$

■

We define the post-AIPA expected payoffs as $\Pi^A(z,\Delta(z,Z_1),Z_1) = \max_\Delta q\left[V^p(z,\Delta,Z_1) - c(\Delta)\right] + (1-q)V^c$.

**Lemma 5:** The patenting threshold declines post-AIPA, $z^{p,A} < z^p$.

*Proof.* The pre-AIPA patenting threshold was defined by the point at which $V^c = \Pi^A(z^p,\Delta(z^p,Z_0),Z_0)$. We know that from Lemma 4, $V^c = \Pi(z^p,\Delta(z^p,Z_0),Z_0) < \Pi^A(z^p,\Delta(z^p,Z_1),Z_1)$ (note, the threshold is being held constant). Rewriting this,

$$V^c = qV^p(z^p,\Delta(z^p,Z_0),Z_0)+(1-q)V^c - c(\Delta(z^p,Z_0)) < \max_{\Delta} q\left[V^p(z^p,\Delta,Z_1) - c(\Delta)\right] + (1-q)V^c = \Pi^A(z^p,\Delta(z^p,Z_0),Z_0).$$

Since $\Pi(z^p,\Delta(z^p,Z_1),Z_1)$ is increasing in $z^p$(Corollary 1) and since it exceeds $V^c$, we must have that $z^{p,A} < z^p$ and the patenting threshold must decrease. ■

**Lemma 6:** If $q \approx 1$ and $V^p_{23}(z,\Delta,Z_1) < 0$, then scope will decline post-AIPA. For any $q < 1$, if $V^p_{23}(z,\Delta,Z_1) > 0$, scope will increase.

*Proof.* The pre-AIPA first order condition for scope is given by

$$qV^p_2(z,\Delta,Z_0) = c_1(\Delta).$$

Suppose $q \to 1$ and $V^p_{23} < 0$. Then $Z_1 > Z_0$ implies $\lim_{q\to 1} qV^p_2(z,\Delta,Z_0) > V^p_2(z,\Delta,Z_1)$. Since $V^p_{22} < 0$, i.e. $V^p$ is concave in all arguments, and $c_{11} > 0$, then to restore equality, $\Delta$ must decline. Thus $\Delta^A(z,Z_1) < \Delta(z,Z_0)$ in this limiting case.

Now suppose $V^p_{23} > 0$, then, unambiguously,

$$V^p_2(z,\Delta,sZ_1) > qV^p_2(z,\Delta,Z_0).$$

Since $V^p_{22} < 0$, i.e. $V^p$ is concave in all arguments, and $c_{11} > 0$, then to restore equality, $\Delta$ must increase. Thus $\Delta^A(z,Z_1) > \Delta(z,Z_0)$ when $V^p_{23} > 0$. ■

## A.3 Citations

Let $s$ denote search effort of the existing stock of knowledge. We assume greater values of search effort $s$ imply better draws of ideas. This theory of a citation captures the notion of 'standing on the shoulder of giants.' In this section, we assume ideas are drawn from $z \sim F(z,s)$. We assume

$F$ is Pareto, with lower bound on the support of $s$. For technical reasons, we restrict functional forms and payoffs so that $z^p >> s$. We assume that citations and search effort are determined prior to drawing the idea quality. Search effort comes at a convex, twice differentiable cost of $\mu(s)$. Thus the pre-AIPA value at $t = 0$ is given by,

$$V(Z_0) = \left[\int_{z^p}^{\infty} \{\max_{\Delta} qV^p(z, \Delta, Z_0) + (1-q)V^c - c(\Delta)\} f(z, s)dz + V^c F(z^p, s)\right] - \mu(s)$$

**Lemma A1.** Under the assumption that (i)$F(x, s)$ is Pareto with tail index $\alpha$, lower bound $s$, (ii) $\mu(s) = \frac{1}{\beta+1}s^{\beta+1}$, and (iii) $\beta > \alpha - 1$, then $s$ in an increasing function of expected payoffs $\Pi(z, \Delta, Z_0)$.

*Proof.* What makes the problem tractable is that $z_p$, $\Delta$, and $Z_0$ are not functions of $s$.[35] We rewrite the value pre-AIPA using only the pdf of $z$ as this will make subsequent comparative statics simpler:

$$V(Z_0) = \left[\int_{z^p}^{\infty} \{\max_{\Delta} qV^p(z, \Delta, Z_0) + (1-q)V^c - c(\Delta)\} f(z, s)dz + \int_{-\infty}^{z^p} V^c f(z, s)dz\right] - \mu(s).$$

Taking first order conditions with respect to $s$ and applying Leibniz's rule yields,

$$\left[\int_{z^p}^{\infty} \{\max_{\Delta} qV^p(z, \Delta, Z_0) + (1-q)V^c - c(\Delta)\} f_2(z, s)dz + \int_{-\infty}^{z^p} V^c f_2(z, s)dz\right] = \mu_1(s).$$

Since $F(x, s)$ is Pareto, we have $F(x, s) = 1 - \left(\frac{s}{x}\right)^{\alpha}$ for $x \geq s$, zero otherwise. We assume $\alpha > 1$. Then the pdf of $F(x, s)$ is given by $f(x, s) = \frac{\alpha s^{\alpha}}{x^{\alpha+1}}$ if $x \geq s$, zero otherwise. Taking the derivative of the cdf and pdf with respect to $s$ yields $F_2(x, s) = -\alpha\frac{s^{\alpha-1}}{x^{\alpha}}$ and $f_2(x, s) = \frac{\alpha^2 s^{\alpha-1}}{x^{\alpha+1}}$. We can then solve for $s$

$$\left[\int_{z^p}^{\infty} \{\max_{\Delta} qV^p(z, \Delta, Z_0) + (1-q)V^c - c(\Delta)\} \frac{\alpha^2 s^{\alpha-1}}{z^{\alpha+1}}dz + \int_{-\infty}^{z^p} V^c \frac{\alpha^2 s^{\alpha-1}}{z^{\alpha+1}}dz\right] = s^{\beta}$$

$$\left[\int_{z^p}^{\infty} \Pi(z, \Delta, Z_0)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{-\infty}^{z^p} V^c \frac{\alpha^2}{z^{\alpha+1}}dz\right] = s^{\beta-\alpha+1}.$$

Therefore, as long as $\beta > \alpha - 1$, $s$ in an increasing function of expected payoffs $\Pi(z, \Delta, Z_0)$. ■

**Lemma A2:** Under the assumption that (i)$F(x, s)$ is Pareto with tail index $\alpha$, lower bound $s$, (ii) $\mu(s) = \frac{1}{\beta+1}s^{\beta+1}$, and (iii) $\beta > \alpha - 1$, then $s$ is increasing post-AIPA.

*Proof.*

The first order condition for $s$ pre-AIPA is

[35] In particular, $z^p$ is independent of $s$ and pinned down by comparing $\max_{\Delta} qV^p(z, \Delta, Z_0) + (1-q)V^c - c(\Delta)$ and $V^c$, conditional on the idea draw.

$$\left[\int_{z^p}^{\infty} \Pi(z,\Delta,Z_0)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{-\infty}^{z^p} V^c \frac{\alpha^2}{z^{\alpha+1}}dz\right] = s^{\beta-\alpha+1}.$$

The first order condition for $s$ post-AIPA is

$$\left[\int_{z^{p,A}}^{\infty} \Pi^A(z,\Delta^A,Z_1)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{-\infty}^{z^{p,A}} V^c \frac{\alpha^2}{z^{\alpha+1}}dz\right] = \left(s^A\right)^{\beta-\alpha+1}.$$

Thus

$$\begin{aligned}
\left(s^A\right)^{\beta-\alpha+1} - s^{\beta-\alpha+1} &= \left[\int_{z^{p,A}}^{\infty} \Pi^A(z,\Delta^A,Z_1)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{-\infty}^{z^{p,A}} V^c \frac{\alpha^2}{z^{\alpha+1}}dz\right] - \left[\int_{z^p}^{\infty} \Pi(z,\Delta,Z_0)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{-\infty}^{z^p} V^c \frac{\alpha^2}{z^{\alpha+1}}dz\right] \\
&= \left[\int_{z^p}^{\infty} \Pi^A(z,\Delta^A,Z_1)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{z^{p,A}}^{z^p} \Pi^A(z,\Delta^A,Z_1)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{-\infty}^{z^{p,A}} V^c \frac{\alpha^2}{z^{\alpha+1}}dz\right] \\
&\quad - \left[\int_{z^p}^{\infty} \Pi(z,\Delta,Z_0)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{z^{p,A}}^{z^p} V^c \frac{\alpha^2}{z^{\alpha+1}}dz + \int_{-\infty}^{z^{p,A}} V^c \frac{\alpha^2}{z^{\alpha+1}}dz\right] \\
&= \left[\int_{z^p}^{\infty} \left(\Pi^A(z,\Delta^A,Z_1) - \Pi(z,\Delta,Z_0)\right)\frac{\alpha^2}{z^{\alpha+1}}dz + \int_{z^{p,A}}^{z^p} \left(\Pi^A(z,\Delta^A,Z_1) - V^c\right)\frac{\alpha^2}{z^{\alpha+1}}dz\right].
\end{aligned}$$

By Lemma 5, and the definition of $z^p$, it must be the case that the right-hand side is positive. Thus, search intensity increases post-AIPA. ■

We interpret our model as predicting that citations – in particular forward citations – increase post-AIPA.

**Discussion of forward vs backward citations.** Consider repeating our model economy so that there are two cohorts of firms. The first cohort ($t = 1$) is identical to the post-AIPA environment in the text. The second cohort ($t = 2$) is also identical to the post-AIPA environment, except they take the prior cohort's public knowledge $Z_1$ as a state variable, and their patenting produces new public knowledge $Z_2$. The value functions for both cohorts are given by,

$$\begin{aligned}
t = 1: \quad & V^A(Z_0) = \int_{z^{p,A}}^{\infty} \{q\left[\max_{\Delta} V^p(z,\Delta,Z_1) - c(\Delta)\right] + (1-q)V^c\}dF(z) + V^c F(z^{p,A}), \quad && s.t. \quad Z_1 = Z_0 + q\int_{z^{p,A}}^{\infty} z dF(z). \\
t = 2: \quad & V^A(Z_1) = \int_{z^{p,A}}^{\infty} \{q\left[\max_{\Delta} V^p(z,\Delta,Z_2) - c(\Delta)\right] + (1-q)V^c\}dF(z) + V^c F(z^{p,A}), \quad && s.t. \quad Z_2 = Z_1 + q\int_{z^{p,A}}^{\infty} z dF(z).
\end{aligned}$$

The model definition of a citation is conceptually closer to a backward citation. But backward citations from the future cohort ($t = 2$ patents) are forward citations to the prior cohort ($t = 1$ patents). Payoffs to patenting increase in all cohorts ($t = 1, 2$) since AIPA is a permanent news shock to all future patenting cohorts. Then $s$ has the interpretation of forward citations to the prior cohort (equivalently forward citations per patent), and $s$ unambiguously increases. Thus, we argue that forward citations rise post-AIPA.

## A.4 Endogenous Duplicates

Suppose the probability of duplication increases as scope increases (.e.g, greater scope, claims, and breadth imply a greater probability of overlap with another patent). We therefore assume that $q(\Delta)$ is a well-behaved decreasing function of scope. If $q(\Delta)$ is chosen so that the payoff

$W(z, \Delta, sZ_0) = q(\Delta)V^p(z, \Delta, sZ_0) + (1 - q(\Delta))V^c$ remains strictly increasing, concave, and continues to statisfy the Inada conditions, then our prior comparative statics in Lemma 1 apply. Is it reasonable to restrict payoffs to be strictly concave in scope? These restriction guarantee positive and finite scope, which is observed in practice.

**Lemma A4.** Assume sufficient conditions on $q(\cdot)$ are met to guaranteee an unique interior solution. Under local convexity of $q$ (e.g. $q_{11} < 0$ around the optimum) as well as substitutability between $\Delta$ and $Z_0$ (e.g. $V^p_{23} < 0$), scope unambigously declines when public knowledge expands. If $V^p_{23} > 0$, then the impact of public knowledge on scope is ambiguous.

*Proof.* The first order condition is given by

$$q(\Delta)V^p_2(z, \Delta, Z_0) + q_1(\Delta)V^p(z, \Delta, Z_0) - q_1(\Delta)V^c = c_1(\Delta).$$

Differentiating with respect to $Z_0$,

$$\Delta_2 \left[q_1(\Delta)V^p_2(z, \Delta, Z_0) + q(\Delta)V^p_{22}(z, \Delta, Z_0) + q_{11}(\Delta)V^p(z, \Delta, Z_0) - q_{11}(\Delta)V^c + q_1(\Delta)V^p_2(z, \Delta, Z_0) - c_{11}(\Delta)\right] = -\left[q(\Delta)V^p_{23}(z, \Delta, Z_0) + q_1(\Delta)V^p_3(z, \Delta, Z_0)\right].$$

We can then characterize the scope as a function of public knowledge:

$$\Delta_2 = \frac{[q(\Delta)V^p_{23}(z, \Delta, Z_0) + q_1(\Delta)V^p_3(z, \Delta, Z_0)]}{[c_{11}(\Delta) - q_1(\Delta)V^p_2(z, \Delta, Z_0) - q(\Delta)V^p_{22}(z, \Delta, Z_0) - q_{11}(\Delta)[V^p(z, \Delta, Z_0) - V^c] - q_1(\Delta)V^p_2(z, \Delta, Z_0)]}.$$

Under local convexity of $q$ (e.g. $q_{11} < 0$) as well as substitutability between $\Delta$ and $Z_0$ (e.g. $V^p_{23} < 0$), scope unambigously declines when public knowledge expands. If $V^p_{23} > 0$, then the impact of public knowledge on scope is ambiguous. ■

# B Validation of Similarity Measures

To validate our technological similarity measure, we draw on the institutional feature that citations made by the EPO are classified based on the relationship between the cited and citing patents. The most important types of citations are X- and Y-citations, which account for 21% and 20% of the citations, respectively, according to the categorization in PATSTAT. An X- or Y-citation indicates that at least one claim in the citing patent cannot be considered novel or does not involve an inventive step, either taking the cited patent alone or combining it with other cited documents. Thus, X- and Y-citations are particularly relevant to the citing patent. In comparison, other citations merely provide general background information. For example, the most common citations (comprising 49% of all citations) are A-citations, which merely define the general state of the art. Therefore, if our similarity measure captures the fundamental proximity of the patented invention, we should observe a higher 'Sim' (defined in Section 4.1) for X- and Y-citations than for other types of citations.

For further comparison, we also match the cited or the citing patent involved in an X- or Y-citation to a random patent. The matching is based on application quarter, IPC 4-digit code, and grant status. We then construct 'Sim' for the fake citation pairs (comprising of actual cited and matched citing patents or matched cited and actual citing patents) and expect 'Sim' to be even lower.

We start from citation pairs with the cited patents filed from 1998 to 2003 and citing patents from 1998 to 2009. This restriction yields a sample of 141,582 X-citations and 50,445 Y-citations. We label them as "Important Citations". We then keep the citing patent in the important citations constant and construct its similarity to its other types of citations (labeled as "Background Citations") and to a patent matched to the cited patent (labeled as "Fake Citations"). Given that the EPO makes only four backward citations per application, on average, the requirement for the citing patent to have at least one important citation and one background citation filed from 1998 to 2003 reduces the sample substantially. For this reason, we summarize and compare 'Sim' for the whole sample of important citations and the reduced sample, separately. In a similar vein, we also compare 'Sim' across important citations, background citations, and fake citations, while keeping the cited patent fixed.

We report the results in Table B.1. Panel A shows that the average 'Sim' for the entire sample of important citations is 0.594, more than three times larger than that of fake citations. When we focus on the subsample in which the citing patent has at least one important citation and one background citation, we observe a slight decrease in 'Sim' from important citations to background citations, a further decrease from background citations to fake citations. The pattern is similar when we compare 'Sim' while keeping the cited patent fixed, as reported in Panel B. To sum, the evidence here demonstrates that our similarity measure exhibits variations that are consistent with the technological overlap identified by patent examiners at the EPO.

Table B.1: Validation of patent classification-based patent similarity measure

*This table reports the average pairwise similarity between different pairs of citing-cited patents. Important references refer to X- and Y-citations, as designated by the EPO, which indicate that at least one claim in the citing patent cannot be considered novel or does not involve an inventive step. Background references refer to other types of actual citations. Fake references involve either the citing or the cited patent is a matched random patent to the actual one based on application quarter, IPC 4-digit code, and grant status. Panel A (B) compares Similarity ('Sim' as defined in Section 4.1) across important citations, background citations, and fake citations, while keeping the citing (cited) patent fixed using two-tailed t-tests. ***, **, and * stand for statistical significance at the 1%, 5%, and 10% level, respectively.*

| **Panel A: Similarity ('Sim') b/w Important, Background, and Fake References (Keep Citing Patent Fixed)** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | **All Important References** | | | **Require the citing patent in the important reference to have at least one background references filed from 98 to 03** | | | | | |
| | Important | Fake | Dif(1-2) | Important | Background | Dif(4-5) | Fake | Dif(4-7) | Dif(5-7) |
| # | 191,358 | 191,358 | | 32,141 | 32,141 | | 32,141 | | |
| Mean | 0.594 | 0.139 | 0.455*** | 0.615 | 0.592 | 0.024*** | 0.147 | 0.468*** | 0.444*** |
| S.E. | 0.001 | 0.001 | 0.001 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| **Panel B: Similarity ('Sim') b/w Important, Background, and Fake References (Keep Cited Patent Fixed)** | | | | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | **All Important References** | | | **Require the cited patent in the important reference to be cited as background references by patents filed from 98 to 09** | | | | | |
| | Important | Fake | Dif(1-2) | Important | Background | Dif(4-5) | Fake | Dif(4-7) | Dif(5-7) |
| # | 191,490 | 191,490 | | 133,982 | 133,982 | | 133,982 | | |
| Mean | 0.594 | 0.160 | 0.434*** | 0.596 | 0.579 | 0.017*** | 0.162 | 0.434*** | 0.417*** |
| S.E. | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |

# C Event Study Analysis

The main text graphically examines knowledge diffusions and patent characteristics before and after AIPA, as presented in Figures 3-6. Here we pursue a parametric regression approach that removes the pre-trends apparent in the data. The regression model is specified as follows:

$$Outcome_{it} = \alpha_0 + \sum_{\tau \in Post} \beta_\tau \mathbf{I}(Month_t = \tau) + \alpha_1 Month_t + \alpha_2 \mathbf{I}(EarlyGrant_i) + \alpha_3 + TechFE + \epsilon_{it} \tag{2}$$

where $i$ indicates the application filed in calendar month $t$.[36] The dependent variables are the forward citations counts, citation lags, technology similarity, renewal rates, patent originality, claims, and abandonment rates. Following Gross et al. (2020), we include a function of a continuous variable of application month ($f(Month)$), as well as a set of dummy variables indicating each month in the post-AIPA regime (37 dummies, each indicating a month from December 2000 to December 2003). The term $f(Month)$ controls for the pre-trend, while the coefficient on the dummies ($\beta_t$) capture the lagged effect of AIPA at different horizons relative to the pre-trend. In the baseline regressions, $f(Month)$ is a linear function of the calendar month.[37] To quantify AIPA's immediate and longer-term effects, we take the average of the estimated $\beta_t$ over different horizons and compute the corresponding standard errors using the delta method.

Model 2 controls for whether patents are granted before 18 months ($I(EarlyGrant_i)$, since such patents are *de facto* untreated by AIPA. We also add a dummy variable ($I(OptOut_i)$) indicating patents that opt out of the pre-grant publication requirement. USPC technology fixed effects are added to control for time-invariant industry characteristics. IPC and USPC codes do not map one-to-one, and so we use USPC codes in our domestic analysis for more accurate controls of technology class.

Panel A of Table C.1 formalizes the graphical forward citation analysis. The dependent variable is the natural logarithm of one plus the number of forward citations, counted in windows ranging from three to ten years after disclosure. We find that the average coefficient on the post-AIPA time dummies (Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$)), which measures the average post-AIPA deviation to the pre-trend, is significantly positive across all four citation measures. Economically, these estimates suggest that post-AIPA patents receive 3.8%-19% more forward citations, on average, than the predicted citations based on the pre-trend. The pre-trend is captured by the coefficient on $Month_t$, which is significantly negative, consistent with the downward trend evident in Figure 3. The coefficient on $\mathbf{I}(OptOut_i)$ is of similar magnitude as the average AIPA effect (Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$)) but has the opposite sign, suggesting the knowledge diffusion of opt-out

[36] Since AIPA became effective on November 29, 2000, we allocate 2,536 applications filed on November 29-30, 2000 to December 2000 so that each month is classified as either pre- or post-AIPA throughout the paper.

[37] In robustness checks, we add second- or third-order polynomials of *Month* to control for potential non-linear pre-trends and find that our inferences are not changed appreciably.

patents do not effectively change. With the caveat that quality-driven selection into opt-outs may be driving the finding, this lends additional support to the idea that pre-grant patent disclosure under AIPA, rather than macro trends or confounding events, enhances knowledge diffusion.

Panel B of Table C.1 conducts regression analysis on the speed of knowledge diffusion using Model 2. The results suggest that the delay to receive one to seven forward citations decreased by 11.6%-20.1%, relative to the predicted delay based on the pre-trend. The coefficient on $\mathbf{I}(OptOut_i)$ is significantly positive, ranging from 17.5%-27.6%, which more than offsets the average AIPA effect. Again, with the caveat of quality-driven selection, this is consistent with the idea that it is pre-grant patent disclosure, rather than macro trends or confounding events, that hastens knowledge diffusion.

As Figure 4 shows, a linear pre-trend does not fit the pre-AIPA patents' citation lags very well; hence, we add the second- or third-order polynomial of the calendar month and report the results in Appendix Table C.2. The coefficients on both $Month^2$ and $Month^3$ are statistically significant, suggesting that the inclusion of non-linear pre-trends is warranted. Nevertheless, we still observe a statistically significant drop in citation lags in the post-AIPA period, although the economic magnitude varies with the degree of polynomials included. The estimated drop in citation delays is 30.0%-41.4% (11.0%-16.3%) when we model the pre-trend using a quadratic (cubic) function.

Panel C of Table C.1 reports the regression results for technology similarity. We find that technology similarity at the 50th and 75th percentiles increase, suggesting that more distant patents become more similar post-AIPA. In contrast, similarity at the 90th and 95th percentile decreases, suggesting more differentiation among the closest patents post-AIPA and corroborating the graphic evidence in Figure 5.

Table C.3 formalizes the graphical analysis of patent characteristics as plotted in Figure 6. Column 1 shows that renewal rates fell, suggesting that inventors pursued patenting for less-valuable ideas post AIPA. Column 2 confirms the graphic evidence in Figure 6 that inventors patent less-original ideas after AIPA. Columns 3-5 show that the number of total allowed claims or independent claims decreases and the average independent claim length increases (more narrowly delineated scope, per Kuhn and Thompson (2019)), providing consistent evidence showing that inventors pursued smaller scopes in their patents. The last column, Column 6, shows a 5.2% (=0.013/0.247) decrease in the abandonment rate relative to the pre-AIPA period.

## Table C.1: AIPA's effect: event study analysis

*This table reports Ordinary Least Squares (OLS) regression estimates of AIPA's effect on forward citations (Panel A), citation lags (Panel B), and technology similarity (Panel C) using the following specification:*

$$Outcome_{it} = \alpha_0 + \sum_{\tau \in Post} \beta_\tau \mathbf{I}(Month_t = \tau) + \alpha_1 Month_t + \alpha_2 \mathbf{I}(EarlyGrant_i) + \alpha_3 \mathbf{I}(OptOut_i) + TechFE + \epsilon_{it}$$

*where $i$ indicates a patent application filed in month $t$. We include 'Month' (a continuous variable of the calendar application month) and a set of dummy variables indicating each month in the post-AIPA period ($\boldsymbol{I}(Month_t = \tau)$). We include technology class fixed effects (3-digit USPC code). Standard errors are clustered by the application month. We obtain the pre-post difference by taking the mean of the estimates of $\beta_\tau$, and its associated standard errors are computed using the delta method. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively.*

**Panel A: Number of citations**

| | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
|---|---|---|---|---|
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.038*** | 0.113*** | 0.164*** | 0.190*** |
| | (0.006) | (0.007) | (0.007) | (0.007) |
| Month | -0.003*** | -0.005*** | -0.006*** | -0.006*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Early Grant (d) | -0.016*** | -0.034*** | -0.045*** | -0.058*** |
| | (0.002) | (0.003) | (0.003) | (0.003) |
| Opt-Out (d) | -0.071*** | -0.135*** | -0.158*** | -0.164*** |
| | (0.006) | (0.006) | (0.006) | (0.006) |
| USPC FE | Yes | Yes | Yes | Yes |
| Observations | 1,107,656 | 1,107,204 | 1,104,670 | 1,089,865 |
| Adj R-squared | 0.135 | 0.148 | 0.154 | 0.161 |

**Panel B: Citations lags**

| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
|---|---|---|---|---|
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.201 *** | -0.143*** | -0.126*** | -0.116*** |
| | (0.007) | (0.006) | (0.006) | (0.006) |
| Month | 0.001*** | 0.001*** | 0.000** | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Early Grant (d) | -0.317*** | -0.218*** | -0.184*** | -0.169*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Opt-Out (d) | 0.276*** | 0.212*** | 0.189*** | 0.175*** |
| | (0.007) | (0.006) | (0.005) | (0.005) |
| USPC FE | Yes | Yes | Yes | Yes |
| Observations | 990,975 | 772,256 | 604,596 | 482,676 |
| Adj R-squared | 0.103 | 0.126 | 0.130 | 0.130 |

**Panel C: Technology similarity**

| | (1) Similarity 50th Pctile | (2) Similarity 75th Pctile | (3) Similarity 90th Pctile | (4) Similarity 95th Pctile |
|---|---|---|---|---|
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.006*** | 0.006*** | -0.006*** | -0.014*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Month | 0.000*** | 0.000* | -0.000*** | -0.000*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Early Grant (d) | -0.001 | -0.001 | 0.001 | 0.005*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Opt-Out (d) | -0.002 | 0.010*** | 0.038*** | 0.052*** |
| | (0.001) | (0.002) | (0.002) | (0.002) |
| USPC FE | Yes | Yes | Yes | Yes |
| Observations | 1,106,975 | 1,106,975 | 1,106,975 | 1,106,975 |
| Adj R-squared | 0.290 | 0.238 | 0.179 | 0.151 |

Table C.2: Event study analysis of citation lags with non-linear trends

*This table reports the before-and-after analyses of AIPA's effect with non-linear pre-trends. The regressions are specified as follows:*

$$Outcome_{it} = \alpha_0 + \sum_{\tau \in Post} \beta_\tau \mathbf{I}(Month_t = \tau) + \alpha_1 f(Month_t) + \alpha_2 \mathbf{I}(EarlyGrant_i) + \alpha_3 \mathbf{I}(OptOut_i) + TechFE + \epsilon_{it}$$

*where $i$ indicates a patent application filed in month $t$. The specification is the same as Table C.1, except that the linear pre-trend is replaced by non-linear trends specified by second-order polynomials of Month in Panel A and third-order in Panel B. Standard errors are clustered by the application month. We obtain the pre-post difference by taking the mean of the estimates of $\beta_\tau$, and its associated standard errors are computed using the delta method. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively.*

| **Panel A: With second-order polynomials of Month** | | | | |
|---|---|---|---|---|
| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.414*** | -0.330*** | -0.315*** | -0.300*** |
| | (0.025) | (0.022) | (0.022) | (0.023) |
| Month | -0.155*** | -0.135*** | -0.137*** | -0.133*** |
| | (0.018) | (0.015) | (0.015) | (0.016) |
| Month$^2$ | 0.000*** | 0.000*** | 0.000*** | 0.000*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Early Grant (d) | -0.316*** | -0.218*** | -0.183*** | -0.168*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Opt-Out (d) | 0.276*** | 0.212*** | 0.189*** | 0.174*** |
| | (0.007) | (0.006) | (0.005) | (0.005) |
| Observations | 990,975 | 772,256 | 604,596 | 482,676 |
| R-squared | 0.103 | 0.126 | 0.130 | 0.130 |

| **Panel B: With third-order polynomials of Month** | | | | |
|---|---|---|---|---|
| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.163*** | -0.124*** | -0.117*** | -0.110*** |
| | (0.008) | (0.007) | (0.007) | (0.007) |
| Month | -3.339*** | -2.709*** | -2.594*** | -2.471*** |
| | (0.264) | (0.227) | (0.226) | (0.241) |
| Month$^2$ | 0.007*** | 0.006*** | 0.005*** | 0.005*** |
| | (0.001) | (0.000) | (0.000) | (0.000) |
| Month$^3$ | -0.000*** | -0.000*** | -0.000*** | -0.000*** |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Early Grant (d) | -0.316*** | -0.218*** | -0.183*** | -0.168*** |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Opt-Out (d) | 0.276*** | 0.212*** | 0.189*** | 0.174*** |
| | (0.007) | (0.006) | (0.005) | (0.005) |
| Observations | 990,975 | 772,256 | 604,596 | 482,676 |
| R-squared | 0.103 | 0.126 | 0.130 | 0.130 |

Table C.3: Event study analysis of patent characteristics

*This table reports OLS regression estimates of AIPA's effect on renewal rates, patent originality, patent claims, and abandonment rates. The regressions are specified as follows:*

$$Outcome_{it} = \alpha_0 + \sum_{\tau \in Post} \beta_\tau \mathbf{I}(Month_t = \tau) + \alpha_1 Month_t + \alpha_2 \mathbf{I}(EarlyGrant_i) + \alpha_3 \mathbf{I}(OptOut_i) + TechFE + \epsilon_{it}$$

*where $i$ indicates a patent application filed in month $t$. The specification is the same as in Table C.1. Standard errors are clustered by the application month. We obtain the pre-post difference by taking the mean of the estimates of $\beta_\tau$, and its associated standard errors are computed using the delta method. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively.*

| | (1) 3.5-yr Renewal (d) | (2) Originality | (3) Claims | (4) Independent Claims | (5) Independent Claim Word Length | (6) Abandon (d) |
|---|---|---|---|---|---|---|
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.014*** | -0.010*** | -0.665*** | -0.120*** | 8.22*** | -0.013*** |
| | (0.002) | (0.002) | (0.106) | (0.016) | (0.717) | (0.003) |
| Month | 0.001*** | 0.001*** | 0.045*** | 0.003*** | -0.223*** | -0.000*** |
| | (0.000) | (0.000) | (0.003) | (0.000) | (0.019) | (0.000) |
| Early Grant (d) | -0.006*** | -0.049*** | -3.033*** | -0.592*** | 5.018*** | |
| | (0.001) | (0.001) | (0.047) | (0.006) | (0.333) | |
| Opt-Out (d) | -0.005** | 0.015*** | 1.787*** | 0.222*** | 5.012*** | -0.000 |
| | (0.002) | (0.002) | (0.101) | (0.014) | (0.612) | (0.000) |
| USPC FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,107,656 | 1,083,275 | 1,107,656 | 1,100,495 | 1,102,757 | 1,536,346 |
| R-squared | 0.027 | 0.124 | 0.034 | 0.058 | 0.062 | 0.002 |

## Table C.4: Event study analyses by subsamples, part I

*This table conducts pre-post comparisons of US patents in terms of forward citations, citation lags, and technology similarity. We exclude patents that opt out of the 18-month disclosure requirement and split the sample into patents with or without EP equivalents (Panel A and B) or any foreign equivalents (Panel C and D). For brevity, we only report 'Pre-Post-Dif $\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$', which indicates the impact of AIPA relative to the pre-trend. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively.*

| **Panel A. US patents with EP equivalents** | | | | |
|---|---|---|---|---|
| **Forward Citations** | (1)<br>Log 3-Yr. Forward Cites | (2)<br>Log 5-Yr. Forward Cites | (3)<br>Log 7-Yr. Forward Cites | (4)<br>Log 10-Yr. Forward Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.048***<br>(0.010) | 0.125***<br>(0.011) | 0.180***<br>(0.012) | 0.207***<br>(0.012) |
| Observations | 306,299 | 306,192 | 305,635 | 302,882 |
| R-squared | 0.155 | 0.169 | 0.176 | 0.185 |
| **Citation Lags** | (1)<br>Log Months to 1 Cite | (2)<br>Log Months to 3 Cites | (3)<br>Log Months to 5 Cites | (4)<br>Log Months to 7 Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.207***<br>(0.016) | -0.153***<br>(0.015) | -0.134***<br>(0.013) | -0.126***<br>(0.013) |
| Observations | 273,298 | 213,306 | 168,401 | 136,062 |
| R-squared | 0.117 | 0.144 | 0.148 | 0.148 |
| **Technology Similarity** | (1)<br>Similarity 50th Pc-tile | (2)<br>Similarity 75th Pc-tile | (3)<br>Similarity 90th Pc-tile | (4)<br>Similarity 95th Pc-tile |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.003*<br>(0.002) | 0.008***<br>(0.003) | 0.001<br>(0.003) | -0.004<br>(0.004) |
| Observations | 306,266 | 306,266 | 306,266 | 306,266 |
| R-squared | 0.258 | 0.225 | 0.173 | 0.146 |

| **Panel B. US patents without EP equivalents** | | | | |
|---|---|---|---|---|
| **Forward Citations** | (1)<br>Log 3-Yr. Forward Cites | (2)<br>Log 5-Yr. Forward Cites | (3)<br>Log 7-Yr. Forward Cites | (4)<br>Log 10-Yr. Forward Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.033***<br>(0.007) | 0.107***<br>(0.009) | 0.156***<br>(0.010) | 0.181***<br>(0.010) |
| Observations | 752,922 | 752,814 | 752,193 | 747,939 |
| R-squared | 0.133 | 0.146 | 0.152 | 0.156 |
| **Citation Lags** | (1)<br>Log Months to 1 Cite | (2)<br>Log Months to 3 Cites | (3)<br>Log Months to 5 Cites | (4)<br>Log Months to 7 Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.199***<br>(0.012) | -0.139***<br>(0.010) | -0.123***<br>(0.010) | -0.111***<br>(0.010) |
| Observations | 674,357 | 525,764 | 410,741 | 326,668 |
| R-squared | 0.102 | 0.125 | 0.129 | 0.129 |
| **Technology Similarity** | (1)<br>Similarity 50th Pc-tile | (2)<br>Similarity 75th Pc-tile | (3)<br>Similarity 90th Pc-tile | (4)<br>Similarity 95th Pc-tile |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.008***<br>(0.001) | 0.005***<br>(0.002) | -0.009***<br>(0.002) | -0.017***<br>(0.003) |
| Observations | 752,431 | 752,431 | 752,431 | 752,431 |
| R-squared | 0.301 | 0.244 | 0.181 | 0.150 |

## Table C.5: Event study analyses by subsamples, part II

*This table conducts pre-post comparisons of U.S. patents in terms of forward citations, citation lags, and technology similarity. We exclude patents that opt out of the 18-month disclosure requirement and split the sample into patents with or without EP equivalents (Panel A and B) or any foreign equivalents (Panel C and D). For brevity, we only report 'Pre-Post-Dif $\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$', which indicates the impact of AIPA relative to the pre-trend. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively.*

| **Panel C. US patents with foreign equivalents** | | | | |
|---|---|---|---|---|
| **Forward Citations** | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.043*** | 0.119*** | 0.173*** | 0.200*** |
| | (0.009) | (0.011) | (0.011) | (0.012) |
| Observations | 574,481 | 574,345 | 573,602 | 569,789 |
| R-squared | 0.129 | 0.142 | 0.148 | 0.155 |
| **Citation Lags** | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.203*** | -0.142*** | -0.119*** | -0.109*** |
| | (0.014) | (0.013) | (0.012) | (0.011) |
| Observations | 513,663 | 398,487 | 310,833 | 247,529 |
| R-squared | 0.102 | 0.126 | 0.131 | 0.131 |
| **Technology Similarity** | (1) Similarity 50th Pc-tile | (2) Similarity 75th Pc-tile | (3) Similarity 90th Pc-tile | (4) Similarity 95th Pc-tile |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.006*** | 0.008*** | -0.002 | -0.007** |
| | (0.001) | (0.002) | (0.003) | (0.003) |
| Observations | 574,414 | 574,414 | 574,414 | 574,414 |
| R-squared | 0.292 | 0.238 | 0.180 | 0.149 |

| **Panel D. US patents without foreign equivalents** | | | | |
|---|---|---|---|---|
| **Forward Citations** | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.034*** | 0.107*** | 0.153*** | 0.178*** |
| | (0.009) | (0.012) | (0.013) | (0.013) |
| Observations | 484,740 | 484,661 | 484,226 | 481,032 |
| R-squared | 0.151 | 0.165 | 0.170 | 0.174 |
| **Citation Lags** | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.200*** | -0.145*** | -0.135*** | -0.125*** |
| | (0.013) | (0.009) | (0.010) | (0.010) |
| Observations | 433,994 | 340,582 | 268,309 | 215,198 |
| R-squared | 0.110 | 0.134 | 0.138 | 0.138 |
| **Technology Similarity** | (1) Similarity 50th Pc-tile | (2) Similarity 75th Pc-tile | (3) Similarity 90th Pc-tile | (4) Similarity 95th Pc-tile |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.006*** | 0.004** | -0.009*** | -0.019*** |
| | (0.002) | (0.002) | (0.002) | (0.003) |
| Observations | 484,283 | 484,283 | 484,283 | 484,283 |
| R-squared | 0.292 | 0.243 | 0.182 | 0.154 |

# D Twin Analysis Robustness

We conduct several robustness exercises to validate or to strengthen our twin analysis in this appendix.

## D.1 Trend of US-EP Twins

Figure D.1 show that U.S. patentees' propensity to file for foreign or EP parallel applications remained stable throughout the period.

## D.2 Validation of Twin US-EP Patents

Table D.1 provides examples of 10 randomly selected pairs of US-EP twins with links to their original documents on Google Patents. We further compute pairwise similarity between the U.S. and EP patent for each US-EP twin in our sample. Overall, we find that the "twins" are indeed extremely similar, with IPC-7 digit measures yielding perfect correlation (e.g. cosine similarity=1.0).

## D.3 Alternative Regression Specifications

Table D.2 estimates Model (1) in the main text with alternate fixed effects: (1) technology class fixed effects measured by 4-digit IPC codes, and (2) applicant fixed effects. We find very similar results. In particular, applicant fixed effects serve as a powerful test showing that fixed unobserved heterogeneity among applicants is unlikely to drive our results.

Table D.3 estimates Model (1) in the main text with an additional linear trend by one-digit IPC technology class (8 classes in total, represented by A-H, respectively) for both USPTO and EPO patents. Again, we find very similar results to the baseline estimates.

Lastly, we examine and rule out the possibility that the post-AIPA increase in citations to U.S. patents comes from a simple migration of citations from EP patents to their U.S. equivalents after AIPA. This test is reported in Table D.4. The variable 'Common Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$)' captures the change in citations for EP twins (with the subsequent interaction of 'US (d)' and post-AIPA month dummies capturing the additional citations for U.S. patents, labelled as 'AIPA effect ($\frac{1}{37}\sum_{\tau \in Post} \gamma_\tau$)'). 'Common Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$)' is positive, suggesting that European twin patents did not receive a reduction in citations post AIPA. This result corroborates the lack-of-migration-hypothesis as shown in Panel B of Table 4, showing that the driver of increased citations comes from U.S. patents, not a migration of EP citations to U.S. equivalents.

## D.4 Excluding the Year Before AIPA

To deal with potential anticipation effects, Table D.5 excludes the year before AIPA (i.e., from November 29, 1999 to November 28, 2000). We find very similar results to our baseline estimates.

## D.5 Placebo AIPA Date

In Table D.6, we use the pre-AIPA period (January 1998-November 2000) and pretend that the date of AIPA was July 1st, 1999, the middle point of the pre-AIPA period. The results yield much smaller, insignificant coefficients, providing us with confidence that our baseline results are not driven by spurious trends.

## D.6 Placebo Sample: U.S. Twins Granted Before 18 months

In this section, we provide an alternative placebo test. Specifically, to support our knowledge diffusion assumptions, we restrict the sample to the set of twins that had early disclosure in the US via early grant (granted before 18 months). Table D.7 of the Appendix estimates AIPA's effect using the subsample of twins for which the US twin was granted at or within 18 months from application. For this subsample, we see little to no effect of AIPA on forward citations and a slight decrease in citation lags (about 20% of the magnitude we see in the entire twin sample). We view this as convincing evidence that no differential knowledge diffusion is occurring when follow-on investors can learn about U.S. patents early on.

## D.7 Citations to US-JP twins

Figure D.2 plots the monthly average forward citations to US-JP twins. We observe divergent trends between US and JP equivalents, justifying the need to control for country-specific trend used in the analyses discussed in Section 6.1.

Figure D.1: U.S. patentees' propensity to file for foreign or EP parallel applications

*The figure below plots the percent of patents filed at the USPTO that file parallel applications in the EPO or any foreign patent office. All U.S. applications filed from 1998 to 2003 that are eventually granted by mid-2014 are included. Foreign or EP parallel applications, identified from the patent family table from PATSTAT, are required to be filed within 18 months of the application of their U.S. equivalents.*

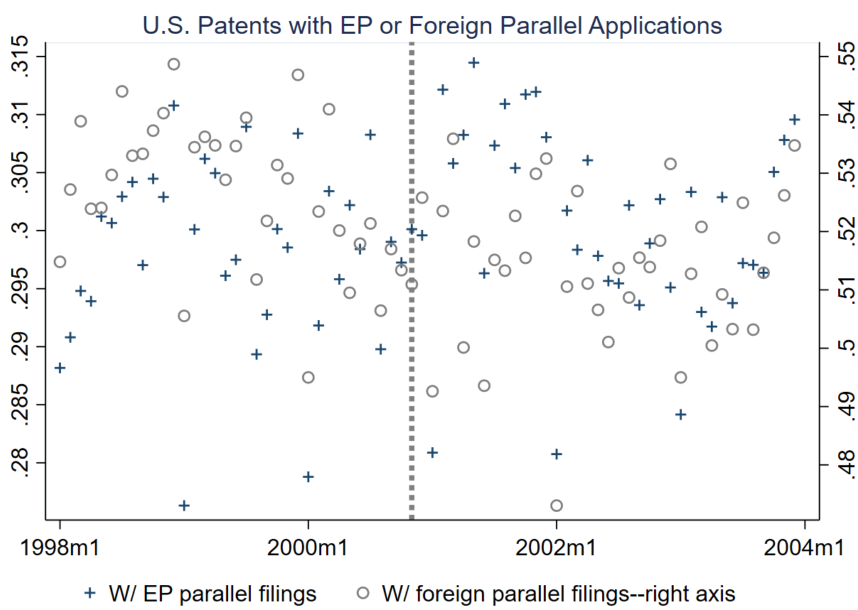


Figure D.2: Citations to US-JP "twins" before and after AIPA

*The figures plot the monthly average number of forward citations (excluding self-citations) to U.S. patents and their equivalent "twins" at the Japanese Patent Office filed during 1998-2003. Forward citations are counted cumulatively 3/5/7/10 years after patent disclosure (i.e., publication date for patents with pre-grant publications or grant date for those without). We normalize the average by its value at the beginning of the sample period. Citations data are obtained from the USPTO and PATSTAT. The vertical line represents AIPA's effective date (November 29, 2000).*

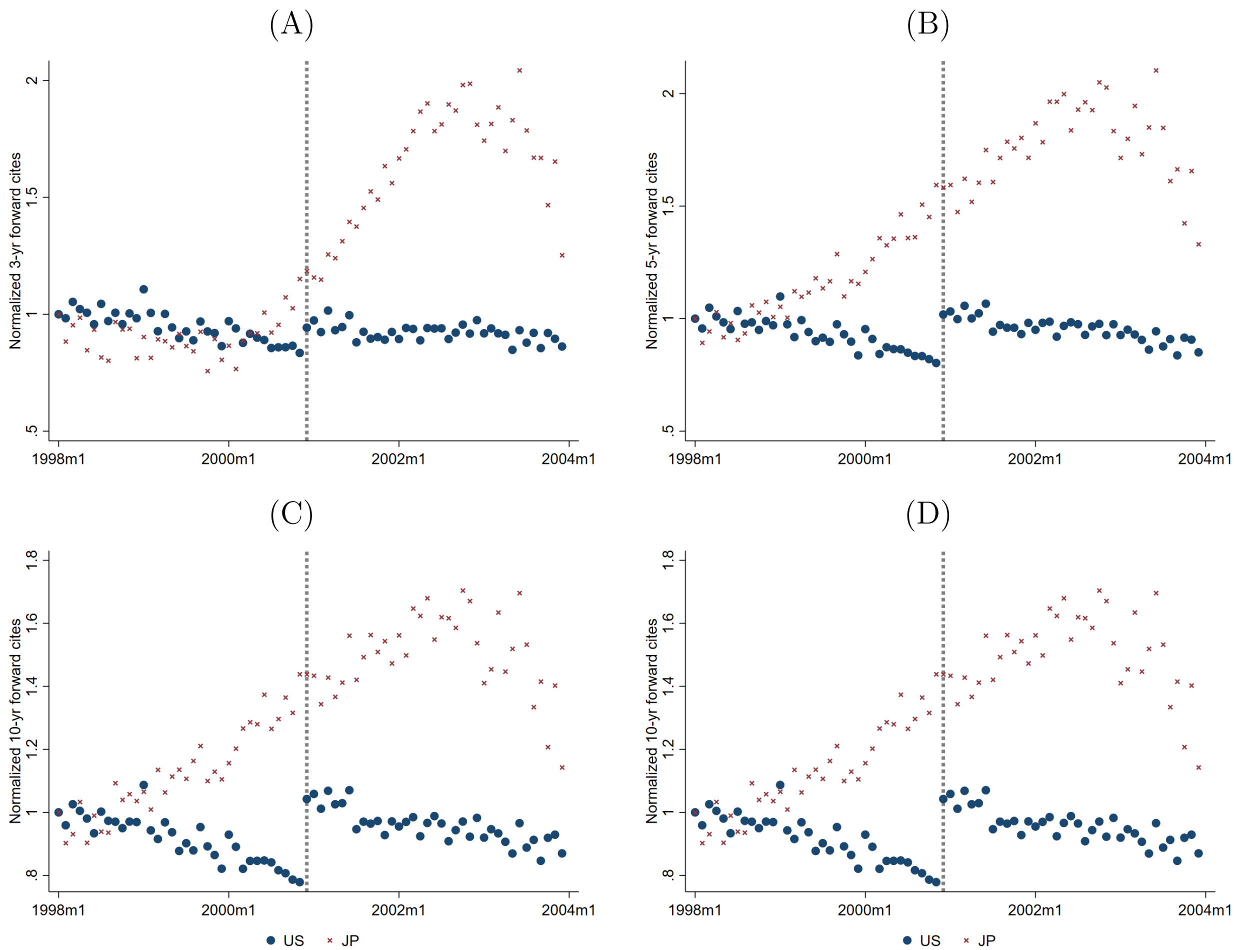


Table D.1: Equivalence of patents in the same family

*Panel A of this table lists 10 random pairs of US-EP twins in our sample, hyperlinked to the original documents available on Google Patents. Panel B compares the average pair-wise similarity for each US-EP twin to its matched pair, which is constructed by matching the EP equivalent to a random patent filed in the EPO in the same quarter and with the same 4-digit IPC code, and grant status. We then compute the IPC-based cosine similarity or text-based similarity between the US patent and its EP equivalent or its EP match, respectively. We report the mean and median similarity for the two types of pairs (US-EP and US-EP Match) and test the significance of the difference in mean using paired t-test and the difference in median using sign test.*

**Panel A. Examples of US-EP twins**

| Examples | EP Patent | US Patent |
|---|---|---|
| Example 1 | EP1409998 | US7060292 |
| Example 2 | EP0983918 | US6182526 |
| Example 3 | EP1232093 | US6715408 |
| Example 4 | EP0846671 | US6417408 |
| Example 5 | EP0931368 | US6422890 |
| Example 6 | EP1202689 | US6261254 |
| Example 7 | EP1573981 | US7286481 |
| Example 8 | EP1160300 | US6559056 |
| Example 9 | EP1435016 | US6488372 |
| Example 10 | EP1155319 | US6782909 |

**Panel B. Within-twin similarity**

| | **Mean** | | | **Median** | | |
|---|---|---|---|---|---|---|
| | US-EP | US-EP Match | Difference | US-EP | US-EP Match | Difference |
| Similarity (7-digit IPC) | 0.971 | 0.564 | 0.407*** | 1.000 | 0.555 | 0.445*** |
| (s.e.) | (0.000) | (0.001) | (0.001) | (.) | (0.001) | |
| Similarity (Google Text) | 0.971 | 0.533 | 0.438*** | 0.989 | 0.522 | 0.467*** |
| (s.e.) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | |

## Table D.2: Alternative fixed effect specifications

*This table reports the DID analyses of AIPA's effect on forward citations, citation lags, and technology similarity using the US-EP twin sample with alternative fixed effect specifications. We include the same set of control variables as our baseline regressions but only report the coefficients on the main variable of interest ('Post AIPA × US (d)') for brevity. Panel A controls for technology class fixed effects (4-digit IPC codes) and Panel B controls applicant fixed effects (first applicant). Standard errors are clustered by application month for U.S. and EP patents, separately. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

| **Panel A: Technology class fixed effects** | | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Forward Citations** | | | | |
| | Log 3-Yr. Forward Cites | Log 5-Yr. Forward Cites | Log 7-Yr. Forward Cites | Log 10-Yr. Forward Cites |
| Post AIPA × US (d) | -0.017*** | 0.052*** | 0.096*** | 0.133*** |
| | (0.005) | (0.006) | (0.006) | (0.006) |
| Observations | 669,556 | 668,113 | 659,888 | 647,767 |
| Adjusted R-squared | 0.268 | 0.309 | 0.332 | 0.357 |
| **Citation Lags** | | | | |
| | Log Months to 1 Cite | Log Months to 3 Cites | Log Months to 5 Cites | Log Months to 7 Cites |
| Post AIPA × US (d) | -0.177*** | -0.185*** | -0.176*** | -0.147*** |
| | (0.007) | (0.008) | (0.011) | (0.014) |
| Observations | 454,223 | 318,579 | 239,384 | 188,486 |
| Adjusted R-squared | 0.127 | 0.127 | 0.129 | 0.129 |
| **technology Similarity** | | | | |
| | Similarity 50th Pctile | Similarity 75th Pctile | Similarity 90th Pctile | Similarity 95th Pctile |
| Post AIPA × US (d) | 0.011*** | 0.009*** | -0.008*** | -0.013*** |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Observations | 668,713 | 668,713 | 668,713 | 668,713 |
| Adjusted R-squared | 0.306 | 0.222 | 0.162 | 0.142 |

| **Panel B: Applicant fixed effects** | | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Forward Citations** | | | | |
| | Log 3-Yr. Forward Cites | Log 5-Yr. Forward Cites | Log 7-Yr. Forward Cites | Log 10-Yr. Forward Cites |
| Post AIPA × US (d) | -0.003 | 0.069*** | 0.115*** | 0.151*** |
| | (0.005) | (0.006) | (0.006) | (0.007) |
| Observations | 638,902 | 637,582 | 629,758 | 618,261 |
| Adjusted R-squared | 0.319 | 0.363 | 0.386 | 0.409 |
| **Citation Lags** | | | | |
| | Log Months to 1 Cite | Log Months to 3 Cites | Log Months to 5 Cites | Log Months to 7 Cites |
| Post AIPA × US (d) | -0.207*** | -0.220*** | -0.209*** | -0.187*** |
| | (0.008) | (0.008) | (0.010) | (0.011) |
| Observations | 421,321 | 287,231 | 211,446 | 163,862 |
| Adjusted R-squared | 0.158 | 0.169 | 0.181 | 0.192 |
| **Technology Similarity** | | | | |
| | Similarity 50th Pctile | Similarity 75th Pctile | Similarity 90th Pctile | Similarity 95th Pctile |
| Post AIPA × US (d) | 0.011*** | 0.008*** | -0.009*** | -0.013*** |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Observations | 637,772 | 637,772 | 637,772 | 637,772 |
| Adjusted R-squared | 0.177 | 0.193 | 0.195 | 0.192 |

## Table D.3: Country-technology specific time trend

*This table reports the DID analyses of AIPA's effect on forward citations, citation lags, and technology similarity after controlling for country-specific-technology linear time trends using the US-EP twin sample. The regression specification is the same as our main specification except that we add a linear trend by one-digit IPC technology class (8 classes in total, represented by A-H, respectively) for both USPTO and EPO, respectively. We only report the coefficients on the main variable of interest ('Post AIPA × US (d)') for brevity. Standard errors are clustered by application month for U.S. and EP patents, separately. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

| **Forward Citations** | | | | |
|---|---|---|---|---|
| | (1)<br>Log 3-Yr. Forward Cites | (2)<br>Log 5-Yr. Forward Cites | (3)<br>Log 7-Yr. Forward Cites | (4)<br>Log 10-Yr. Forward Cites |
| Post AIPA × US (d) | -0.028 | 0.068* | 0.111*** | 0.135*** |
| | (0.030) | (0.034) | (0.037) | (0.040) |
| Country-Tech Trend | Yes | Yes | Yes | Yes |
| Observations | 669,446 | 667,785 | 658,722 | 642,212 |
| Adjusted R-squared | 0.469 | 0.530 | 0.562 | 0.589 |

| **Citation Lags** | | | | |
|---|---|---|---|---|
| | (1)<br>Log Months to 1 Cite | (2)<br>Log Months to 3 Cites | (3)<br>Log Months to 5 Cites | (4)<br>Log Months to 7 Cites |
| Post AIPA × US (d) | -0.194*** | -0.234*** | -0.241*** | -0.212*** |
| | (0.037) | (0.029) | (0.029) | (0.031) |
| Country-Tech Trend | Yes | Yes | Yes | Yes |
| Observations | 317,650 | 149,333 | 81,338 | 49,396 |
| Adjusted R-squared | 0.280 | 0.321 | 0.331 | 0.341 |

| **Technology Similarity** | | | | |
|---|---|---|---|---|
| | (1)<br>Similarity 50th Pctile | (2)<br>Similarity 75th Pctile | (3)<br>Similarity 90th Pctile | (4)<br>Similarity 95th Pctile |
| Post AIPA × US (d) | 0.004*** | 0.003 | -0.004* | -0.005** |
| | (0.001) | (0.002) | (0.002) | (0.002) |
| Country-Tech Trend | Yes | Yes | Yes | Yes |
| Observations | 666,998 | 666,998 | 666,998 | 666,998 |
| Adjusted R-squared | 0.624 | 0.686 | 0.699 | 0.701 |

Table D.4: DID analysis with linear pre-trends

*This table reports the DID analyses of AIPA's effect on forward citations (Panel A), citation lags (Panel B), and technology similarity (Panel C) using the US-EP twin sample. The regressions are specified as follows:*

$$Outcome_{ijt} = \alpha_0 + \alpha_1 \mathbf{I}(US_j) + f(Month_t) + \delta W_j + \sum_{\tau \in Post} \beta_t \mathbf{I}\{Month_t = \tau\} + \sum_{\tau \in Post} \gamma_t \mathbf{I}(US_j) \times \mathbf{I}\{Month_t = \tau\} + \epsilon_{ijt}$$

*where $j$ indicates the patent application belonging to family $i$ and filed in month $t$. $\mathbf{I}(US_j)$ indicates whether the patent is filed in the USPTO. We include a linear pre-trend in $f(Month_t)$ and a set of dummy variables indicating each month in the post-AIPA period to identify the deviations from the pre-trend. We control for the same set of patent characteristics ($W_j$) as the baseline DID analyses in Tables 4-6. For brevity, we only report 'Common Pre-Post Difference' and 'AIPA-Effect', computed as the mean of $\beta_t$ and $\gamma_t$, respectively. The associated standard errors are computed using the delta method. Standard errors are clustered by application month for U.S. and EP patents, separately. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

| **Panel A: Forward Citations** | | | | |
|---|---|---|---|---|
| | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
| Common Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.028 | 0.025 | 0.035 | 0.038* |
| | (0.018) | (0.021) | (0.021) | (0.021) |
| AIPA Effect ($\frac{1}{37}\sum_{\tau \in Post} \gamma_\tau$) | -0.036*** | 0.030*** | 0.078*** | 0.113*** |
| | (0.008) | (0.010) | (0.010) | (0.010) |
| Observations | 670,142 | 669,708 | 668,373 | 659,620 |
| Adj R-squared | 0.215 | 0.256 | 0.281 | 0.305 |

| **Panel B: Citation Lags** | | | | |
|---|---|---|---|---|
| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Common Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | 0.008 | 0.007 | 0.017 | 0.001 |
| | (0.021) | (0.016) | (0.016) | (0.018) |
| AIPA Effect ($\frac{1}{37}\sum_{\tau \in Post} \gamma_\tau$) | -0.148*** | -0.171*** | -0.173*** | -0.153*** |
| | (0.010) | (0.007) | (0.008) | (0.010) |
| Observations | 454,497 | 318,794 | 239,577 | 188,668 |
| Adj R-squared | 0.080 | 0.048 | 0.033 | 0.027 |

| **Panel C: Technology Similarity** | | | | |
|---|---|---|---|---|
| | (1) Similarity 50th Pc-tile | (2) Similarity 75th Pc-tile | (3) Similarity 90th Pc-tile | (4) Similarity 95th Pc-tile |
| Common Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post} \beta_\tau$) | -0.000 | 0.006 | 0.013*** | 0.010* |
| | (0.002) | (0.004) | (0.005) | (0.005) |
| AIPA Effect ($\frac{1}{37}\sum_{\tau \in Post} \gamma_\tau$) | 0.011*** | 0.007*** | -0.012*** | -0.017*** |
| | (0.001) | (0.002) | (0.002) | (0.002) |
| Observations | 454,497 | 318,794 | 239,577 | 188,668 |
| Adj R-squared | 0.010 | 0.013 | 0.012 | 0.011 |

## Table D.5: Excluding the year before AIPA

*This table reports the DID analyses of AIPA's effect on forward citations (Panel A), citation lags (Panel B), and technology similarity (Panel C) using the US-EP twin sample after excluding patents filed in the year before AIPA became effect (i.e., from November 29, 1999 to November 28, 2000). We include the same set of control variables as our baseline DID regressions in Tables 4-6 but only report the coefficients on the main variable of interest ('Post AIPA × US (d)') for brevity. All regressions include patent family fixed effects and application month fixed effects. Standard errors are clustered by application month for U.S. and EP patents, separately. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

| **Panel A: Forward Citations** | | | | |
|---|---|---|---|---|
| | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
| Post AIPA × US (d) | -0.001 | 0.072*** | 0.118*** | 0.165*** |
| | (0.008) | (0.009) | (0.009) | (0.010) |
| Observations | 502,500 | 501,136 | 493,613 | 482,157 |
| Adj R-squared | 0.456 | 0.520 | 0.554 | 0.579 |

| **Panel B: Citation Lags** | | | | |
|---|---|---|---|---|
| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Post AIPA × US (d) | -0.270*** | -0.305*** | -0.304*** | -0.287*** |
| | (0.011) | (0.011) | (0.014) | (0.015) |
| Observations | 240,710 | 113,524 | 61,643 | 37,172 |
| Adj R-squared | 0.285 | 0.323 | 0.332 | 0.343 |

| **Panel C: Technology Similarity** | | | | |
|---|---|---|---|---|
| | (1) Similarity 50th Pctile | (2) Similarity 75th Pctile | (3) Similarity 90th Pctile | (4) Similarity 95th Pctile |
| Post AIPA × US (d) | 0.015*** | 0.012*** | -0.008*** | -0.013*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Observations | 499,856 | 499,856 | 499,856 | 499,856 |
| Adj R-squared | 0.609 | 0.674 | 0.688 | 0.689 |

## Table D.6: Placebo AIPA Date

*This table reports the placebo DID analyses of AIPA's effect on forward citations, citation lags, and technology similarity using the US-EP twin sample. We only use the pre-AIPA period (January 1998-November 2000) and split this sample into two subsamples in the middle. We set 'Placebo Post-AIPA' to be one if the patent is filed on and after July 1st, 1999. We include the same set of control variables as our baseline DID regressions in Tables 4-6. We only present the coefficients on the interaction term 'Placebo Post-AIPA × US (d)' for brevity. Standard errors are clustered by application month for U.S. and EP patents, separately. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

| **Forward Citations** | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
|---|---|---|---|---|
| Placebo Post-AIPA × US (d) | -0.022 | -0.034 | -0.042 | -0.029 |
| | (0.023) | (0.028) | (0.030) | (0.033) |
| Observations | 318,472 | 317,998 | 314,749 | 306,537 |
| Adjusted R-squared | 0.430 | 0.490 | 0.520 | 0.546 |

| **Citation Lags** | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
|---|---|---|---|---|
| Placebo Post-AIPA × US (d) | 0.021 | 0.012 | 0.010 | 0.003 |
| | (0.024) | (0.019) | (0.019) | (0.020) |
| Observations | 155,139 | 73,813 | 41,116 | 25,551 |
| Adjusted R-squared | 0.227 | 0.266 | 0.279 | 0.295 |

| **Technology Similarity** | (1) Similarity 50th Pc-tile | (2) Similarity 75th Pc-tile | (3) Similarity 90th Pc-tile | (4) Similarity 95th Pc-tile |
|---|---|---|---|---|
| Placebo Post-AIPA × US (d) | 0.003*** | 0.001 | 0.000 | -0.000 |
| | (0.001) | (0.002) | (0.002) | (0.002) |
| Observations | 315,301 | 315,301 | 315,301 | 315,301 |
| Adjusted R-squared | 0.656 | 0.728 | 0.751 | 0.752 |

## Table D.7: Restricting to twins with US patents granted within 18 months

*This table reports the DID analyses of AIPA's effect on forward citations and citation lags using the US-EP twins where the US twin is granted within 18 months. The regression specification is the same as the main analysis except that we exclude the interaction term 'Early Grant × US (d)'. We only present the coefficients on the interaction term 'Post-AIPA × US (d)' for brevity. Standard errors are clustered by the application month for U.S. and EP patents, separately. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively*

| **Forward Citations** | | | | |
|---|---|---|---|---|
| | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
| Post AIPA × US (d) | -0.032*** | -0.013 | -0.005 | 0.023* |
| | (0.010) | (0.012) | (0.012) | (0.014) |
| Observations | 87,731 | 87,685 | 87,272 | 86,624 |
| Adjusted R-squared | 0.451 | 0.512 | 0.544 | 0.572 |
| **Citation Lags** | | | | |
| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Post AIPA × US (d) | -0.008 | -0.070*** | -0.086*** | -0.060** |
| | (0.017) | (0.017) | (0.024) | (0.029) |
| Observations | 39,268 | 17,211 | 8,541 | 4,842 |
| Adjusted R-squared | 0.301 | 0.357 | 0.369 | 0.407 |

# E Adjusting for Continuations

In this section, we proxy the invention date using the priority date (rather than application date used in our main analysis). The priority date adjusts for continuation, divisional, and continuation-in-parts and is arguably closer to the date of invention than the application date. In Table E.1, citation lag is measured as the average time between the patent priority dates of the focal patent and its first 1/3/5/7 forward citation(s). We find significant reductions in citation lags, similar to our baseline results.

To further demonstrate the robustness of our results to continuation filings, in Table E.2, we repeat our baseline analysis after excluding continuation patents. Specifically, from PATSTAT's application continuation table, we identify continuation patents if they have one of the following relationship with prior patent applications: continuation (pursue additional claims to an invention disclosed in an earlier application by the applicant, i.e., parent application), continuation-in-part (add subject matter not disclosed in the parent application, but repeat a substantial portion of the parent's specification), and division (claim a distinct or independent invention "carved out" of the parent application). Our sample contains 63,365 (9.45%) continuation filings. Excluding these patents yields very similar results to our baseline.

Table E.1: Continuation patent filings

*This table reports DID estimates of AIPA's effect on a modified measure of citation lag using the US-EP twin sample. To account for the influence of continuation filings, the dependent variable is measured as the average time between the patent priority date (rather than application date used in our main analysis) of the focal patent and its first 1, 3,5,7 forward citation(s). Standard errors are clustered by the application month for U.S. and EP patents, separately. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively.*

| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
|---|---|---|---|---|
| US (d) | -0.592*** | -0.409*** | -0.357*** | -0.344*** |
| | (0.010) | (0.009) | (0.009) | (0.011) |
| Post AIPA × US (d) | -0.160*** | -0.219*** | -0.219*** | -0.221*** |
| | (0.011) | (0.012) | (0.014) | (0.018) |
| Granted (d) | -0.191*** | -0.163*** | -0.125*** | -0.123*** |
| | (0.014) | (0.011) | (0.014) | (0.017) |
| Early Grant (d) | -0.401*** | -0.323*** | -0.273*** | -0.235*** |
| | (0.015) | (0.017) | (0.021) | (0.025) |
| Early Grant × US (d) | 0.273*** | 0.144*** | 0.098*** | 0.089** |
| | (0.020) | (0.025) | (0.031) | (0.034) |
| Observations | 276,967 | 132,159 | 73,394 | 45,623 |
| Adj R-squared | 0.294 | 0.320 | 0.336 | 0.345 |

## Table E.2: Continuation patent filings

*This table address the concern that continuation filings affect our results. We drop continuation filings from the sample (US-EP twins) and repeats our analysis of forward citations, citation lags, and technological similarity in Panels A-C, respectively. Standard errors are clustered by the application month for U.S. and EP patents, separately. ***, **, and * stand for statistical significance based on two-sided tests at the 1%, 5%, and 10% level, respectively.*

**Panel A. Forward Citations**

| | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
|---|---|---|---|---|
| US (d) | 0.755*** | 0.927*** | 1.029*** | 1.145*** |
| | (0.006) | (0.007) | (0.008) | (0.009) |
| Post AIPA × US (d) | -0.002 | 0.073*** | 0.126*** | 0.168*** |
| | (0.008) | (0.009) | (0.010) | (0.010) |
| Granted (d) | 0.175*** | 0.217*** | 0.245*** | 0.282*** |
| | (0.006) | (0.007) | (0.007) | (0.008) |
| Early Grant (d) | 0.370*** | 0.453*** | 0.500*** | 0.549*** |
| | (0.009) | (0.011) | (0.012) | (0.013) |
| Early Grant × US (d) | -0.421*** | -0.514*** | -0.569*** | -0.637*** |
| | (0.010) | (0.014) | (0.017) | (0.019) |
| Observations | 573,787 | 573,392 | 571,967 | 563,478 |
| Adj R-squared | 0.456 | 0.517 | 0.545 | 0.569 |

**Panel B. Citation Lags**

| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
|---|---|---|---|---|
| US (d) | -0.338*** | -0.156*** | -0.096*** | -0.069*** |
| | (0.009) | (0.007) | (0.007) | (0.008) |
| Post AIPA × US (d) | -0.289*** | -0.321*** | -0.314*** | -0.297*** |
| | (0.012) | (0.010) | (0.013) | (0.015) |
| Granted (d) | -0.189*** | -0.162*** | -0.129*** | -0.115*** |
| | (0.009) | (0.010) | (0.013) | (0.017) |
| Early Grant (d) | -0.338*** | -0.288*** | -0.242*** | -0.210*** |
| | (0.012) | (0.014) | (0.019) | (0.024) |
| Early Grant × US (d) | 0.088*** | 0.024 | 0.003 | -0.012 |
| | (0.023) | (0.025) | (0.031) | (0.034) |
| Observations | 271,708 | 127,687 | 69,313 | 41,774 |
| Adj R-squared | 0.252 | 0.297 | 0.310 | 0.329 |

**Panel C. Technology Similarity**

| | (1) Similarity 50th Pc-tile | (2) Similarity 75th Pc-tile | (3) Similarity 90th Pc-tile | (4) Similarity 95th Pc-tile |
|---|---|---|---|---|
| US (d) | 0.023*** | 0.049*** | 0.061*** | 0.056*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Post AIPA × US (d) | 0.012*** | 0.010*** | -0.009*** | -0.014*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Granted (d) | 0.011*** | 0.021*** | 0.012*** | 0.005*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Early Grant (d) | 0.001 | 0.007*** | 0.008*** | 0.006*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Early Grant × US (d) | -0.006*** | -0.020*** | -0.018*** | -0.007*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Observations | 570,879 | 570,879 | 570,879 | 570,879 |
| Adj R-squared | 0.606 | 0.675 | 0.689 | 0.690 |

# F Results without the Harhoff Adjustment

In the main text, following Harhoff et al. (2006), we adjust for patent equivalents when counting forward citations for EPO patents. Specifically, if a future EPO patent cites a U.S. patent but not its EPO equivalent, the adjustment will count it as one citation for the EPO equivalent as well. Table F.1 repeats our main analysis without this adjustment. We find stronger results for forward citations, and moderately weaker results for citation lags at longer horizons.

Table F.1: Undo Harhoff Adjustment

*This table reports DID estimates of AIPA's effect on forward citations in Panel A and on citation lags in Panel B after undoing the Hahoff adjustment. The regression specifications are similar to our baseline DID regressions. All regressions include patent family fixed effects and application month fixed effects. Standard errors are clustered by application month for U.S. and EP patents, separately. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

**Panel A: Forward Citations**

| | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
|---|---|---|---|---|
| US (d) | 0.874*** | 1.116*** | 1.259*** | 1.408*** |
| | (0.007) | (0.009) | (0.010) | (0.011) |
| Post AIPA × US (d) | 0.031*** | 0.092*** | 0.133*** | 0.171*** |
| | (0.009) | (0.011) | (0.012) | (0.013) |
| Granted (d) | 0.211*** | 0.262*** | 0.293*** | 0.336*** |
| | (0.007) | (0.008) | (0.008) | (0.010) |
| Early Grant (d) | 0.506*** | 0.631*** | 0.696*** | 0.768*** |
| | (0.010) | (0.012) | (0.013) | (0.015) |
| Early Grant × US (d) | -0.526*** | -0.658*** | -0.725*** | -0.809*** |
| | (0.013) | (0.018) | (0.021) | (0.023) |
| Observations | 669,877 | 668,218 | 659,154 | 642,640 |
| Adjusted R-squared | 0.394 | 0.449 | 0.477 | 0.503 |

**Panel B: Citation Lags**

| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
|---|---|---|---|---|
| US (d) | -0.394*** | -0.214*** | -0.150*** | -0.121*** |
| | (0.009) | (0.008) | (0.008) | (0.009) |
| Post AIPA × US (d) | -0.254*** | -0.294*** | -0.292*** | -0.274*** |
| | (0.011) | (0.011) | (0.013) | (0.014) |
| Granted (d) | -0.201*** | -0.178*** | -0.147*** | -0.139*** |
| | (0.013) | (0.010) | (0.013) | (0.016) |
| Early Grant (d) | -0.393*** | -0.342*** | -0.298*** | -0.255*** |
| | (0.013) | (0.014) | (0.018) | (0.022) |
| Early Grant × US (d) | 0.149*** | 0.079*** | 0.059* | 0.031 |
| | (0.022) | (0.024) | (0.030) | (0.033) |
| Observations | 317,853 | 149,423 | 81,400 | 49,424 |
| Adj R-squared | 0.272 | 0.309 | 0.316 | 0.328 |

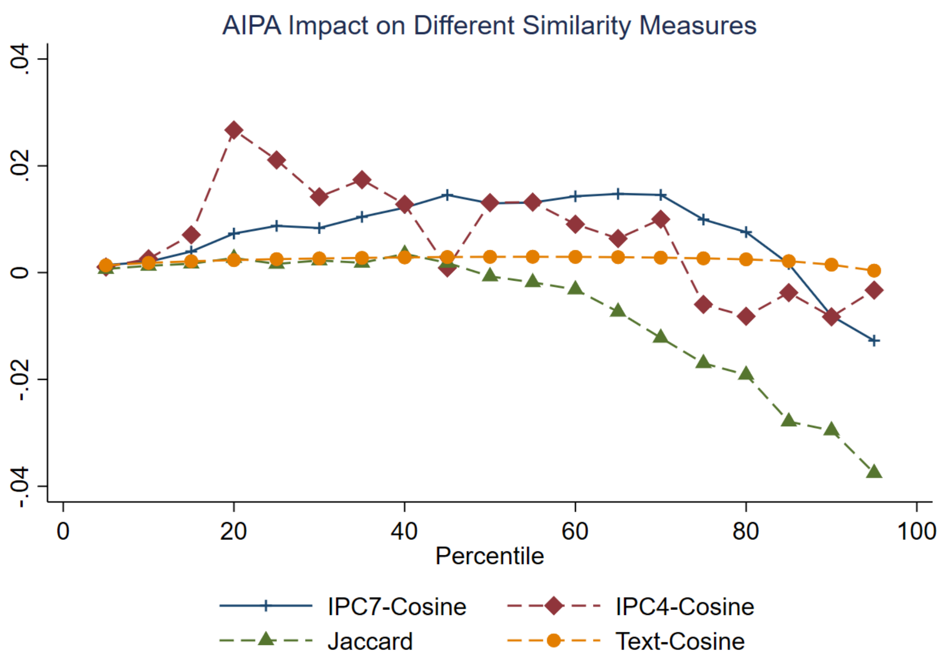
# G Alternate Patent Similarity Measures

In Appendix Figure G.1 (with corresponding Table G.1), we consider alternate measures of technology similarity, including Jaccard similarity, and alternate levels of aggregation, including IPC-4 digit codes. We find very similar results across all percentiles to our baseline results which use cosine-distance for IPC-7 digit codes. We also construct a text-based cosine similarity by using word-embedding vectors, which encode information from the patent text with a machine learning model developed by Google.[38] The Google text-based measures of similarity yield similar results at lower percentiles (below p75, roughly), but exhibit no significant reduction in similarity at higher percentiles. It might be caused by issues when classifying large bodies of text into sparse vectors. The Google patent text project provides vectors of 64 dimensions and we found the measure to be too blunt – that is, it does not sufficiently differentiate among technologically close patents. We found much less variance in similarity (the 50th percentile text-based similarity has a mean of 0.501 and a standard deviation of 0.088) using the text-based measure versus measures based on IPC-4 or IPC-7 cosine distance (the 50th percentile IPC7-based similarity has a mean of 0.099 and a standard deviation of 0.185).

[38] The data is downloaded from Google Patents Public Datasets on BigQuery. See a brief description about the data at https://cloud.google.com/blog/topics/public-datasets/google-patents-public-datasets-connecting-public-paid-and-private-patent-data

Figure G.1: AIPA effect on alternative similarity measures

*This figure plots the estimated AIPA effect on different proxies for technology similarity measured at different percentiles from the 5th percentile to the 95th percentile. The estimated AIPA effect is the coefficient on the interaction term ('Post AIPA × US (d)') where a positive (negative) coefficient indicates an increase (decrease) in technology similarity. Refer to Table 6 notes for a description of the regression specifications.*



## Table G.1: Alternative similarity measures

*This table reports DID estimates of AIPA's effect on technological similarity. Panel A uses Jaccard similarity based on 7-digit IPC codes, Panel B uses cosine similarity based on 4-digit IPC codes, and Panel C uses cosine similarity based on text-based vectors (word embeddings provided by Google Patents). Standard errors are clustered by application month for U.S. and EP patents, separately. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

**Panel A. Jaccard similarity based on 7-digit IPC codes**

| | (1) Jaccard Similarity 50th Pctile | (2) Jaccard Similarity 75th Pctile | (3) Jaccard Similarity 90th Pctile | (4) Jaccard Similarity 95th Pctile |
|---|---|---|---|---|
| US (d) | -0.005*** | 0.002*** | 0.005*** | 0.010*** |
| | (0.000) | (0.001) | (0.001) | (0.001) |
| Post AIPA × US (d) | -0.001*** | -0.017*** | -0.029*** | -0.037*** |
| | (0.000) | (0.001) | (0.002) | (0.002) |
| Granted (d) | 0.006*** | 0.010*** | 0.008*** | 0.005*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Early Grant (d) | 0.003*** | 0.002 | 0.005*** | 0.006*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Early Grant × US (d) | -0.002** | 0.000 | 0.003** | 0.006*** |
| | (0.001) | (0.001) | (0.002) | (0.002) |
| Family FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 670,150 | 670,150 | 670,150 | 670,150 |
| Adj R-squared | 0.889 | 0.900 | 0.905 | 0.903 |

**Panel B. Cosine similarity based on 4-digit IPC codes**

| | (1) Cosine Similarity 50th Pctile | (2) Cosine Similarity 75th Pctile | (3) Cosine Similarity 90th Pctile | (4) Cosine Similarity 95th Pctile |
|---|---|---|---|---|
| US (d) | 0.005*** | 0.012*** | -0.022*** | -0.013*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Post AIPA × US (d) | 0.011*** | -0.002* | -0.008*** | -0.003*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Granted (d) | 0.025*** | 0.016*** | 0.013*** | 0.008*** |
| | (0.002) | (0.001) | (0.001) | (0.001) |
| Early Grant (d) | -0.016*** | -0.008*** | 0.010*** | 0.008*** |
| | (0.003) | (0.002) | (0.002) | (0.001) |
| Early Grant × US (d) | 0.011*** | -0.001 | -0.014*** | -0.014*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Family FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 670,150 | 670,150 | 670,150 | 670,150 |
| Adj R-squared | 0.889 | 0.900 | 0.905 | 0.903 |

**Panel C. Text-based similarity**

| | (1) Text Similarity 50th Pctile | (2) Text Similarity 75th Pctile | (3) Text Similarity 90th Pctile | (4) Text Similarity 95th Pctile |
|---|---|---|---|---|
| US (d) | 0.00931*** | 0.00669*** | 0.00295*** | 0.00170*** |
| | (0.000176) | (0.000198) | (0.000221) | (0.000234) |
| Post AIPA × US (d) | 0.00296*** | 0.00268*** | 0.00149*** | 0.000380 |
| | (0.000193) | (0.000216) | (0.000233) | (0.000253) |
| Granted (d) | 0.00595*** | 0.00822*** | 0.0107*** | 0.0125*** |
| | (0.000266) | (0.000348) | (0.000440) | (0.000477) |
| Early Grant (d) | -0.00134*** | -0.00108*** | -0.000862* | -0.000531 |
| | (0.000283) | (0.000358) | (0.000439) | (0.000486) |
| Early Grant × US (d) | -0.00464*** | -0.00743*** | -0.00829*** | -0.00741*** |
| | (0.000344) | (0.000408) | (0.000505) | (0.000566) |
| Family FE | Yes | Yes | Yes | Yes |
| Month FE | Yes | Yes | Yes | Yes |
| Observations | 670,212 | 670,212 | 670,212 | 670,212 |
| Adj R-squared | 0.809 | 0.765 | 0.754 | 0.781 |

# H R&D Investments around AIPA

## H.1 Public Company R&D Summary Statistics

Table H.1 provides summary statistics for the Compustat firms analyzed in Table 12 after imposing different sales filters. Trimming the sample by excluding firms with sales less than 10/20/50 million dollars affects 'R&D to Sales Ratio' but it has minimal impact on 'Log R&D'.

## H.2 Time Trend of AIPA Effects on R&D

Figure H.1 plots log(R&D+1), after winsorizing at the 1% level, over time around AIPA for those in the top and the bottom AIPA exposure quintile, respectively. We require firms to have sales greater than $10m to be in the sample, similar to the criteria described in Section 7. The two groups appear to follow similar trends prior to AIPA and then diverge thereafter. We view this as evidence that pre-trends are unlikely to explain our main results.

## H.3 Heterogeneous R&D Effects

Table H.2 examines AIPA's heterogeneous effects on corporate R&D investments along the dimensions of technology cycle, opt-out ratio, and technology similarity. Instead of computing a focal firm's AIPA exposure as the median application-grant delay across *all* technology classes (weighted by the firm's patent share in every technology class), we compute AIPA exposure *within* certain subset of technologies. In particular, we compute a focal firm's AIPA exposure as the median application-grant delay among the subset of technology classes with above/below median: (1) length of technology cycle, (2) opt-out rates, and (3) technology similarity.

Columns (1)-(3) show that the AIPA effect on R&D concentrates on technology classes with a long technology cycle (as measured the average gap of the application dates of citing patents to cited ones, aggregated by the technology class of citing patents), across different sample size restrictions. This evidence is consistent with the intuition that timely disclosures are more meaningful to inventors in fast-moving fields. Columns (4)-(6) show that the AIPA effect is mainly driven by technology classes with higher opt-out ratio, whose 18-month disclosures presumably contain more valuable information.[39] In Columns (7)-(9), we find that the AIPA effect is stronger in areas with low similarity, where knowledge diffusion is more limited to begin with.

[39] Despite that more patents opt out of the 18-month disclosure requirement in these classes, the total information reveled in 18-month disclosures can be more valuable to follow-on inventors in the same field. The mere fact of high opt-out ratio indicates that typical patent filings in these classes indeed contain important information. Note that patents are mandated to be disclosed in 18 months when they have parallel applications, which include most valuable patents.

Figure H.1: Log R&D for the top and bottom quintile of AIPA exposure

*This figure plots average log(R&D+1), after winsorizing at the 1% level, around AIPA for those in the top and the bottom AIPA exposure quintile, respectively.*

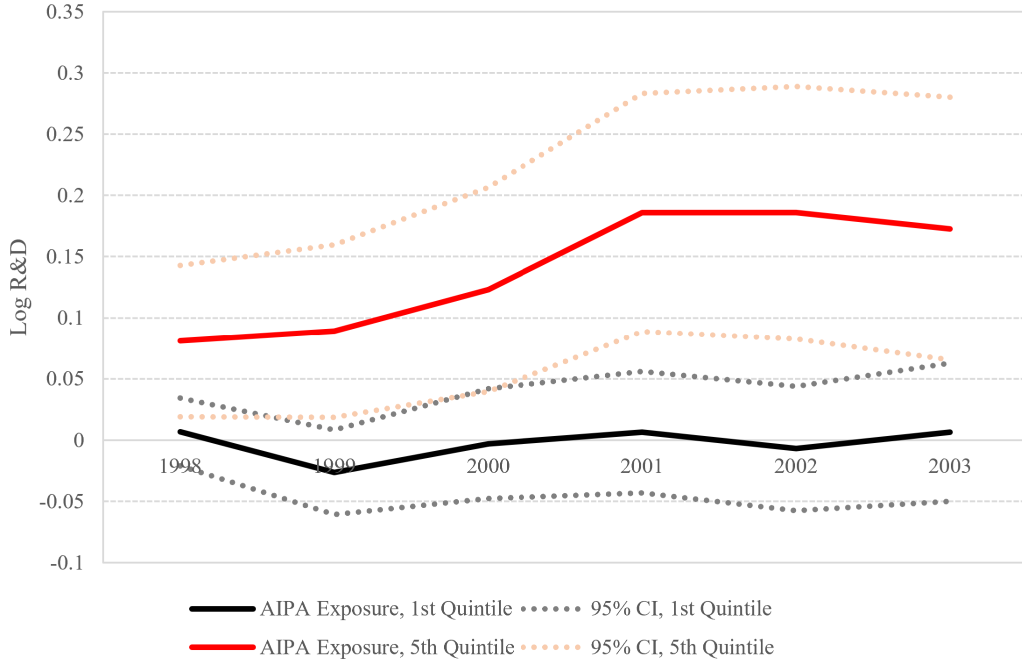



Table H.1: Compustat R&D Summary Statistics

*This table provides summary statistics for the Compustat firms after imposing different sales filters.*

| **Variable** | **Sample** | **Mean** | **SD** | **N** |
|---|---|---|---|---|
| R&D to Sales Ratio | All | 1.664 | 49.911 | 20170 |
| | Sales ≥10m | 0.127 | 0.401 | 17978 |
| | Sales ≥50m | 0.066 | 0.139 | 13693 |
| Log R&D | All | 1.865 | 1.828 | 20391 |
| | Sales ≥10m | 1.935 | 1.89 | 17978 |
| | Sales ≥50m | 2.096 | 2.049 | 13693 |
| AIPA Exposure (in years) | All | 1.399 | 1.391 | 20391 |
| | Sales ≥10m | 1.414 | 1.395 | 17978 |
| | Sales ≥50m | 1.467 | 1.381 | 13693 |

Table H.2: AIPA's effect on R&D investment: Heterogeneous Exposures

*This table reports the impact of AIPA on R&D investment by U.S. public companies. The dependent variable is natural logarithm of R&D expenses. The sample size varies across columns as we impose different size restrictions. Post AIPA (d) is a dummy indicator equal one if the fiscal year ends after the effective date of AIPA (November 29, 2000) and zero otherwise. AIPA Exposure is the firm-specific exposure to the AIPA shock measured by the median application-grant delay (in years) by USPC technology class weighted by the firm's patent share in each class. We differentiate technology classes based on technology cycle (the median citation lag for each USPC during the three-year pre-AIPA period) in Columns 1-3, opt-out rate in Columns 4-6 (the opt-out rate for each USPC in the three years post AIPA), and technology similarity in Columns 7-9 (average similarity for each USPC in the three years pre AIPA). The portfolio weight is based on the firm's patent portfolio filed from January 1, 1998 to November 28, 2000 (inclusive). The exposure variable is set to be zero for firms without any patents. We include the same set of control variables as the baseline regressions but do not report their coefficients for brevity. Standard errors are clustered by firm. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Exposure of X | USPCs with Long Tech Cycle | | | USPCs with Low Opt-out Rate | | | USPCs with High Similarity | | |
| Exposure of Y | USPCs with Short Tech Cycle | | | USPCs with High Opt-out Rate | | | USPCs with Low Similarity | | |
| Sample | Sales≥10m | Sales≥20m | Sales≥50m | Sales≥10m | Sales≥20m | Sales≥50m | Sales≥10m | Sales≥20m | Sales≥50m |
| Dependent Variable | Log R&D | Log R&D | Log R&D | Log R&D | Log R&D | Log R&D | Log R&D | Log R&D | Log R&D |
| Post AIPA (d) | -0.031* | -0.035* | -0.051** | -0.022 | -0.025 | -0.039* | -0.027 | -0.031 | -0.047** |
| | (0.018) | (0.019) | (0.021) | (0.018) | (0.019) | (0.022) | (0.018) | (0.019) | (0.022) |
| Post AIPA (d) × AIPA Exposure of X | -0.000 | 0.001 | 0.009 | 0.012 | 0.008 | 0.007 | 0.003 | 0.002 | 0.005 |
| | (0.010) | (0.011) | (0.013) | (0.009) | (0.010) | (0.012) | (0.010) | (0.011) | (0.013) |
| Post AIPA (d) × AIPA Exposure of Y | 0.037*** | 0.036*** | 0.032*** | 0.019** | 0.020** | 0.020** | 0.029*** | 0.029*** | 0.028*** |
| | (0.008) | (0.008) | (0.009) | (0.008) | (0.008) | (0.010) | (0.008) | (0.009) | (0.010) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 17,336 | 15,780 | 13,141 | 17,336 | 15,780 | 13,141 | 17,336 | 15,780 | 13,141 |
| R-squared | 0.967 | 0.967 | 0.968 | 0.967 | 0.967 | 0.968 | 0.967 | 0.967 | 0.968 |

# I Excluding Software Patents

To address concerns that aggregate movements induced by the dot-com boom and bust affect our results, we exclude software patents as defined by Graham and Vishnubhakat (2013) and find similar results in Table I.1 for both event study of pre-post comparisons (Panel A) and the twin study of difference-in-differences analyses (Panel B). Note that the definition in Graham and Vishnubhakat (2013) is broad as it nearly identifies a third of all patents post-1995 as software patents.

Table I.1: Excluding software patents

*We replicate Table C.1 excluding software patents as defined in Graham and Vishnubhakat (2013). Standard errors are clustered by application month for U.S. and EP patents, separately. ***, **, and * indicate 1%, 5%, and 10% significance levels, resp.*

| | | | | |
|---|---|---|---|---|
| **Panel A. US before-after comparison** | | | | |
| **Forward Citations** | | | | |
| | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post}\beta_\tau$) | 0.001*** | 0.062*** | 0.107*** | 0.129*** |
| | (0.008) | (0.010) | (0.102) | (0.012) |
| Observations | 721,939 | 721,788 | 721,026 | 716,407 |
| R-squared | 0.115 | 0.129 | 0.138 | 0.147 |
| **Citation Lags** | | | | |
| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post}\beta_\tau$) | -0.155*** | -0.104*** | -0.090*** | -0.085*** |
| | (0.011) | (0.009) | (0.007) | (0.007) |
| Observations | 629,367 | 467,762 | 350,254 | 267,963 |
| R-squared | 0.078 | 0.097 | 0.105 | 0.110 |
| **Technology Similarity** | | | | |
| | (1) Similarity 50th Pctile | (2) Similarity 75th Pctile | (3) Similarity 90th Pctile | (4) Similarity 95th Pctile |
| Pre-Post Difference ($\frac{1}{37}\sum_{\tau \in Post}\beta_\tau$) | 0.010*** | 0.009*** | -0.004* | -0.014*** |
| | (0.012) | (0.002) | (0.002) | (0.029) |
| Observations | 721,532 | 721,532 | 721,532 | 721,532 |
| R-squared | 0.322 | 0.261 | 0.194 | 0.161 |
| **Panel B. US-EP Twin Sample** | | | | |
| **Forward Citations** | | | | |
| | (1) Log 3-Yr. Forward Cites | (2) Log 5-Yr. Forward Cites | (3) Log 7-Yr. Forward Cites | (4) Log 10-Yr. Forward Cites |
| Post AIPA × US (d) | -0.039*** | 0.026*** | 0.072*** | 0.112*** |
| | (0.007) | (0.009) | (0.009) | (0.010) |
| Observations | 469,793 | 468,714 | 462,740 | 452,039 |
| R-squared | 0.424 | 0.492 | 0.527 | 0.555 |
| **Citation Lags** | | | | |
| | (1) Log Months to 1 Cite | (2) Log Months to 3 Cites | (3) Log Months to 5 Cites | (4) Log Months to 7 Cites |
| Post AIPA × US (d) | -0.239*** | -0.258*** | -0.251*** | -0.237*** |
| | (0.011) | (0.011) | (0.017) | (0.019) |
| Observations | 222,997 | 99,491 | 51,668 | 30,170 |
| R-squared | 0.261 | 0.297 | 0.305 | 0.318 |
| **Technology Similarity** | | | | |
| | (1) Similarity 50th Pctile | (2) Similarity 75th Pctile | (3) Similarity 90th Pctile | (4) Similarity 95th Pctile |
| Post AIPA × US (d) | 0.014*** | 0.006*** | -0.006*** | -0.011*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Observations | 467,519 | 467,519 | 467,519 | 467,519 |
| R-squared | 0.648 | 0.707 | 0.711 | 0.712 |

Additional Analysis for Referees – Responses Only

"Patent Publication and Innovation" by Deepak Hegde, Kyle Herkenhoff, and Chenqi Zhu

Not be included in publication

February 11, 2022