RIETI Discussion Paper Series 22-E-074

# Language Barriers and the Speed of Knowledge Diffusion

**HIGHAM, Kyle**
Institute of Innovation Research, Hitotsubashi University

**NAGAOKA, Sadao**
RIETI

Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 22-E-074
**August 2022**

# Language barriers and the speed of knowledge diffusion[1]

Kyle Higham[2][1] and Sadao Nagaoka[2,3]

[1]Institute of Innovation Research, Hitotsubashi University, Japan
[2] Research Institute of Economy, Trade and Industry (RIETI), Japan
[3] Faculty of Economics, Tokyo Keizai University, Japan

Abstract

While language barriers are well-known obstacles to knowledge diffusion, quantitative research on this topic is sparse. In this work, we attempt to fill this gap by providing causal evidence on their effects on the speed of knowledge diffusion by exploiting the introduction of pre-grant publications by the American Inventors Protection Act (AIPA) in 2000. We find that this policy significantly accelerated, relative to Japanese inventors, US inventors' use of Japan-originating technical knowledge in their patents. Our analysis controls for biases of patent citations as proxies of knowledge flow, including preference for citing local prior art. Consistent with incentives for translation, this acceleration is much larger for small firms and the firms with little investment in the Japanese market. Consistent with high uncertainty of foreign patents before translation, we see much larger effects of the AIPA on the patent applications with higher quality. Our findings suggest that pre-grant publication provides a significant public good for cumulative innovation through earlier translations of foreign patents.

Key words: language barriers, knowledge diffusion, citations, public good

JEL classification: O31,O32, O34, H44



[1] This paper is part of the results of the project "Building innovation capability and incentive: evidence from micro data" at the Research Institute of Economy, Trade and Industry (RIETI). The members of the study group and the RIETI Discussion Paper Review Group provided many useful comments on this paper. We would like to express our gratitude for them.

[2] higham@iir.hit-u.ac.jp

# 1 Introduction

Language barriers have been a hindrance to communication since time immemorial. While the English language has become established as a lingua franca in international scientific and technological settings (Ammon, 2001; Tardy, 2004), these barriers persist for many in the global research ecosystem despite technological advances such as automated translation, and globalization more generally (Gibbs, 1995; Ferguson et al., 2011; Bowker and Ciro, 2019).

Significant heterogeneity exists in a country's ability and willingness to diffuse and absorb information to and from other countries, particularly with respect to technical knowledge. Indeed, some countries are generally more outward-facing than others with respect to communicating their scientific and technological advances. As they pertain to language barriers, intensity of knowledge flows can be moderated by a variety of factors, from inherent linguistic distances to other languages (Dickens, 2018) and effective foreign-language education (Nunan, 2003), to the perceived audience of a research output (Bentley, 2015), researchers' incentives to publish for an international audience in the first place (Shibayama and Baba, 2015), and the resources available to do so (Salager-Meyer, 2008). However, despite the central role that knowledge flows and spillovers play within contemporary innovation ecosystems and for economic growth more generally (Griliches, 1992; Grossman and Helpman, 1991; Coe and Helpman, 1995; Park, 1995; Aghion and Jaravel, 2015), there is very little existing research studying the ways in which language barriers causally impact these phenomena. Further, detailed knowledge about how language barriers may directly affect follow-on innovation is vital for a complete understanding of global technological development (Branstetter, 2000; Seck, 2012; Luintel and Khan, 2017).

In an initial step toward a deeper understanding of this potential friction, we ask two broad questions in the current work. First, insofar as patent citation information can tell us about international knowledge diffusion, are we able to observe accelerated diffusion upon the removal of a language barrier? Second, conditional on observing an acceleration, we may expect that the language barrier previously had heterogeneous impacts on access to knowledge. This heterogeneity likely varies across several dimensions, including the characteristics of the technology being diffused (or not), and the incentives and resources for translation by the firms and inventors who could use the knowledge. As such, we ask: which characteristics of both follow-on innovators and the inherent properties of the knowledge itself moderate the acceleration of knowledge diffusion? The answers to this question will provide insight into the nature of frictions in knowledge diffusion due to language barriers that still remain in place and inform any potential policy responses.

To tackle these questions, we consider a change in patent policy in the United States that suddenly gave the English-speaking world accelerated access to technological information produced in Japan — the American Inventors Protection Act (AIPA). For simplicity, we focus on knowledge spillovers to firms and inventors based in the United States, relative to those based in Japan. The Japanese language is an ideal setting for studying this phenomenon, particularly when considering knowledge diffusion to the US at the time of the policy change, for two reasons. First, the policy change we consider took place in 2000 while machine translation of technical documents was still in its infancy, hindered by both domain-specific terminology and the specific and complex sentence structure of patent claims (Cavalier, 2001). Second, Japanese is rarely spoken in the United States, so the language barrier is insurmountable for most US-based inventors unless extensive translations of the Japanese publications are available.. Third, Japan is a technologically advanced society with a significant amount of important technical knowledge being produced, particularly in fast-moving industries around the time of the policy change (Narin, 1995; Odagiri and Goto, 1996).

Our empirical strategy centers on the pure twin sample of patent families filed only at the Japanese Patent Office (JPO) and the United States Patent and Trademark Office (USPTO), by Japanese firms, to directly assess the effect of language barriers on knowledge diffusion. Before the AIPA, all of these patent families were published in Japanese 18 months after their first filing, but the English versions were not available until their USPTO grant. Indeed, the idea that US inventors often faced a language barrier with respect to foreign knowledge was even mentioned as part of the motivation for implementing the AIPA in the first place:

> *While our foreign competitors are able to see the latest U.S. patent technology in their native languages barely six months after a U.S. inventor files a patent application in their country, the reverse is not true. U.S. researchers and investors are denied the opportunity to learn what their foreign competitors are working on until a U.S. patent issues.*
>
> \- House Report 106-287 Part 1[1]

After the AIPA, the USPTO implemented a pre-grant publication program that also published applications, in English, 18 months after first filings. This change provides a natural experiment to which we apply a difference-in-differences (DID) framework to assess the causal effect of the AIPA on the acceleration of knowledge diffusion of US-based inventors relative to Japan-based inventors based on the timing of first USPTO applicant-originating citations to patents in our twin sample from each group.

Addressing our first question, we find that a significant acceleration of knowledge flow from Japan to the US occurred after the language barrier is removed, a result based purely on applicant citations. In order to assess whether inventors' preference for citing local prior art (unrelated to the removal of language barriers) affected our results, we

[1] See `https://purl.fdlp.gov/GPO/gpo28104`.

analyzed the effects of the AIPA for the group of the triadic patents—before the AIPA was implemented, the European Patent Office (EPO) published English applications 18 months after priority—finding no such acceleration for this group. Thus, language barriers drive our results from the pure twin sample. Further, we check whether examiner citations mimic the effects observed for applicant citations, which would be the case if the examiner citations carried information about knowledge flow between inventors, as much of the existing literature implicitly assumes. However, we find a slight *de*celeration that may be attributed to substitution effects due to US applicants getting access to, and citing, prior art that was previously behind the language barrier.

To address our second question, we conduct triple-difference and split-sample DID analyses on two groups of potential mediators of the acceleration effect. The first group, which we refer to as 'invention-specific' mediators, are related to properties of the knowledge, the quality of which is unknown before translation.We consider two primary invention-specific moderators: a patent's impact in the Japanese domestic market, and the recency of the antecedents of invention described in the patent. While the results are somewhat nuanced, we generally find that acceleration of knowledge across the language barrier was confined to high-quality patents, and was more pronounced for inventions that integrate relatively new knowledge on average.

The second group of variables, which we name 'appropriability' mediators, refer to the capability of the citing firm to appropriate the returns from its translations of the Japanese patent disclosures before their US publications. In this work, we consider two appropriability mediators: the R&D scale of the citing firm and whether the citing firm has past experience patenting in the Japanese market, both of which will help a firm to profit from its translation investments. We find that acceleration of knowledge diffusion from Japan is almost entirely confined to small firms and firms with minimal investment in the Japanese market.

The remainder of this work is structured as follows. Section 2 first covers the diverse set of background literature required to contextualise our approach and findings, then situates the current work relative to this literature and formulates hypotheses. Section 3 describes our data sources, cohort construction, and gives a detailed description of the variables of interest alongside the same for our controls. Section 4 gives an overview of our empirical methods while Section 5 presents the results of our analyses. Section 7 concludes and briefly discusses some implications of our results as well future directions.

# 2 Background and hypotheses

## 2.1 Literature review

Past work on knowledge diffusion in both international and domestic contexts is myriad and diverse. For our purposes, we focus on three distinct aspects of this literature while reserving for Section 2.2 the growing literature on the specific policy context that we exploit. We first cover the set of evidence that justify our use of patent citations as a measure of knowledge diffusion in the first place. We then review the use of patents and their citations for studying international knowledge diffusion and highlight the strikingly small subset of those works that additionally consider the role of language barriers in this context. Lastly, we briefly introduce some pertinent works on the heterogeneous effects of changes in the timing of technical disclosure on follow-on innovators and innovation more broadly.

### 2.1.1 Patent citations and knowledge diffusion

Much of the extant research that uses patent citations as an indicator of knowledge flow either relies on the fact that previous work has done so, and direct external validation exercises are exceedingly rare.[2] Indeed, historically, indirect 'elimination' methods have been popular means to identify knowledge flows. These methods attempt to control for any potential causes of citation that are not spillover-related, this allows one to infer that any signal that remains may reflect spillovers in some real sense. For example, when knowledge spillovers are anchored to geography (Audretsch and Feldman, 1996; Feldman, 1999; Audretsch and Feldman, 2004), industrial agglomeration patterns need to be removed, as incidental citations (i.e., those unrelated to knowledge spillovers) are more likely to be geographically localised when similar firms cluster together (Rosenthal and Strange, 2003). These elimination methods have largely arrived at similar results: there is usually some anomalous signal (e.g., citation localisation) remaining after controlling for incidental factors(Jaffe et al., 1993; Singh and Marx, 2013; Murata et al., 2014). While there are important caveats in the use of patent citations for measuring knowledge diffusion (Breschi and Lissoni, 2001; Alcacer and Gittelman, 2006; Kuhn et al., 2020), this large body of literature combined with the smaller-scale survey evidence suggests that, on average, patent citations remain a noisy indicator of real knowledge diffusion.

In their recent review, Belderbos and Mohnen (2020) consider knowledge spillovers in international contexts and note several broadly applicable pros and cons to using patent citations to measure them, two of which are of particular importance to the current work.

[2] The only survey, to our knowledge, that directly asks inventors about their knowledge inputs with reference to specific patent-patent citations is that of Jaffe et al. (2000), who found that while the information content of any given citation is relatively low, backward citations likely provide a noisy signal about aggregate knowledge flows.

First, not all inventions are patented and therefore any measure of spillovers using patent citations are prone to bias towards industries with a higher propensity to patent, and this may differ between countries. This leads to most work on spillovers being accompanied by some variation of the implicit caveat: "conditional on a technology being patented..." This remains true for the current work. Second, citations, even those submitted by the applicant, are an imperfect measure of knowledge spillovers (Kuhn et al., 2020). In this work, we attempt to mitigate this problem by restricting the citations to those more likely to reflect knowledge flows, and focus on changes in citation-derived statistics in response to the event of early disclosure, rather than using absolute volumes.

### 2.1.2 International knowledge diffusion and language barriers

There are a variety of mechanisms that allow inventors to access foreign knowledge, and a variety of ways to measure these flows and their implications. Because we rely on patent citations as a reasonable (if noisy) way of measuring these knowledge flows, we will briefly explore the most relevant findings of past literature that use citations in this way, with a focus on the US-Japan context where possible. For more general overviews of international knowledge diffusion, we refer readers to the excellent reviews of Branstetter (2000), Keller (2004), and Montobbio and Kataishi (2014).

A prominent class of models used to study international knowledge diffusion was implemented to full effect using patent citations in Jaffe and Trajtenberg (1996) and Jaffe and Trajtenberg (1999), building directly on earlier theoretical and empirical work (Caballero and Jaffe, 1993; Jaffe et al., 1993). These models consider flows of citations between 'cells' defined by time, technological field, and the location of the citing and cited inventors. Knowledge flow to a particular cell is then modelled as a combination of a knowledge diffusion process which increases the citations rate (which can also be represented by a cumulative advantage process (Higham et al., 2017)) and a knowledge obsolescence which causes the citation rate to drop as the technology ages. Notably, Jaffe and Trajtenberg (1999) observe that citations to, from, and within Japan are consistently outliers in the data. With respect to modal citation lags, within-Japan citations come much faster than any other intra-country citations, and citations from Japan also arrive faster than other cross-country citations. Indeed, US-based inventors were found to cite Japanese inventors faster than their compatriots. The relevance of Japanese technology to US inventors, at least historically, further justifies the use of the Japanese language as an important case study for examining the relationship between language barriers and knowledge diffusion. It has also been noted recently that, with respect to the knowledge diffusion that can be measured using patent citations, increases in the rate of international knowledge diffusion in the past three decades has been largely driven by citations to inventions with their origins in either US or Japan (Berkes et al., 2021). In that work, Japan's technological influence is particularly great from around 1995 to 2010—perfectly lining up with the introduction of the AIPA and years that followed.

There are very few past works that explicitly and empirically consider the impact of language barriers on international knowledge diffusion, and even fewer that do so using patent citations. When language is considered, it is usually in the form of a simple same-language dummy (e.g., MacGarvie, 2005) or relatively shallow analyses that finds that, indeed, language barriers do add friction to knowledge diffusion processes (e.g., Peri, 2005; Hegde et al., 2022). There are still fewer past works that consider mediating factors that may mitigate or exacerbate the impact these language barriers on knowledge spillovers.[3]

One early article, due to Keller (2001), attempts to dissect distinct knowledge spillover channels that lead to their attenuation with geographic distance. While it is found that trade patterns explain the majority of this affect, the author also find that foreign direct investment (FDI) and language barriers explain much of the remainder, in roughly equal parts. Peri (2005) came to a similar conclusion using a gravity-model approach. In two related works, restricted to the European context, Maurseth and Verspagen (2002) and Fischer et al. (2009) find that on top the strong effects of distance, borders, and technological compatibility between European regions, a shared language still has a significantly positive impact on the likelihood of a cross-region citation. Similarly, MacGarvie (2005) assesses some preconditions under which international knowledge diffusion occurs at a global scale. The author finds that cross-border citations increase when countries are geographically and technologically proximate, when there is no language barrier, and when there is high levels of foreign direct investment.

These results are somewhat surprising in retrospect. That is, in past studies of international knowledge spillovers in the context of technological innovation, so much effort has been spent untangling the FDI-spillover relationship, or the distance/border-spillover relationship, while the language-related effects that are found by everyone who considers them are usually considered an obvious friction that should simply be controlled for. In contrast, we view language differences as persistent impediment to knowledge diffusion that deserves to be studied in more detail.

Further, the fact that the current work only considers knowledge diffusion between Japan and the US context highlights an important distinction between this work and many others on the topic of international knowledge diffusion: these countries so distant that most geographical-distance effects are likely to be irrelevant to the probability of knowledge diffusion (c.v. Peri, 2005; Sin, 2018). Instead, our empirical approach rests primarily on the idea that patents are a useful source of information for inventors. While we acknowledge that this is not a settled debate, there is a slowly accumulating evidence base that this is true to some extent (Jaffe et al., 2000; Giuri et al., 2007; Ouellette, 2012;

[3]Of course, real language barriers are not binary variables but instead vary in size depending on a variety of interrelated factors (Schomaker and Zaheer, 2014; Brannen et al., 2017), including the linguistic distance between languages and the languages commonly spoken in any given country (Melitz, 2008; Melitz and Toubal, 2014).

Furman et al., 2021), and particularly for fast moving technology categories of the kind for which Japan was a world-leader during the time period we consider (Ouellette, 2017).

One recent work has made initial inroads into the language-spillover relationship by exploiting the advent of machine translation of Chinese patents into English by Google Patents (Büttner et al., 2022). Using Korean patents as a control group (which weren't translated until a later date), the authors find a significantly positive increase in forward citations to the treated documents following the change. Digging deeper into the mediators of this effect, the strongest effects were found for sectors in which China is very active, university patents (which are argued to be more readable and thus more amenable to machine translation), cited patents with figures, citing inventor teams containing no Asian names, and small follow-on innovators, while also finding no effect on examiner citations.

While the event, measured outcome, and geographical context are different than those we consider in the current work, we note two more subtle, but important, distinctions between Büttner et al. (2022) and the current work. First, we focus on pre-grant publications in English as a disclosure event that facilitates knowledge diffusion, which allows us to measure the causal effect of language barriers on the *speed* of new knowledge diffusion by looking at the sources of first citations. In contrast, Büttner et al. (2022) consider translations of granted patents up to 18 years old, which limits the causal inference to the extent of *additional* knowledge diffusion after machine translation (by which time the translated patents may be obsolete). Moreover, Büttner et al. (2022) compare the changes in levels of the forward citation in the 5 years before and after the event (10 years in total), which makes it difficult to separate the effect of machine translation from concurrent effects, in particular, the improvement of invention quality by Chinese firms.

### 2.1.3 Accelerated technical disclosure

The disclosure function of technical information is considered one of the core functions of the patent system and has been studied from a range of perspectives, theoretical and empirical (Denicolò and Franzoni, 2003; Murray and O'Mahony, 2007; Seymore, 2009; Hall et al., 2014). The timing of technical disclosure has received much attention in recent years, though before the AIPA there was little empirical data available to study what happens when new and relevant technical information becomes accessible to inventors in a more timely fashion. As this work also relies on the AIPA, we provide an overview of the related empirical work in the next section. In the most relevant theoretical work, Aoki and Prusa (1996) model expected innovation outcomes between the Japanese pre-grant publication system and US pre-grant secrecy. They argue that pre-grant publication allows for more coordination between patenting firms, which results in less duplicative R&D but smaller inventive steps. After all, without knowledge of competitors' technology, firms may feel that they should make larger inventive steps and produce innovations of higher quality in order to remain competitive, but there is a higher risk of duplication. In the Japanese system, where firms had more knowledge of competitors' technology, relative technological positions were more transparent and firms could adjust their R&D spending accordingly.

## 2.2 The AIPA as a natural experiment

In this work, we make use of a change in US patent policy called the American Inventor's Protection Act of 1999, henceforth referred to as the AIPA. While the AIPA made several significant changes to US patent law, the change of interest is the introduction of pre-grant publication of patent applications to the USPTO 18 months after their first filing, which was implemented on November 29, 2000. This change brought the United States in line with all other major patent offices at the time, including the JPO, which already published applications after 18 months and presents a quasi-experiment that we are able to exploit. Notably, we are able to use the fact that patent families that were only ever filed in both the JPO and USPTO have an interesting property: before the AIPA they were published in Japanese only, but after the AIPA, they are published in both English and Japanese at approximately the same time. That is, before the AIPA, equivalent technical knowledge was disclosed in Japanese before it was disclosed in English, and often at a significant lag. As such, changes in the speed of knowledge diffusion immediately following the implementation of the AIPA may give us direct insight into the effect of language barriers on that diffusion more generally.

We are certainly not the first to use the implementation of the AIPA as an event study for learning about the relationship between the timing of accelerated technical disclosure and innovation, broadly defined. Indeed, before the AIPA even took effect, researchers began to consider its potential effects on patenting strategy and incentives to carry out R&D (Johnson et al., 2003). More recent studies have been able to use the AIPA to assess the real impacts of accelerated publication. A first look at the longer-term consequences of the AIPA on patenting strategy was conducted by Graham and Hegde (2015). In particular, they look at the determinants of a firm's choice to keep a patent application secret (which is an option if there are no parallel applications abroad), along several dimensions.

Despite none of the patents in the cohorts studied in the current work being eligible for secrecy, several findings from Graham and Hegde (2015) emerge as potentially relevant for understanding what drives firms' reactions to the option of early-disclosure. First, they find that secrecy-eligible patent applications filed by small firms are just as likely to be kept secret as those filed by large firms (about one in six). Second, they find that patents that are kept secret have unusually small scope. Third, they find that large firms pay maintenance fees for longer when the patent was secret, while small firms do the opposite. That is, it seems that the secrecy kept by small firms through the avoidance of pre-grant patent publication is not valued highly. Finally, patents issuing from secrecy gain significantly less citations than their published counterparts, across all firm sizes.

Subsequent work has added some colour to the above findings but, by and large, a surface-level interpretation supports the idea that firms generally benefit from early disclosure of patent applications. For example, one way an economic benefit is conveyed is through the timing of patent licensing, which was been found to have significantly accelerated in the biomedical industry and of particular importance to small firms (Hegde and Luo, 2018). Other work looking at the effect of the AIPA on capital markets has found that earlier disclosure plays an important role in the diffusion of information to shareholders (and would-be investors) of innovating firms, as these parties can obtain more timely information about a firm's recent R&D activities and outcomes (Hegde et al., 2018). Firm-level heterogeneities in the impact of the AIPA can also arise based on the timing of technical disclosure that a firm was accustomed to pre-AIPA. Kim and Valentine (2021) consider the effect of changes in the levels of knowledge 'spill-ins' and 'spill-outs' that a firm experiences when the AIPA is implemented. The authors find that firms that historically experienced (or induced) longer grant lags for their patents relative to their competitors invested less in R&D after the AIPA but those firms that experienced shorter granted lags increased their R&D spending, concomitantly affecting their respective patenting rates. These effects were found to be stronger for industries where patents are an important source of technical information for follow-on innovators (Ouellette, 2012).

In one of the most empirically relevant articles for the purposes of the current work, Hegde et al. (2022) conduct a broad swath of analyses on the impact of the AIPA on follow-on innovation. This work implements a DID framework around the AIPA (using European patents as controls) and finds that patents receive more citations; the first, third, fifth, and seventh citations occur faster; technologically similarity increases between dissimilar patents and reduces between highly similar patents; and patents became narrower but less likely to be abandoned. The authors additionally conduct the former two sets of analyses on a set of Japan-US patent equivalents, which reveals that the change in the speed and extent of citation flows was even greater for this cohort relative to the cohort of Europe-US equivalents. They also point out the possibility that a language barrier may explain the difference. Based on the assumptions that applicants and examiners are more likely to search locally for prior art (disclosure in foreign patent office is less used) and that examiner citations can indicate knowledge flow to inventors they interpret these findings as showing that the AIPA has brought significantly more knowledge diffusion.

Our work, while closely related to their work, differs fundamentally in that we do not depend on the above two assumptions for identifying the policy effect, instead assesses their validity directly. First, we do not use examiner citations since examiners often cite early prior art which the inventors cannot observe (Okada and Nagaoka, 2020). The tendency for examiners to cite local prior art in their examination would identify an artificial acceleration of knowledge diffusion due to the AIPA. Second, we identify the effect of the AIPA on knowledge flow by comparing the behaviors of the US inventors and the Japanese inventors, using their US applicant citations. The US inventors are the treatment group and the Japanese inventors are the control group, and only the latter are likely to understand pre-grant Japanese-language publications from the JPO. That is, we rely exclusively on USPTO applicant citations, which avoids the assumption that citations made in different patent offices are comparable.[4] Although a potentially stronger preference for US inventors to cite the US prior art than that for the Japanese inventors can introduce a bias into this kind of analysis, we assess this potential bias and control for it in our estimation.

Relatedly, Baruffaldi and Simeth (2020) measure the impacts of the AIPA on knowledge diffusion within the United States, based on a similar identification assumption as Hegde et al. (2022). While the primary results in this work are derived from regression discontinuity analyses, a supplementary difference-in-differences analysis is also conducted with a group of patent families filed only in Europe serving as the control group; the authors obtain results consistent with those of Hegde et al. (2022) for this cohort. They also find that early disclosure has a larger impact on 'discrete' technologies than on 'complex' technologies (see, e.g., Cohen et al., 2000; Von Graevenitz et al., 2011), and for technologically proximate inventors who can quickly absorb the new information (particularly for those who use Patent Attorney Service Firms to improve prior art search). The increased knowledge diffusion doesn't appear to have a significant (domestic) geographical dimension.

In research that explicitly models citation lag distributions (à la Caballero and Jaffe (1993)) before and after the AIPA and focusing only on applicant citations, Okada and Nagaoka (2020) consider a raft of factors influencing the timing of the inflection point of the distribution. This inflection point is directly determined by the respective properties of technological diffusion and obsolescence functions that are reflected in patent citations. The authors consider applicant citations to a purely domestic cohort (patents with no foreign equivalents and, therefore, with an option to opt out of pre-grant publication) and observe large technological heterogeneity in the impact of the AIPA, moderated by the relative length of pre-AIPA examination lags and the general speed of the technological frontier. Notably, the average effect appears to be primarily driven by acceleration of

[4] In fact, while our estimations based on triadic patents suggest that disclosures at the EPO are used by US-based inventors as a knowledge source, it is important to make a conceptual distinction between inventors' preference to cite local prior art and the constraints on their search strategies (which are generally unobservable). Also, we note that citations across offices are not necessarily comparable, and generally have different generation mechanisms (Bacchiocchi and Montobbio, 2010; Higham and Yoshioka Kobayashi, 2022).

knowledge diffusion in a select few technological areas, with most technologies seeing significantly lower acceleration.

Lastly, one of the stated goals of the AIPA was to reduce wasteful R&D duplication.[5] Lück et al. (2020) observe that this appears to have been achieved by showing that its implementation reduced the number of patent applications blocked by prior art of which the applicant was likely unaware. These effects were also found at the EPO — after all, patents granted by the USPTO also constitute prior art for applications filed at the EPO.

## 2.3 Hypotheses to be tested

We expect that the introduction of pre-grant publications to the US patent system significantly accelerated US inventors' access and use of Japan-originating technical knowledge, relative to Japanese inventors. Thus, we have the following Hypothesis 1 in terms of the timing of citations:

> H1: The introduction of pre-grant publications to the US patent system significantly reduced the citation lag for citations from US inventors, relative to Japanese inventors, reflecting accelerated knowledge flow to the US inventors.

In testing the above Hypothesis, we will conduct DID estimation, based on a pure twin sample of patent families filed by Japanese firms only at the JPO and the USPTO, to directly assess the effect of language barriers on knowledge diffusion. Before the AIPA, all of these patent families were published in Japanese 18 months after their first filing, but the English versions were not available until their USPTO grant.

In addition, we assess and control for the other effects of USPTO pre-publication on citation lag that are not driven by knowledge diffusion. Several mechanisms could lead to such a shock to the citation lag. For example, the fact that an application is also filed in the United States signals that the invention is important in the US market, which is very likely to increase the frequency of citation from US patents. In addition, under US patent law, applicants are obliged to disclose known prior art to the Patent Office. If there exist both applications published only in Japan and applications published in both Japan and the US, and either could serve as prior art, the latter is likely to be chosen since additional explanation or a translation is required when submitting prior art in a language other than English.[6] Therefore, the application published in Japan and the U.S. has a higher probability of being cited. Both effects apply not only to Japan-based inventors but also to US-based inventors.

For assessing and controlling such potential biases, we use triadic patent families, which we define as those patents with equivalents filed in all three triadic offices (USPTO,

[5] H. Rept. 106-287 Part 1, see `https://purl.fdlp.gov/GPO/gpo28104`.

[6] Section 609, CFR 1.98 (a)(3) of the USPTO Manual of Patent Examining Procedure `https://www.uspto.gov/web/offices/pac/mpep/s609.html`.

JPO, and EPO) and granted at least at the USPTO (we cannot observe rejected applications at the USPTO before the AIPA). The only essential difference between the pure twin patents and the triadic patents is that there is the EPO publication in English, published 18 months after first filing, in the latter case.

If the EPO pre-grant publication is as good a knowledge source as the USPTO pre-grant publication for the US-based inventors, there will be no acceleration of knowledge flow due to the AIPA. As such, a DID estimation based on the triadic patents can identify the above biases, and the differencing the two DID estimates from the pure twins and from the triadic patents (i.e., a triple difference specification) can identify the true effect of the AIPA on the speed of knowledge diffusion. On the other hand, if the EPO pre-grant publication is ignored by US inventors (their search is completely local), the two DID results will be identical and the triple difference would provide us with an underestimate of the effect on the speed of knowledge diffusion. That is, in this extreme case where US-based inventors were not looking at any foreign prior art previously, even those published in English, the acceleration of citation observed for triadic patents consists of both a knowledge flow effect of the AIPA and the effect of a preference for citing local prior art, as in the case for twin patents. At the other extreme, if EPO publications were as useful and accessible as a USPTO publication for the US-based inventors, the acceleration captures only the effect of the preference for citing local prior art. Thus, the effect size estimated from a triple difference specification (the difference of the twin patents and the triadic patents) becomes significant if the US inventors only search local prior art.

Our second Hypothesis assesses whether inventors search beyond local prior art:

> H2: If the US-based inventors search significantly beyond local publications, the implementation of the AIPA had a much larger acceleration effect on the twin patent cohort that on the triadic patent cohort.

If H2 is confirmed, we will implement the triple difference estimation described above to check for the effect of the AIPA on knowledge diffusion.

At the USPTO, examiners often add further citations to those disclosed by the inventors. These additions are made for a variety of reasons, including justifications for rejection when an application is deemed non-novel or too obvious (Higham and Yoshioka Kobayashi, 2022). There is even mixed evidence that examiners could make citations which capture missing knowledge flows that were omitted (Lampe, 2012; Kuhn et al., 2021). Such omissions may not be intentional—inventors are not always aware of the entire technological landscape surrounding their inventions. This is especially true for new prior art, since such prior art remains secret until their disclosure. The fact that examiner citations tend to come sooner than applicant citations (see, e.g., Okada and Nagaoka, 2020), particularly once super-citing patents are removed, indicate that this latter effect is significant. A primary reason that examiner citations come 'earlier' is because they are usually added a long time after the citing application was filed, and relevant prior art has been disclosed during this period between filing and examination (or, further, between initial invention development and examination). This leads us to our third hypothesis:

> H3: Because examiner citations to recent prior art are unlikely to capture knowledge flows between inventors, and are usually added a long time after the initial filing, we will not observe a strong acceleration of examiner citations from US-based applications relative to Japan-originating inventions upon the implementation of the AIPA.

Lastly, language barriers are less absolute than secrecy, since they can be largely overcome through a variety of means, such as translation or hiring multilingual researchers. Thus, the effect of language barriers on knowledge diffusion is heterogeneous among innovators, and depends on whether investments for overcoming the barriers are profitable. We derive two hypotheses on these heterogeneities, starting with a simple conceptual model describing the incentives and ability of firms to undertake such investments.

We first assume that the economic effect of translating a foreign language patent application is proportional to the product of the size of complementary asset of the firm $i$, $S_i$, and the quality (determined by attributes such as technological importance, relevance, and timeliness) of the knowledge contained within a foreign patent $j$, $z_j$, which is stochastic and unknown before translation. Given the cost, $c_t$, of translation and the cost $c_u$ of incorporating the new knowledge into a subsequent invention, firm $i$ will invest in the translation of a foreign language patent application if and only if the expected return is positive, that is, if $v_{i,j} = (\alpha S_i z_j - c_u) - c_t > 0$. For simplicity, let us assume that the distribution of $z_j$ is independent of $S_i$. Then, the above equation implies that a firm with large complementary assets will translate all relevant foreign language translations, while a firm with small complementary assets translate none of them. This leads to two important and testable implications.

A firm with large complementary assets (large firms and firms which have special assets to use the foreign technology) are likely to translate all relevant foreign language patent applications, so that the additional effect of the AIPA-driven 'translation' of these applications through pre-grant publication does not exist for such a firm. Larger firms may also have access to more knowledge channels or other means of accessing the knowledge generated in the Japanese market (see, e.g., Almeida, 1996), which are complementary to their translation incentive. Following this logic, we can form our fourth hypothesis:

> H4: The AIPA had the largest impact on firms that were small and had few resources to access Japan-originating technical knowledge before the policy change, or were less likely to be directly involved in or otherwise involved to the Japanese innovation ecosystem.

On the other hand, for small firms which do not translate the foreign language patent applications on their own, we expect the additional effect of AIPA-driven public translation will increase with patent quality. That is, pre-AIPA, a high-quality patent may only have been 'discovered' (and subsequently cited) via individual translation by chance whereas, post-AIPA, this inefficiency is removed and high-quality inventions (those with $\alpha S_i z_j - c_u > 0$) are cited quickly after their initial disclosure. As such, we construct our fifth hypothesis:

> H5: The AIPA had the largest effect on citation timing for high-quality patents.

# 3 Data

## 3.1 Sources

To form our base cohort of patent families, their citations, and most of their metadata, we use the PATSTAT database. This includes WIPO technology fields as defined by Schmoch (2008), and both DOCDB and INPADOC family identifiers. For an overview of these family types, we direct readers to the excellent review of Martinez (2010).

To this set, we add harmonised inventor and assignee data for both citing and cited patents from the USPTO's PatentsView platform.[7] Because all cited families have a US equivalent (by construction), and all citing patents are USPTO patents themselves, we can use this metadata to calculate firm characteristics and obtain inventor locations (derived from their address as listed the US patent grant).[8]

## 3.2 Cohort construction

We first form a core of cited patents that stretch across the policy change, so that a difference-in-differences (DID) model can be used to study the impact of the policy change. This sample consists of DOCDB (simple) patent families that contain equivalents that were only ever applied for in both Japan and the US.[9] As we are studying knowledge flow from Japan, these cited families only list Japan-based inventors on the front page of their US equivalent(s). For consistency across the policy change, the US equivalents within the included families must be granted, but this restriction is not applied to Japanese equivalents, as non-granted applications at the JPO are public both before and after the policy change. To mitigate the effects of mid- to long-term trends in application and citation behaviours, we only include patent families that filed their US

[7] http://www.patentsview.org/

[8] The harmonisation process used by the USPTO is described in detail at `https://patentsview.org/disambiguation`.

[9] This sample accounts for around 2.4% of all Japanese patent applications but 5.9% of the patent applications in the top 5% with respect to inventor citations received.

equivalent 182 days (26 weeks) either side of the policy change. This set of potentially cited patents constitutes our 'focal cohort'.

We then consider citations to this sample from granted US patents which are identified as having purely US-based inventors or purely Japan-based inventors. For our baseline analyses, we only include applicant-added citations to the families in our cited sample, though we conduct a comparison using examiner-added citations that otherwise fit the same criteria. Firm-level self-citations (with assignees identified by the USPTO) are removed, along with any remaining within-family citations at the INPADOC (extended) family level (of which there are very few) and any duplicated between-family citations (resulting from, for example, citations copied from patent to child applications). We also include inter-office citations to our focal cohort; that is, citations on USPTO patents made to the JPO equivalents of the families in this cohort. Finally, we restrict the analysis to those firms that are able to be treated in the first place—citing firms must display evidence of R&D activities before the policy change. This is enforced by ensuring citing firms all have at least one patent granted by the USPTO before the policy change.

Priority date-to-priority date citation lags are calculated for each citing-cited family pair, then any citations that occur at a lag of more than 10 years after the cited patent's priority date are excluded to ensure all patents have the same amount of time to receive citations.[10] As we are only comparing the timing of first citations from Japan- and US-based inventors, we remove all but the first citation from each source to each cited patent. The sample size statistics before and after this process are displayed in Table 1. This process is conducted after any additional restrictions are placed on the cohort that depend on the specific analysis being conducted. These restrictions are explained in detail in the relevant results subsection.

In our main sample, we are left with cited patents that were only ever filed at the JPO and USPTO, and having received USPTO applicant citations from granted patents filed by both Japan-based and US-based inventor teams. Both of these teams had access to the pre-grant application in Japanese before and after the policy change, and both were able to access the English version after the policy change. By restricting our analysis to citations made by USPTO patents, we do not have to concern ourselves with any legal differences between patent systems that might create differences in the legal meaning and strategic use of citations made on patents in different jurisdictions (Higham and Yoshioka Kobayashi, 2022).

To check for effects induced by preferences for citing the prior art of the local patent office, we also consider a cohort of triadic patents filed by Japan-based inventors, for which the EPO application was published in English before and after the policy change.

[10]The priority date of a patent is the date when it was first filed at any patent office party to the Paris Convention for the Protection of Industrial Property (1883). Further, because this date can be used as an effective filing date for the purposes of subsequent filings at other patent offices (or continuations at the same office), only one priority dates exists for each DOCDB patent family.

Table 1: Baseline sample sizes.

| Citing Country | Total citations Pre | Total citations Post | First citations Pre | First citations Post |
|---|---|---|---|---|
| US | 1693 | 3081 | 565 | 820 |
| | (565) | (820) | (565) | (820) |
| Japan | 1892 | 3341 | 565 | 820 |
| | (565) | (820) | (565) | (820) |

Notes: Unbracketed values are citation counts, while bracketed values indicate number of distinct cited patents. To be consistent with analyses later in this work, citing patents making over 100 citations are removed from these calculations.

That is, the only change potentially-citing US-based inventors experienced with respect to triadic patents was that the pre-grant publication was now published in English by the EPO, in addition to the Japanese-language equivalent published by the JPO. This comparison will allow us to assess how a potentially higher preference of the US inventors to cite local (USPTO-published) prior art affects the impact of the AIPA on citation lags at the USPTO, even when language barriers are not in place (due to the EPO pre-grant publication).

We acknowledge that our final cohorts and citation sets are quite small by the standard of patent data-based analyses, even compared those which use a related empirical strategy within the AIPA context (Hegde et al., 2022). This is done to both avoid complex empirical strategies that require more built-in assumptions, and to obtain as clear a signal as possible for observing any language-barrier effects.

## 3.3 Variables

In this section we introduce the variables that we use to explore the effects of the language barrier on knowledge diffusion. All variables that follow are normalised by their standard deviation, such that coefficients in all regressions reflect the effect on the dependent variable of a one standard deviation change in each independent variable. We categorise all independent variables introduced in this section by whether they are related to the intrinsic properties of an invention or its appropriability by follow-on innovators. Note that invention-specific variables, as they are properties of the cited patents, do not change when all but the first citations are removed, but variables related to appropriability do change during this process. The summary statistics of the former variables are presented in Table 2, while those of the latter are presented in Table 3. The dependent variable, citation lag, is included in the latter table, as are the mediators described in detail in Section 6.

### 3.3.1 Dependent variable

We use the time delay to the first citation as our dependent variable. Earlier disclosure of knowledge will naturally lead to earlier follow- on innovation, when the cited patents are a source of knowledge. Earlier disclosure of knowledge can also enable an early recognition of duplications. The time delay to the first citation is a direct measure to capture these effects of early disclosure on cumulative innovations. It is at these early times that we expect language barriers to play the most significant role.

From a practical standpoint, using first citations allows us more easily test for mediating factors of the language barrier on citation lag at the citing-family level. This allows us to directly integrate the properties of citing patents (including characteristics of the citing inventors and firms) into our analyses and control for factors that may affect citation lags at the citation level, such as the size of the firm assigned to the citing patent. This is not possible when one uses a measure like citation counts, which necessarily have to depend on cited patent-level statistics to control for these effects. Lastly, first citations are non-parametric (c.v. Okada and Nagaoka, 2020) and do not diminish the sample to a size which causes nuanced effects to be swallowed by noise (as would be the case for the second or other subsequent citations).[11]

Specifically, our dependent variable is calculated as the logarithm of the lag measured from 18 months after the priority date of a patent family in our focal cohort to the priority date of the first family that cites it, by geographical source. Priority dates are as close as possible to the invention date, but have little reason to think that the distribution of time lags between hypothetical invention dates and the priority dates are different for cited and citing patents (though any systematic level differences will not affect our DID estimates). As such, the priority-to-priority lag is likely to be close to the invention-to-invention lag, should something so concrete even exist in the latter case. This lag is then offset by an 18 month period during which the cited patent application is still secret at the JPO for all cited families. Therefore, we believe our dependent variable can, on average, accurately measure changes in the speed of follow-on innovation (conditional on follow-on innovators applying for, and eventually being granted, a patent).

As shown in Table 2, most first citations to our cohort occur within four years of the standard 18-month pre-grant publication. However, some patents receive their first citation (from either country) after much longer times. To mitigate the influence of these longer lags, and to emphasise a proportional change in lag rather than an absolute change (which may be highly technology or country-specific) we use log-transformed citation lag, with the 18 month shift of the lag distribution resulting in a minimum lag of 0 days before the transformation. In effect, this 'starts the clock' on the timing of incoming citations at the disclosure date of the cited patent and the logarithmic transformation

[11] We note that these particular benefits are specific to our preferred empirical strategy.

effectively places a lower weight on the same absolute difference in citation arrival time as a particular invention ages. Empirically, this transformation results in regression coefficients that should be interpreted as the percentage change in the lag that can be attributed to a one-standard deviation change in the associated independent variable.

#### 3.3.2 DID-relevant independent variables

For identification of a treatment effect in a difference-in-differences framework, we first specify the treated group and the timing of the treatment. To restate, we are interested in the diffusion of knowledge produced in Japan among the US inventors, and how this changed when the AIPA accelerated the publication of this knowledge in the English language. US-based inventors are the treatment group and Japan-based inventors are the control group within our framework. Citation lags from US-based inventors to patents filed after the implementation of the AIPA are, therefore, our treated observations, while those from Japan-based inventors are our control observations. Note that because we condition our sample on having a Japanese equivalent (and therefore a Japanese pre-grant publication), these cited patents will often have a priority date that is earlier than their USPTO filing date. However, both the JPO and USPTO publish pre-grant publications 18 months after their priority date (not necessarily the USPTO filing date).

The timing of the treatment is the date that the pre-grant publication provision of the AIPA came into effect: November 29, 2000. To include more detailed time fixed-effects than simple before change/after change dummies, we split this one-year cohort into week-long windows. Mechanically, the date of the policy change is set as the start of a window despite the fact that it was a Wednesday. Our time fixed-effects are then week-fixed effects, though we note that the treatment effects that we observe are not sensitive to the size of the time window.

### 3.4 Controls

We include both invention-specific and appropriability controls in all analyses, to control for certain properties of both cited and citing patents. Any potential controls that might plausibly be endogenous to the policy change, such as the technological distance between citing and cited patents (Hegde et al., 2022), are not considered. Log-transformations are conducted when we expect a diminishing marginal effect as the variable becomes larger.

**Invention-specific Controls**

*Originality.* In past literature, the number and distribution of different fields that a patent draws on has been used as a basic indicator of originality (Trajtenberg et al., 1997; Hall et al., 2001). When originality is related to the radical-ness or cumulative-ness of a new invention, as appears to be the case (Baron and Delcamp, 2012; Verhoeven et al., 2016),

Table 2: Summary statistics: Invention-specific variables

| | All focal families | | | |
|---|---|---|---|---|
| | Mean | SD | Median | IQR |
| ***Controls*** | | | | |
| Originality | .564 | .261 | .64 | .331 |
| Science dependence | .143 | .722 | 0 | 0 |
| ***Mediators*** | | | | |
| JPO citations | 2.38 | 13.4 | 1 | 2 |
| Field pace (days) | 1929 | 1146 | 1685 | 1287 |
| Field temporal spread (days) | 917 | 866 | 665 | 916 |
| Observations | 1385 | | | |

Notes: Variables characterising the focal cohort. Notes: SD=Standard deviation; IQR=inter-quartile range. All variables constructed as described in main body. To be consistent with analyses later in this work, citing patents making over 100 citations are removed from these calculations.

we may reasonably expect this property to affect the time it takes to build on (and cite) a given invention. Originality is calculated as a modified Herfindahl–Hirschman index as first described by Higham et al. (2019) wherein all technological categories assigned to each cited patent are considered in the calculation.[12]

*Science dependence.* With a similar intuition as for originality, patents that have a close relationship to published scientific articles may reflect their novelty or their distance from the scientific 'frontier', which could affect how difficult they are to build on or around. The relationship between this property and citation lag, in turn, may be moderated by the industry in which the invention was produced or the discreteness of the technology (Corsino et al., 2019). We calculated science dependence as the log-transformed count of all citation a patent makes to scientific articles, from both their specification and in their list of non-patent literature citations (Marx and Fuegi, 2020, 2022).

*Primary technological fields of cited families.* These are included as fixed effects at the primary WIPO family level (Schmoch, 2008).

### Citing-patent Controls

*Citing inventors.* A large number of inventors on a citing patent may indicate both the size of the investment made by the assignee in the invention process as well as the size of the knowledge base on which the patent is built (assuming inventors within a team generally have non-redundant knowledge bases to draw from). This variable is the

[12]For this purpose, we use 3-digit Cooperative Patent Classification codes.

Table 3: Summary statistics: Citing-patent variables

| | All citations | | | | First citation pairs | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Median | IQR | Mean | SD | Median | IQR |
| ***Dependent variable*** | | | | | | | | |
| Citation lag (days) | 1919 | 854 | 1848 | 1436 | 1607 | 815 | 1401 | 1239 |
| ***Controls*** | | | | | | | | |
| Citing inventors | 2.69 | 1.79 | 2 | 3 | 2.65 | 1.81 | 2 | 2 |
| ***Mediators*** | | | | | | | | |
| R&D scale | 880 | 1268 | 279 | 1278 | 888 | 1275 | 277 | 1283 |
| Observations | 10007 | | | | 2770 | | | |

Notes: Variables characterising the forward citations to the focal cohort (appropriability view). Notes: SD=Standard deviation; IQR=inter-quartile range. Variables constructed as described in main body. To be consistent with analyses later in this work, citing patents making over 100 citations are removed from these calculations.

log-transformed count of inventors on each citing patent.

*Primary technological fields of citing families.* These are included as fixed effects at the primary WIPO family level (Schmoch, 2008).

# 4 Methods

While the design of treatment and control groups are different, our basic empirical approaches are similar to that of previous work (Hegde et al., 2022; Baruffaldi and Simeth, 2020; Büttner et al., 2022), in that we implement a relatively basic difference-in-differences model to measure the effect of a potential change in a citation-based measure to a focal cohort of patent, induced by an exogenous change in accessibility to patent information. However, we note that only (Hegde et al., 2022) does so at the citation level (the advantages of which are discussed in Section 3.3.1).

The baseline DID model, where the dependent variable is the citation lag (measured as described in Section 3.3.1) between citing patent family $j$ to an earlier, Japan-originating, patent family $i$ that is part of our focal cohort, can be written as

$$\begin{aligned} Lag_{j\to i} &= \alpha + \beta_1\, US_j + \beta_2\, Time_i + \beta_3\, US_j \times postAIPA_i \\ &+ \boldsymbol{\gamma}\cdot \boldsymbol{u_i} + \boldsymbol{\lambda}\cdot \boldsymbol{v_j} + \epsilon \ \ , \end{aligned} \tag{1}$$

where $US_j$ and $Time_i$ are citation source (a dummy equal one for US-based inventors) and time fixed effects for the focal cohort, respectively. The interaction term $US_j \times postAIPA_i$ is the variable of interest, which equals one for the citations coming from US-based inventors to patent families with priority date after the policy change—the treated group. $\boldsymbol{u_i}$ and $\boldsymbol{v_j}$ are vectors of control variables or technology-field fixed effects for the cited and citing families, respectively.

Equation 1 can then be augmented to a triple-difference model with a language-barrier mediator of interest $M$:

$$
\begin{aligned}
Lag_{j\to i} = \alpha &+ \beta_1\, US_j + \beta_2\, Time_i + \beta_3\, US_j \times postAIPA_i \\
&+ \beta_4\, M + \beta_5\, M \times postAIPA + \beta_6\, M \times US \\
&+ \beta_7\, M \times US \times postAIPA \\
&+ \boldsymbol{\gamma} \cdot \boldsymbol{u_i} + \boldsymbol{\lambda} \cdot \boldsymbol{v_j} + \epsilon \quad . \qquad (2)
\end{aligned}
$$

In this specification, $\beta_7$ captures the effect of the language barrier on treated inventors that can be attributed to $M$, which can represent the size of the citing firm, some measure of underlying technological importance of the cited family, or any other variable of interest. At this stage, we do not restrict $M$ to a particular aspect of the citation (e.g., invention-specific vs. appropriability), because any metadata relating to the citation could reasonably mediate knowledge flow across language barriers. Because treatment is assigned at the citing-country level rather than at random, standard errors for all regressions are clustered at this level.

We also conduct split-sample cross-checks for all mediators. These tests are helpful because they give more readily interpretable view of the mediators' effect on knowledge diffusion across the language barrier. To conduct these tests, the sample is first split into two subsets (above and below median in the case of quantifiable variables, binary for categorical variables) *before* removing all but the first citation from each source. We then run the baseline regression described in Equation 1 on each subset and include the mediator itself as a covariate to account for any remaining variation along this dimension within the subset (for quantifiable variables). The relevant comparison is then between the treatment effects for each subset, which can be contrasted with $\beta_7$ from the triple difference estimation to check whether the results are consistent.

Lastly, we run straightforward tests of the parallel-trend hypothesis for each DID regression and note specifications for which non-parallel pre-event trends are significant. These tests are conducted by augmenting the DID model with a linear time trend multiplied with Heaviside step functions that indicate whether an observation is part of the pre-policy period, as well as dummies indicating whether the citing inventors are US-based. A Wald test is then conducted using the coefficient and standard error of this trend; the resulting p-value is calculated. As such, a significant result for this test indicates a significant difference between the linear pre-trend of the control group (Japan-based) relative to the case group (US-based), and the average effect on the treated is likely to be biased.

# 5 Baseline Results and Robustness Checks

## 5.1 Baseline regression

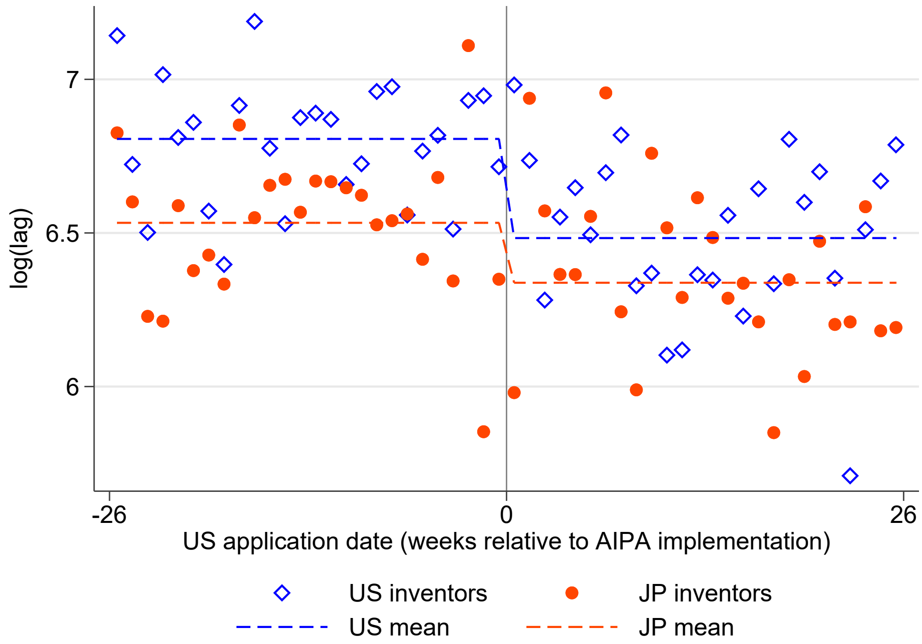
Before any formal analysis is conducted, one important issue related to citation timing must be addressed. In recent times, Kuhn et al. (2020) have noted that the number of applicant citations made by any particular patent has been increasing, apparently for strategic purposes, and a new class of 'super-citing' patents that cite hundreds or thousands of patents as prior art has spawned. The citations made by these patents often have little to do with knowledge diffusion, which has resulted in the decreasing quality of citations as a metric for capturing such phenomena (Kuhn et al., 2020). We address this problem by introducing a threshold that excludes citing patents that make more than a certain number of citations. To do this, we run our DID regressions on several sets of citations with different thresholds applied, as described in Appendix A. We find that restricting our analysis to citing patents that make a maximum of 100 backward citations provides a nice balance between the information content of included citations and ensuring the parallel trend assumption holds. For purposes of comparison and consistency, this threshold is applied to all other regressions in this work.

Before running the formal regression, we plot the raw average data in a way that will assist our interpretation. Figure 1 provides the graphical representation of the average time to first citations (after applying the threshold above) for the cohorts before and after implementation of the AIPA (see Appendix B for construction details of all such figures). First, average (first) citation lags drop for all citation types after the policy change, even for Japan-based inventors, to whom we expect no acceleration of knowledge flows after removal of the language barrier. These findings show clearly that the inventors prefer to cite local (USPTO) prior art when available and such preference has little to do with knowledge flows induced by the AIPA. Since our estimations below are based on a DID analysis, the relevant point is whether these effects are stronger for Japanese-based or US-based inventors, which is addressed in detail in Section 5.2.

We also observe that the first citations from US-based inventors consistently arrive later than their Japan-based counterparts. This observation is entirely consistent with prior work studying international knowledge flows in the Japanese context (Jaffe and Trajtenberg, 1999; Jang et al., 2009) and well-established facts about the dynamics of geographical knowledge localisation more generally (Thompson, 2006; Li, 2014).

Our formal regression results obtained by implementing the DID model described in Section 4 are displayed in Column (1) of Table 4. We find the baseline effect size of the policy change is a 12.9% decline in time to first citation (where time starts 18 months after priority) for US-based inventors relative to Japan-based inventors. Remembering that all patents in our focal cohort, even those filed before the AIPA was enacted, are

Figure 1: Average time to first citation for focal cohort




Notes: Solid lines indicate the (log-transformed) lag to first citation, by source, averaged by week. Dashed lines indicate the same, averaged over the 26-week periods before and after the AIPA came into effect.

published in Japanese 18 months after their priority date, we interpret this result as the average acceleration of knowledge diffusion that is caused by the removal of the language barrier between US- and Japan-based inventors. Lastly, in this baseline regression, we note that the language barrier appears to account for almost one half of the relative delay in knowledge diffusion from Japan to the United States. In sum, we find good preliminary support for Hypothesis 1, and this evidence is built on in Section 5.2.

## 5.2 Triadic patents and baseline robustness

Annual patenting trends and the impact of the AIPA on citation lag, unrelated to knowledge diffusion, office-level localisation of citations may effect Japan- and US-based inventors differently, and independently of effects due to the language barrier.

Annual patenting trends and patent office-level localisation of citations may moderate measurement of the impact of the AIPA differently for Japan- and US-based inventors, and independently of effects due to the language barrier or knowledge diffusion more generally. To address these potential distortions of our baseline results in identifying knowledge diffusion, we incorporate into our analysis a set of citations to each group for a sample of triadic patents. In this sample, there are EPO pre-grant publications in English in addition to the JPO pre-grant publications in Japanese associated with the triadic patents granted in the US. That is, the USPTO pre-grant publications cover only

Table 4: Baseline Regressions

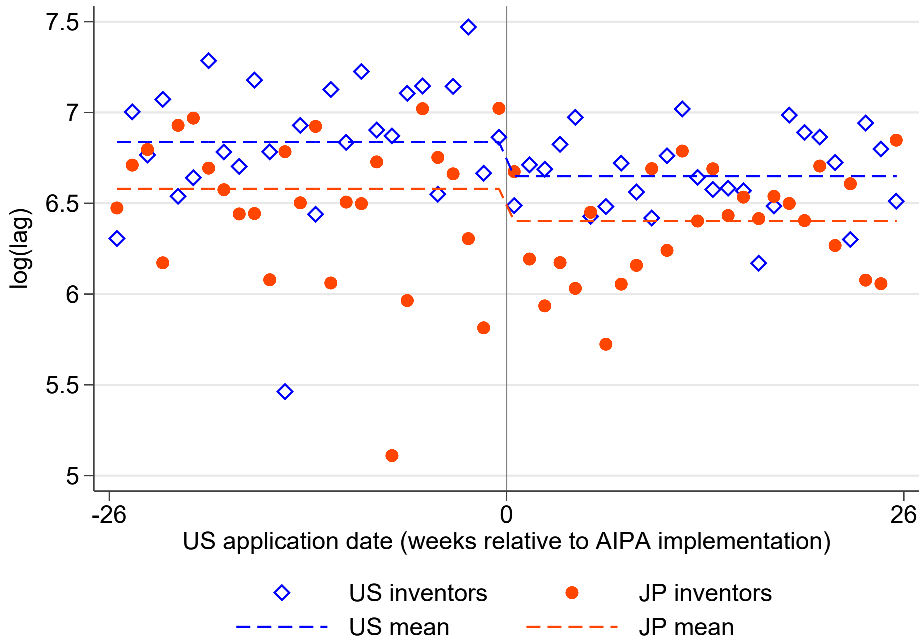
| | (1)<br>Twin Set | (2)<br>Triadic Set | (3)<br>Differenced |
|---|---|---|---|
| US × post-AIPA | -0.129** | -0.006 | -0.016 |
| | (0.013) | (0.386) | (0.235) |
| US × post-AIPA × Twin | | | -0.116** |
| | | | (0.050) |
| US assignee (citing) | 0.277*** | 0.257*** | 0.262** |
| | (0.006) | (0.004) | (0.013) |
| Science dependence | 0.006 | -0.004 | -0.006 |
| | (0.914) | (0.881) | (0.760) |
| Originality | -0.013 | 0.017 | -0.008 |
| | (0.681) | (0.460) | (0.733) |
| Citing inventors | -0.057 | -0.012 | -0.038 |
| | (0.564) | (0.850) | (0.638) |
| Constant | 6.790*** | 6.120** | 6.579** |
| | (0.007) | (0.022) | (0.022) |
| Observations | 2770 | 1542 | 4312 |

Notes: Threshold on the number of backward citations made by citing patents is set to 100. p-values displayed in round brackets. *p<0.1; **p<0.05; ***p<0.01

the inventions already disclosed in both of these publications. Nevertheless, as shown in Figure 2, even when language barriers never existed, we observe a drop in citation lag for both the Japan- and US-based inventors upon implementation of the AIPA. These results likely reflect the significance of inventor preference for citing local (USPTO) prior arts for both the Japan- and US-based inventors. At the same time, the averaged drop in the citation lag for the US inventors is much smaller in Figure 2 than in Figure 1 (0.19 and 0.32, respectively), unlike that for the Japanese inventors (0.18 and 0.19), suggesting that the EPO pre-grant publications in English were also used by the US inventors and any preference for citing the local prior art is only partial.

In order to formally test Hypothesis 2, we can assess the statistical significance of the difference between the US- and Japan-based inventors, observed in Figure 2, using our DID estimates. We test whether the effect of preferences of the US inventors to cite the local prior art of the home office (the USPTO) is significantly larger than those of the Japanese inventors for citing the local prior art of the foreign office (also the USPTO), based on our control and treatment groups of triadic patents. The results are reported in Column (2) of the Table 4. The estimated size of the effect of the AIPA is close to zero,

Figure 2: Average time to first citation to the triadic cohort



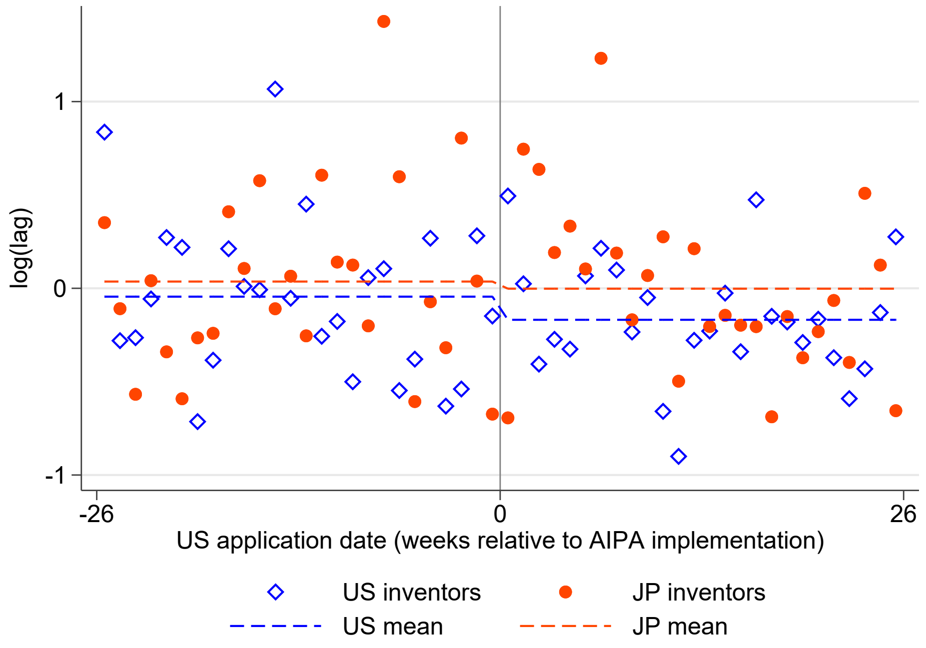

Notes: Solid lines indicate the (log-transformed) lag to first citation, by source, averaged by week. Dashed lines indicate the same, averaged over the 26-week periods before and after the AIPA came into effect.

indicating that there is no significant difference between the US and Japanese inventors in their preferences for citing local (USPTO) prior art. On the other hand, the estimate from the pure twins is highly significant and large. Thus, the results suggest that the US inventors' search is not limited to prior art among USPTO publications before the implementation of AIPA.

One important note regarding the DID results for the triadic patents is a potentially non-parallel pre-trend across the two geographical groups. As shown in Figure 2, there is a slight negative trend for the Japanese inventors' citation lags and a slight positive trend for the US inventors' citation lags, so that the result in Table 4 might represent an overestimate of the effect of Japanese inventor preference for local patents.

However, further differencing of the two DID results can address this problem in the identification of the effect of the AIPA on knowledge diffusion. That is, we are able to use the citations to triadic patents as an independent set of citations from the our 'twin' cohort to check the robustness of our baseline results in the face of potential office-level biases and differing temporal patenting trends. Empirically, because the citations to triadic patents are independent observations from the citations to our pure twin patents, we can combine the twin and triadic citation sets and adding a dummy equal to unity for the twin citations and equal to zero for triadic citations. A triple difference regression can then be implemented, where treated citations are those coming from US-based inventors

Figure 3: Difference in (log-transformed) time to first citation between twin and triadic cohorts




Notes: Solid lines indicate the difference in (log-transformed) lag to first citation, by source, averaged by week. Dashed lines indicate the same, averaged over the 26-week periods before and after the AIPA came into effect.

to patents in the twin cohort after the policy change.

This second difference (twin vs. triadic), shown graphically in Figure 3, is able to control for factors that are common to citations coming into both twin and triadic cohorts, including any general shocks to citation lags at the implementation of the AIPA that are unrelated to the removal of a language barrier (e.g., increased accessibility of applications previously only published before grant in foreign jurisdictions, for US inventors). Figure 3 does not display any obvious non-parallel trends, and additionally shows a marked widening of the gap between the two citation sources.

These data show the difference of (means of) log-transformed values, where the widening of about 0.12 implies a 12% drop in relative citation lag for US-based inventors. Reported in Column (3) of Table 4,[13] we find that this triple-differenced result is only slightly smaller than the baseline DID effect size, and still highly significant. Indeed, both results are in line with the average drop in citation lag displayed in Figure 3. This suggests that differences in the citation behaviour of Japan- and US-based inventors that are unrelated to the language barrier only have minimal impact on our estimated effect size, and that in controlling for them we still observe a significant reduction of relative

[13]While not displayed in Table 4, the appropriate interaction terms between citation source (US vs. Japan), target cohort (twin vs. triadic), and US filing date (before vs. after the AIPA), as well as all of these as standalone variables, are included. None of the variables or interactions omitted from the table have significant coefficients.

citation lag for the US inventors, providing further evidence for Hypothesis 1.

## 5.3 Effect of the AIPA on timing of examiner citations

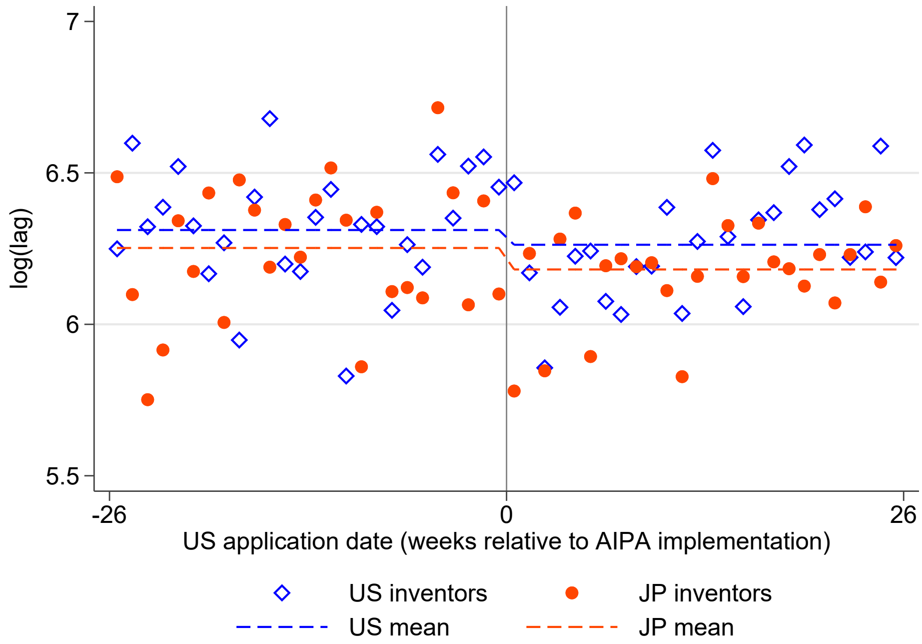
Many past works implicitly assume that examiner citations capture some of the knowledge flow from prior art to the focal invention, or, at the very least, add only noise to this signal. If this is the case, we may observe a similar acceleration of examiner citations from US-based applications relative to Japan-originating inventions upon the implementation of the AIPA. However, such assumption is unlikely to hold for the earliest (shortest-lag) examiner citations to the prior art. This is because inventors are generally not obliged to incorporate and cite prior art which isn't disclosed before their patent application. In contrast, examiner citations are usually added to patents late in the application process; indeed, an applicant to the USPTO might have to wait 1-2 years before an examiner even sees a patent application and, procedurally, will continue to add relevant prior art until grant.[14]

Figure 4 displays averaged US examiner citation lags to treatment and control groups. The average citation lag drops for both groups, although much less significantly than those in Figure 1. Thus, the AIPA allowed an US examiner to cite the Japanese patent applications earlier because of earlier pre-grant publications in English even if they had been published in Japanese by the JPO. Interestingly, the drop in citation lag is slightly larger for the patents by Japanese inventors, despite the expectation that the AIPA should have accelerated knowledge diffusion to the US inventors relative to that to the Japanese inventors.

The formal regression results displayed in Table 5 confirm a small but significant and *positive* treatment effect for examiner citations: a relative increase in the relative citation lag for US-based inventors. There are several potential interpretations for this effect; we suggest this coefficient is a combination of a substitution effect and a removal effect. The former effect can occur because, post-AIPA, *applicants* are now able to cite pre-grant publications that would previously have been secret. Pre-AIPA, some of this prior art would have only been available near the end of the examination process and thus the applicant would not have had a chance to cite it. Thus, many of these short-lag citations have moved from an examiner citation to an applicant citation. This substitution was more likely to be common for patents with US-based inventors than for those with Japan-based inventors, since the AIPA accelerated the English publications of the Japanese patents. To check this intuition, we add further statistical details in Appendix C, wherein

[14]In general, a patent is granted when the examiners determine that an application meets all the necessary requirements. Because prior art is the primary means through which the fulfilment of two of these requirements, novelty and non obviousness, is determined, the examination process is largely complete once the examiner fails to find any further prior art relevant to these requirements. That is, conditional on a patent having been granted, the examiner likely found and cited prior art until very close to the grant date (often 2-4 years after application, and sometimes even longer when measured from the priority date).

Figure 4: Average time to first citation for examiner citations




Notes: Solid lines indicate the (log-transformed) lag to first citation, by source, averaged by week. Dashed lines indicate the same, averaged over the 26-week periods before and after the AIPA came into effect.

Table A2 presents evidence of a greater drop in first examiner citations added to patents from US-based inventors compare to Japan-based inventors.

The 'removal' effect refers to a similar phenomenon, whereby US-based inventors and their attorneys find out about relevant prior art sooner than before and are able to draft around them when they may have been otherwise been grounds for rejection. This behaviour would avoiding a potential examiner citation to that prior art altogether and complement the coordination effect postulated by Aoki and Prusa (1996). Again, this phenomenon would have the effect of preferentially removing short-lag examiner citations post-AIPA.

Lastly, as seen in the coefficient for 'US assignee' in Table 5, the citing-country effect observed for applicant citations (whereby Japan-based inventors generally cited our cohort faster than US-based inventors) is insignificant for examiner citations. This result implies that, while office-level home biases in examiner citations have been reported in the past (e.g., Bacchiocchi and Montobbio (2010)), biases related to inventor location are not observed in our sample, at least for first citations. In summary, we find some evidence supporting hypothesis H3, noting the small potential bias indicated by the parallel trend test

Table 5: Examiner citations to twin cohort

| | (1) |
|---|---|
| US × post-AIPA | 0.022*** |
| | (0.003) |
| US assignee (citing) | 0.058 |
| | (0.105) |
| Science dependence | -0.002 |
| | (0.965) |
| Originality | 0.022 |
| | (0.544) |
| Citing inventors | 0.009 |
| | (0.920) |
| Constant | 6.329** |
| | (0.016) |
| Observations | 4360 |
| Parallel trends p-value | 0.092* |

Notes: Threshold on the number of backward citations made by citing patents is set to 100. p-values displayed in round brackets. *p<0.1; **p<0.05; ***p<0.01

# 6 Mediating factors

With baseline results that reflect a real acceleration in knowledge diffusion caused by the effective removal of a language-barrier, we now turn our eye to factors that may mediate the effective size of the language barrier. The effect of these barriers is moderated by the willingness of an individual firm or inventor to incur the cost of investing in translations of Japanese patent application documents. Having established that the AIPA likely increased the relative speed of initial knowledge diffusion by removing a language barrier between inventors in Japan and the United States, in this section we extend our baseline model for the pure twin sample to explore the ways in which different kinds of follow-on innovations and inventors were affected differently by the AIPA.

As hypothesised in 2.3, we consider two different classes of mediating influences in this work. The first are invention-specific mediators: those underlying characteristics of the knowledge being diffused from a Japanese patent application , which become known to US inventors only after the time of publication in English. Making a translation decision by targeting such characteristics is difficult. The second are appropriability mediators: those properties of the citing US inventors that affect their probability to engage in translations of the Japanese applications before the AIPA. This section is split along these lines.

## 6.1 Invention-specific Mediators

As described in Section 4, we conduct two different analyses for most mediating factors. The first extends the baseline regression to a triple-difference model by incorporating the mediating factor into the model, while the second splits the sample before running the usual DID regression. In this section, we examine three mediating factors, introduced sequentially below. A fourth mediating factor, indicating whether the application was eventually granted by the JPO, may be endogenous to the first citation lag; we therefore include this analysis in Appendix D for those interested.

### 6.1.1 Patent impact in the Japanese market

One relatively accessible aspect of invention quality is whether and how it is used by others.[15] Indeed, Japan is well-known as a technological powerhouse and produces many inventions that are of interest to Western markets,[16] in part due to local competition (Sakakibara and Porter, 2001). As such, the underlying importance of the technology within the Japanese market may moderate the rate at which follow-on innovators based in the US build on these technologies, because targeted translation of important patent applications is difficult. Because many of the technologies produced by Japan may be broadly categorised as 'high-tech' and fast-moving, the language barrier may play a particularly significant role in inhibiting the diffusion of these technologies.

In particular, we can capture the importance of an invention to the Japanese market by looking at in-text citations from follow-on innovations on pure-JPO applications (those corresponding to patent families only ever filed at the JPO) that are filed by Japanese firms (referred to hereafter as 'JPO citations'). Due to the location of these citations, we expect that it is highly unlikely that US-based inventors, attorneys had access to this information or any indicators that might be derived from them. That is, with respect to the timing of citations made by US-based inventors in their USPTO-granted patents, and unlike citations originating with USPTO patents or applications, this quality measure is likely to be significantly exogenous to the first citation lag in the United States.

The triple difference estimates, shown in Column 1 of Table 6, indicate that heterogeneity in the quality of the cited inventions is associated with a significantly accelerated diffusion upon the implementation of the AIPA. Further, our split sample results reinforces the notion of a strong mediating impact of patent quality on the significance of the language barrier.[17] Indeed, the low-quality subset see no significant change in citation

[15] Note, however, that this definition of patent quality has an ambiguous and complex relationship to other dimensions of quality, such as financial value or commercialisation (Higham et al., 2021).

[16] While the economy experienced a notable slowdown at the end of the 20th century (Branstetter and Nakamura, 2003; Arora et al., 2013), Japan's total inventive output continues to be very large in absolute terms (WIPO, 2021).

[17] Approximately half of all focal patents did not receive a single citation after the restrictions described

lag from US-based inventors relative to Japan-based inventors. The high-quality sample, in contrast, shows a highly significant treatment effect: roughly double the magnitude of our baseline treatment effect. As implied by the triple-difference results, this also suggests that treatment effects are practically confined to inventions cited within Japan. It is important to note that all of the patents in our cohort receive at least one USPTO applicant citation from each geographical source—many in the 'low-quality' sample are likely still moderately well-cited patents by USPTO standards. In any case, these results imply that the language barrier has the largest impact on patents which are important for domestic (Japan-based) follow-on innovators which, presumably, are also important for many industries in the US.

#### 6.1.2 Technology field-level dynamics

In addition to patent quality, another intrinsic characteristic that may play a large role in determining the effect of the AIPA on follow-on innovation is the speed at which a technology is moving. To assess this, we construct a measure, the mean citation lag of the *backward* citations made by the US equivalents within our focal cohort, dubbed 'field pace', which aims to capture the approximate pace at which technological developments occur in the field of the citing patent. It is unlikely, for example, that we see a patent for a novel semiconductor cite many 50-year old patents, but this could easily happen for a patent describing a novel kind of office chair. The intuition behind these measures, therefore, is that firms will more readily cite a novel technology when the field is moving very quickly and in a cumulative fashion. As such, we may expect a relative acceleration in citation lag within these fields in response to the AIPA, as this foreign knowledge suddenly becomes more accessible and firms are able to more quickly build on that knowledge to remain competitive. On the other hand, to the extent that such field characteristics are known ex ante, firms are more likely to translate the patents in that field. This tends to reduce the impact of the AIPA.

The triple-difference specification in Column (2) of Table 6 indicate that field pace does not appear to mediate the effect of the AIPA on the relative citation lag. However, upon splitting the sample at the median, we find that those cited families embedded in fast-moving technological fields (i.e., small pace values) see a much larger treatment effect than their counterparts in slow-moving fields.

In sum, we find strong support for hypothesis H4—early translation appears to have had the largest impact on citation behaviour to those patents that had high impact in the Japanese domestic market as well as to those that are part of fast-moving industries.

above. As such, "below the median" refers to these patents. Because all patents in this subset have the 'same' quality we are not able to include the quality measure as a covariate in this subset. This omission is also applied to the 'subset above median' regression so a fair comparison could be made between the subsets.

Table 6: Invention-specific Mediators

| | (1) | (2) |
|---|---|---|
| $M$: | JPO citations | Field Pace |
| $M$ × US × post-AIPA | -0.130*** | 0.027 |
| | (0.004) | (0.136) |
| US × post-AIPA | -0.133** | -0.145** |
| | (0.012) | (0.028) |
| $M$ × post-AIPA | 0.120* | -0.000 |
| | (0.056) | (0.965) |
| $M$ × US | 0.174*** | 0.066** |
| | (0.010) | (0.011) |
| US assignee (citing) | 0.282*** | 0.288** |
| | (0.008) | (0.020) |
| JPO citations (cited) | -0.182*** | |
| | (0.003) | |
| Field pace | | 0.051** |
| | | (0.015) |
| Observations | 2770 | 2728 |
| | ***Subset below median*** | |
| US × post-AIPA | -0.005 | -0.244** |
| | (0.366) | (0.011) |
| US assignee (citing) | 0.161** | 0.307*** |
| | (0.012) | (0.004) |
| Observations | 1318 | 1350 |
| | ***Subset above median*** | |
| US × post-AIPA | -0.228** | -0.045* |
| | (0.025) | (0.071) |
| US assignee (citing) | 0.366*** | 0.261** |
| | (0.009) | (0.019) |
| Observations | 1452 | 1378 |

Notes: Each mediator $M$ in this table corresponds to a property of the cited patent, where $M$ is described at the top of each column. Threshold on the number of backward citations made by citing patents is set to 100. p-values displayed in round brackets. *p<0.1; **p<0.05; ***p<0.01

## 6.2 Appropriability Mediators

In contrast to invention-specific mediators, appropriability mediators are the characteristics of the citing US applicants or inventors. We consider two appropriability mediators in this work: the R&D scale of the citing firm and the citing firm's past Japan-related patenting activity. These mediators are calculated using data characterising firm behaviour before the policy change and used for the entire sample period, so that they are plausibly exogenous. These mediators affect the probabilities for the US applicants to engage in translations from the incentives side, so that they heterogeneously affect the impact of the AIPA.

The empirical framework remains the same as for the invention-specific mediators, but splitting the samples for appropriability mediators introduces a quirk in the sample of citing patents. When splitting the sample by invention-specific variables, the first citations are unaffected, so the citations that make up the subsets were the same as those in the full sample of first citations. However, for appropriability mediators, we consider the first citation from each group of citing assignees, which in many cases introduces new 'first' citations when the first citation received from a particular subset of the assignees is not the first citation from any assignee. The resulting subsamples are still precisely the subsamples of interest, but note that the citations in the sample used for the triple difference specification are not exactly the same as those in each split-sample specification.

### 6.2.1 Citing firm R&D scale

The first appropriability mediator we consider is the R&D scale of the citing firm, the results for which are shown in Column (1) of Table 7. We proxy R&D scale by counting the total number of unique inventors listed on all patents filed by a citing firm in the year 2000. This definition roughly captures the size of a firm's R&D workforce at the time of the policy change, which in turn is mechanically related to both the total number of people employed by a firm (a common measure of firm size) as well as the resources that a firm has dedicated to R&D activities.

In the triple difference model, we observe that the treatment effect is significantly affected by R&D scale (although the significance is low) and the size of the coefficient is large, such that, at the mean, a two-standard deviation increase in R&D scale almost cancels out the treatment effect. Upon splitting the sample at the median, we find that the treatment effect almost disappears completely for the subset of high-R&D scale firms, while it is significantly negative and large (an acceleration of -22.8%) for low-R&D scale firms. The direction of this result is broadly in agreement with similar work conducted in the past (Büttner et al., 2022), and lends itself to the interpretation that smaller (or less R&D-intensive) firms are the main beneficiaries of earlier spill-ins of foreign knowledge caused by the implementation of the AIPA. Conversely, this conclusion also implies that smaller firms may remain at a relative disadvantage with respect to access to foreign-language knowledge today. We also suspect that firm size might interact with cited patent quality in a complex manner—for example, when a strict restriction on resources available for translation (pre-AIPA) lead to quality-based prioritisation of translation activities. We explore this dynamic further in Appendix E.

Further, we suggest that this is an entirely expected result to arrive at. In addition to having high willingness to spend more resources on translation of relevant prior art, larger firms may have access to more channels through which knowledge can flow, such as formal or informal business relationships with Japanese firms, or by virtue of more likely to have subsidiaries in Japan or Japanese-speaking inventors (Branstetter, 2006; Almeida and Phene, 2004; Miguélez, 2018).

It was noted during the passage of the AIPA through the US congress that 'small inventors' may, on the whole, shoulder the majority of deleterious consequences of such a law change.[18] However, the pre-grant publication opt-out clause, included specifically for the small firms that only patent in the United States, is not often used by small firms filing eligible applications (Graham and Hegde, 2015). Indeed, while we cannot speak to the net impact of the AIPA on small inventors, it appears that, with respect to access to foreign-language knowledge, the AIPA was of significant benefit to small firms. Closing the gap between small and large firms in terms of the timing of access to foreign-language knowledge adds a further, if relatively small, benefit of the AIPA to small firms.

#### 6.2.2 Japanese patenting activities of citing firms

Lastly, we consider the effect of involvement that a citing party may have in the Japanese market before the policy change. As a proxy for this behaviour, we construct a binary variable that takes the value 1 if we observe that a US-based firm that cites a patent in our focal cohort filed a patent application at the Japanese patent office at any time before the AIPA was implemented. The intuition behind this mediator is that we may expect that a firm filing for a patent in a foreign jurisdiction has some assets within that country, and thus is more likely invest in translations of patents published in the local language (before the implementation of the AIPA). Thus, such a firm will be less affected by the policy change as they already had some access to this source of knowledge.

Because this variable is a proxy for involvement in the Japanese market, it is only applied to US firms. That is, we consider all assignees that are granted US patents with all Japan-based inventors to be part of the Japanese innovation ecosystem by default. Upon splitting the sample, this choice means that the first Japan-originating citations remain the same as for the baseline regression, but the first US citation may not be the first citation overall, rather the first from their respective subset. Further, when we do not split the sample, there are no direct comparisons between US firms with and without

[18] `https://eagleforum.org/patent/nobel_letter.html`

prior patenting activities in Japan—we only consider the first US-originating and the first Japan-originating citations. Unlike firms' R&D scale, which is defined for both Japan- and US-originating patents, it not clear that we will obtain interpretable results using a triple difference framework in the this case. We avoid this potential pitfall altogether by splitting the sample as described above and running the standard difference-in-difference regression on each, then comparing the treatment effects observed for each subset.

The results of these regressions can be seen in Column (2) of Table 7. We find that the treatment effect for those firms that had not previously shown any interest in the Japanese market through their patenting activities is significantly higher than that found for those firms that had previously filed for a patent in Japan. That is, the removal of the language barrier led to faster follow-on innovation for the former group, though we note that the sample size for this group is quite small.

Further, this result is not entirely independent of that found for R&D-scale above; firms that had not previously filed for a patent in Japan were approximately one standard deviation smaller in terms of their R&D scale than those that had done so. As such, it is difficult to disentangle the interplay between attention to the Japanese market and the resources available to do so, with respect to their contributions to the treatment effects we observe. As such, we find general support for hypothesis H5, in that both smaller firms and those without previous patenting experience in the Japanese market see a much greater acceleration to Japan-originating prior art after the removal of the language barrier.

Table 7: Appropriability Mediators

| | (1) | (2) |
|---|---|---|
| $M$: | R&D intensity | Prior JPO app. |
| $M$ × US × post-AIPA | 0.063* | |
| | (0.094) | |
| US × post-AIPA | -0.151** | |
| | (0.015) | |
| $M$ × post-AIPA | -0.045* | |
| | (0.058) | |
| $M$ × US | 0.104* | |
| | (0.057) | |
| R&D scale (citing) | -0.153** | |
| | (0.017) | |
| US assignee (citing) | 0.224** | |
| | (0.019) | |
| Observations | 2770 | |
| | ***Subset below median*** | ***No prior JPO app.*** |
| US × post-AIPA | -0.318** | -0.177*** |
| | (0.020) | (0.007) |
| US assignee (citing) | 0.232* | 0.505*** |
| | (0.063) | (0.002) |
| Observations | 1020 | 720 |
| | ***Subset above median*** | ***Prior JPO app.*** |
| US × post-AIPA | -0.035 | -0.111** |
| | (0.199) | (0.045) |
| US assignee (citing) | 0.356* | 0.314** |
| | (0.057) | (0.024) |
| Observations | 1102 | 2428 |

Notes: Each mediator $M$ in this table corresponds to a property of the citing patent, where $M$ is described at the top of each column. Threshold on the number of backward citations made by citing patents is set to 100. p-values displayed in round brackets. $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

## 7 Conclusions

Language barriers impede the flow of otherwise freely available knowledge, inhibiting the rate of global technological development and economic growth more generally. In this work, we provide one of the first quantitative assessments of the role of language barriers on knowledge diffusion, focusing on the diffusion between two technological powerhouses with a very high language barrier dividing them.

In order to achieve this, we use a change in US patent policy, the American Inventors Protection Act of 1999, as a discontinuous event to study the effect of the removal of a language barrier on US inventors relative to those inventors based in the originating country, Japan. In particular, we are interested in the acceleration of patent citations made by US-based inventors to Japanese inventions for which patents were applied for only in the US and Japan, relative to Japan-based inventors, in the wake of the AIPA's implementation. Prior to the AIPA, only Japanese-language pre-grant publications existed for these inventions. By focusing on acceleration caused by the AIPA, we can identify the impact that language barriers have on knowledge diffusion, and the factors that mediate this impact.

Using a difference-in-differences (DID) approach, we have probed the existence and mediators of the language barrier on knowledge diffusion along several lines of inquiry. First, we find that the implementation of the AIPA significantly accelerated the diffusion of knowledge generated in Japan to US-based inventors. While part of the acceleration of the citations by the US inventors can be attributed to inventors' preference for citing local prior art (unrelated to the removal of language barriers), such preference of a similar size exists for the Japanese inventors as well—this is established using DID estimations based on a set of triadic patents, which have (pre-AIPA) English pre-grant publications from the EPO.

The average time to the first citation declined by about 13%, around a half of the total citation delay of the US-based inventors relative to the Japan-based inventors. A small part of this acceleration can be attributed to inventors' preference for citing local prior art. A DID estimation based on triadic patents, for which there exists a EPO pre-grant publication in English, is able to capture this bias as well as other non-knowledge flow-related impact of the AIPA on citation lag. As such, we conducted a triple difference estimation using a combined pool of pure twin patents and triadic patents, and obtained essentially the same results as those from the pure twin sample.

The same acceleration was not found for a similar sample using examiner citations and in fact a deceleration was found. Our findings are consistent with recent evidence that examiner citations to most recent prior art are unlikely to capture knowledge flows to inventors at the time of invention.

We then look at invention-specific and appropriability mediators of this effect. In the former case, we find that the AIPA significantly accelerated the diffusion and use of knowledge contained in the highest quality patents as well as from those existing within a fast-paced technological ecosystem. In the latter case, we observe that the acceleration effect appears to be almost entirely confined to firms with the smallest R&D scales, and to those that were unlikely to be involved in the Japanese market before the policy change. These findings support a simple but robust theory that firms with large complementary assets had stronger incentives and better capabilities to translate Japanese patent applications before the AIPA, and that patent quality is uncertain before translation, so that targeting translation at high quality patents is difficult.

This work makes two important contributions to the literature. First, besides providing causal evidence of the role of language barriers on knowledge diffusion, we have also demonstrated the heterogeneity of the impact of language barriers, which are more constraining for small firms and the firms with little involvement in the Japanese market, which would have important policy implications. We have thus demonstrated that the patent system has provided a significant public good through the translation of foreign applications in the pre-grant, lowering barriers to access to this knowledge in a similar way that patent libraries did for USPTO patents in the pre-internet era (Furman et al., 2021). The results also suggest that sophisticated and technologically specialized machine translation of technical information has the potential to have a large impact on the global diffusion of knowledge, and ensuring the easy access by small firms and individual researchers to such services is very important for realizing this potential.

Second, our analysis of knowledge diffusion pays close attention to the mechanisms of citations and controls for citation biases of applicants and examiners, without depending on the assumption that applicants and examiners are more likely to search locally for prior art for our identification (Hegde et al., 2022). Our finding that examiner citations could not mimic the results from applicant citations suggests that the inclusion of examiner citations would provide biased information about knowledge flows to inventors, especially for those from recent prior art. We have also shown the significance of inventor and examiner preference to cite local prior art (at the office level), which may not accompany knowledge flow, and control for them in our estimations.

One avenue of future work that we do not consider empirically in this work is the response of firms, both within and outside the United States, to the AIPA with respect to English-language disclosure. Taking a combinatorial view of invention (Weitzman, 1998; Fleming, 2001), when some inventors have access to more 'public' knowledge than others, then those inventors will have more components with which to conduct technological search (Stuart and Podolny, 1996; Kauffman et al., 2000)—their technological landscape is a superset of that of an inventor without adequate access to useful translations.

Lastly, another avenue of future work may be able to use the AIPA as an event to study the impact of earlier disclosure of purely domestic USPTO applications on countries outside the US, particularly in places where English is more or less commonly spoken.[19]

[19] We note that even within English-language patents, effective levels of disclosure can vary significantly (Fromer, 2008; Ouellette, 2012; Dyer et al., 2020) and, indeed, the technical content is often intentionally obfuscated (Devlin, 2009; Kong et al., 2020).

# References

P. Aghion and X. Jaravel. Knowledge spillovers, innovation and growth. *The Economic Journal*, 125(583):533–573, 2015.

J. Alcacer and M. Gittelman. Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics*, 88(4):774–779, 2006.

P. Almeida. Knowledge sourcing by foreign multinationals: patent citation analysis in the us semiconductor industry. *Strategic Management Journal*, 17(S2):155–165, 1996.

P. Almeida and A. Phene. Subsidiaries and knowledge creation: The influence of the mnc and host country on innovation. *Strategic Management Journal*, 25(8-9):847–864, 2004.

U. Ammon. *The dominance of English as a language of science*. Mouton de gruyter Berlin, 2001.

R. Aoki and T. J. Prusa. Product development and the timing of information disclosure under US and Japanese patent systems. *Journal of the Japanese and International Economies*, 10(3):233–249, 1996.

A. Arora, L. Branstetter, and M. Drev. Going soft: How the rise of software-based innovation led to the decline of Japan's IT industry and the resurgence of Silicon Valley. *Review of Economics and Statistics*, 95(3):757–775, 2013.

D. B. Audretsch and M. P. Feldman. R&D spillovers and the geography of innovation and production. *The American Economic Review*, 86(3):630–640, 1996.

D. B. Audretsch and M. P. Feldman. Knowledge spillovers and the geography of innovation. In *Handbook of Regional and Urban Economics*, volume 4, pages 2713–2739. Elsevier, 2004.

E. Bacchiocchi and F. Montobbio. International knowledge diffusion and home-bias effect: Do USPTO and EPO patent citations tell the same story? *Scandinavian Journal of Economics*, 112(3):441–470, 2010.

J. Baron and H. Delcamp. The private and social value of patents in discrete and cumulative innovation. *Scientometrics*, 90(2):581–606, 2012.

S. H. Baruffaldi and M. Simeth. Patents and knowledge diffusion: The effect of early disclosure. *Research Policy*, 49(4):103927, 2020.

R. Belderbos and P. Mohnen. Intersectoral and international R&D spillovers. *UNU-MERIT Working Papers*, (047), 2020.

P. J. Bentley. Cross-country differences in publishing productivity of academics in research universities. *Scientometrics*, 102(1):865–883, 2015.

E. Berkes, K. Manysheva, and M. Mestieri. Global innovation spillovers and productivity: Evidence from 100 years of world patent data. *Available at SSRN 4121469*, 2021.

L. Bowker and J. B. Ciro. *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Group Publishing, 2019.

M. Y. Brannen, R. Piekkari, and S. Tietze. The multifaceted role of language in international business: Unpacking the forms, functions and features of a critical challenge to mnc theory and performance. In *Language in International Business*, pages 139–162. Springer, 2017.

L. Branstetter. Looking for international knowledge spillovers a review of the literature with suggestions for new approaches. *The economics and econometrics of innovation*, pages 495–518, 2000.

L. Branstetter. Is foreign direct investment a channel of knowledge spillovers? Evidence from Japan's FDI in the United States. *Journal of International Economics*, 68(2): 325–344, 2006.

L. Branstetter and Y. Nakamura. Is japan's innovative capacity in decline? In *Structural Impediments to Growth in Japan*, pages 191–223. University of Chicago Press, 2003.

S. Breschi and F. Lissoni. Knowledge spillovers and local innovation systems: A critical survey. *Industrial and Corporate Change*, 10(4):975–1005, 2001.

B. Büttner, M. Firat, and E. Raiteri. Patents and knowledge diffusion: The impact of machine translation. *Research Policy*, 51(10):104584, 2022.

R. J. Caballero and A. B. Jaffe. How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. *NBER Macroeconomics Annual*, 8:15–74, 1993.

T. Cavalier. Perspectives on machine translation of patent information. *World Patent Information*, 23(4):367–371, 2001.

D. T. Coe and E. Helpman. International R&D spillovers. *European Economic Review*, 39(5):859–887, 1995.

W. M. Cohen, R. R. Nelson, and J. P. Walsh. Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not). *NBER Working Paper*, (w7552), 2000.

M. Corsino, M. Mariani, and S. Torrisi. Firm strategic behavior and the measurement of knowledge flows with patent citations. *Strategic Management Journal*, 40(7):1040–1069, 2019.

V. Denicolò and L. A. Franzoni. The contract theory of patents. *International Review of Law and Economics*, 23(4):365–380, 2003.

A. Devlin. The misunderstood function of disclosure in patent law. *Harv. JL & Tech.*, 23:401, 2009.

A. Dickens. Population relatedness and cross-country idea flows: evidence from book translations. *Journal of Economic Growth*, 23(4):367–386, 2018.

T. Dyer, S. Glaeser, M. H. Lang, and C. Sprecher. The effect of patent disclosure quality on innovation. *Available at SSRN 3711128*, 2020.

M. P. Feldman. The new economics of innovation, spillovers and agglomeration: A review of empirical studies. *Economics of Innovation and New Technology*, 8(1-2):5–25, 1999.

G. Ferguson, C. Pérez-Llantada, and R. Plo. English as an international language of scientific publication: A study of attitudes. *World Englishes*, 30(1):41–59, 2011.

M. M. Fischer, T. Scherngell, and E. Jansenberger. Geographic localisation of knowledge spillovers: Evidence from high-tech patent citations in Europe. *The Annals of Regional Science*, 43(4):839–858, 2009.

L. Fleming. Recombinant uncertainty in technological search. *Management Science*, 47 (1):117–132, 2001.

J. C. Fromer. Patent disclosure. *Iowa L. Rev.*, 94:539, 2008.

J. L. Furman, M. Nagler, and M. Watzinger. Disclosure and subsequent innovation: Evidence from the patent depository library program. *American Economic Journal: Economic Policy*, 13(4):239–70, 2021.

W. W. Gibbs. Lost science in the third world. *Scientific American*, 273(2):92–99, 1995.

P. Giuri, M. Mariani, S. Brusoni, G. Crespi, D. Francoz, A. Gambardella, W. Garcia-Fontes, A. Geuna, R. Gonzales, D. Harhoff, et al. Inventors and invention processes in europe: Results from the patval-eu survey. *Research Policy*, 36(8):1107–1127, 2007.

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

Z. Griliches. The search for R&D spillovers. *Scandinavian Journal of Economics*, 94:29, 1992.

G. M. Grossman and E. Helpman. Trade, knowledge spillovers, and growth. *European Economic Review*, 35(2-3):517–526, 1991.

B. Hall, C. Helmers, M. Rogers, and V. Sena. The choice between formal and informal intellectual property: A review. *Journal of Economic Literature*, 52(2):375–423, 2014.

B. H. Hall, A. B. Jaffe, and M. Trajtenberg. The NBER patent citation data file: Lessons, insights and methodological tools, 2001.

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

D. Hegde, B. Lev, and C. Zhu. Patent disclosure and price discovery. *Available at SSRN 3274837*, 2018.

D. Hegde, K. Herkenhoff, and C. Zhu. Patent publication and innovation. *NBER Working Paper*, (w29770), 2022.

K. Higham and T. Yoshioka Kobayashi. Patent citation generation at the triadic offices: Mechanisms and implications for analysis. *Available at SSRN 4022851*, 2022.

K. Higham, M. Governale, A. Jaffe, and U. Zülicke. Fame and obsolescence: Disentangling growth and aging dynamics of patent citations. *Physical Review E*, 95(4):042309, 2017.

K. Higham, M. Governale, A. Jaffe, and U. Zülicke. Ex-ante measure of patent quality reveals intrinsic fitness for citation-network growth. *Physical Review E*, 99(6):060301, 2019.

K. Higham, G. de Rassenfosse, and A. B. Jaffe. Patent quality: Towards a systematic framework for analysis and measurement. *Research Policy*, 50(4):104215, 2021.

A. B. Jaffe and M. Trajtenberg. Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. *Proceedings of the National Academy of Sciences*, 93(23): 12671–12677, 1996.

A. B. Jaffe and M. Trajtenberg. International knowledge flows: Evidence from patent citations. *Economics of Innovation and New Technology*, 8(1-2):105–136, 1999.

A. B. Jaffe, M. Trajtenberg, and R. Henderson. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108 (3):577–598, 1993.

A. B. Jaffe, M. Trajtenberg, and M. S. Fogarty. Knowledge spillovers and patent citations: Evidence from a survey of inventors. *American Economic Review*, 90(2):215–218, 2000.

S.-L. Jang, S. Lo, and W. Chang. How do latecomers catch up with forerunners? Analysis of patents and patent citations in the field of flat panel display technologies. *Scientometrics*, 79(3):563–591, 2009.

P. H. Jensen, A. Palangkaraya, and E. Webster. Disharmony in international patent office decisions. *Fed. Cir. BJ*, 15:679, 2005.

D. K. Johnson, D. Popp, et al. Forced out of the closet: The impact of the american inventors protection act on the timing of patent disclosure. *RAND Journal of Economics*, 34(1):96–112, 2003.

S. Kauffman, J. Lobo, and W. G. Macready. Optimal search on a technology landscape. *Journal of Economic Behavior & Organization*, 43(2):141–166, 2000.

W. Keller. The geography and channels of diffusion at the world's technology frontier. *NBER Working Paper*, (w8150), 2001.

W. Keller. International technology diffusion. *Journal of Economic Literature*, 42(3): 752–782, 2004.

J. Kim and K. Valentine. The innovation consequences of mandatory patent disclosures. *Journal of Accounting and Economics*, 71(2-3):101381, 2021.

N. Kishi and N. Takahashi. Licensing strategy of japanese firms and competitive advantage. *Annals of Business Administrative Science*, 9:1–12, 2010.

N. Kong, U. Dulleck, A. B. Jaffe, S. Sun, and S. Vajjala. Linguistic metrics for patent disclosure: Evidence from university versus corporate patents. Technical report, National Bureau of Economic Research, 2020.

J. Kuhn, K. Younge, and A. Marco. Patent citations reexamined. *The RAND Journal of Economics*, 51(1):109–132, 2020.

J. Kuhn, K. Younge, and A. Marco. Strategic citation: A reassessment. *The Review of Economics and Statistics*, pages 1–24, 2021.

R. Lampe. Strategic citation. *Review of Economics and Statistics*, 94(1):320–333, 2012.

M. A. Lemley and B. Sampat. Is the patent office a rubber stamp. *Emory LJ*, 58:181, 2008.

Y. A. Li. Borders and distance in knowledge spillovers: Dying over time or dying with age? Evidence from patent citations. *European Economic Review*, 71:152–172, 2014.

S. Lück, B. Balsmeier, F. Seliger, and L. Fleming. Early disclosure of invention and reduced duplication: An empirical test. *Management Science*, 66(6):2677–2685, 2020.

K. B. Luintel and M. Khan. Ideas production and international knowledge spillovers: Digging deeper into emerging countries. *Research Policy*, 46(10):1738–1754, 2017.

M. MacGarvie. The determinants of international knowledge diffusion as measured by patent citations. *Economics Letters*, 87(1):121–126, 2005.

C. Martinez. Insight into different types of patent families. 2010.

M. Marx and A. Fuegi. Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal*, 41(9):1572–1594, 2020.

M. Marx and A. Fuegi. Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations. *Journal of Economics & Management Strategy*, 31(2):369–392, 2022.

P. B. Maurseth and B. Verspagen. Knowledge spillovers in Europe: A patent citations analysis. *Scandinavian Journal of Economics*, 104(4):531–545, 2002.

J. Melitz. Language and foreign trade. *European Economic Review*, 52(4):667–699, 2008.

J. Melitz and F. Toubal. Native language, spoken language, translation and trade. *Journal of International Economics*, 93(2):351–363, 2014.

E. Miguélez. Inventor diasporas and the internationalization of technology. *The World Bank Economic Review*, 32(1):41–63, 2018.

F. Montobbio and R. Kataishi. The international dissemination of technological knowledge. In *Routledge Handbook of the Economics of Knowledge*, pages 175–198. Routledge, 2014.

Y. Murata, R. Nakajima, R. Okamoto, and R. Tamura. Localized knowledge spillovers and patent citations: A distance-based approach. *Review of Economics and Statistics*, 96(5):967–985, 2014.

F. Murray and S. O'Mahony. Exploring the foundations of cumulative innovation: Implications for organization science. *Organization Science*, 18(6):1006–1021, 2007.

F. Narin. Patents as indicators for the evaluation of industrial research output. *Scientometrics*, 34(3):489–496, 1995.

D. Nunan. The impact of English as a global language on educational policies and practices in the Asia-Pacific Region. *TESOL Quarterly*, 37(4):589–613, 2003.

H. Odagiri and A. Goto. *Technology and industrial development in Japan: Building* *capabilities by learning, innovation, and public policy.* Oxford University Press, 1996.

Y. Okada and S. Nagaoka. Effects of early patent publication on knowledge dissemination: Evidence from US patent law reform. *Information Economics and Policy*, 51:100852, 2020.

L. L. Ouellette. Do patents disclose useful information? *Harvard Journal of Law and Technology*, 25(2):545, 2012.

L. L. Ouellette. Who reads patents? *Nature Biotechnology*, 35(5):421–424, 2017.

W. G. Park. International R&D spillovers and OECD economic growth. *Economic Inquiry*, 33(4):571–591, 1995.

G. Peri. Determinants of knowledge flows and their effect on innovation. *Review of Economics and Statistics*, 87(2):308–322, 2005.

S. S. Rosenthal and W. C. Strange. Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, 85(2):377–393, 2003.

M. Sakakibara and M. E. Porter. Competing at home to win abroad: Evidence from Japanese industry. *Review of Economics and Statistics*, 83(2):310–322, 2001.

F. Salager-Meyer. Scientific publishing in developing countries: Challenges for the future. *Journal of English for Academic Purposes*, 7(2):121–132, 2008.

U. Schmoch. Concept of a technology classification for country comparisons. 2008.

M. S. Schomaker and S. Zaheer. The role of language in knowledge transfer to geographically dispersed manufacturing operations. *Journal of International Management*, 20 (1):55–72, 2014.

A. Seck. International technology diffusion and economic growth: Explaining the spillover benefits to developing countries. *Structural Change and Economic Dynamics*, 23(4): 437–451, 2012.

S. B. Seymore. The teaching function of patents. *Notre Dame L. Rev.*, 85:621, 2009.

S. Shibayama and Y. Baba. Impact-oriented science policies and scientific publication practices: The case of life sciences in Japan. *Research Policy*, 44(4):936–950, 2015.

I. Sin. The gravity of ideas: How distance affects translations. *The Economic Journal*, 128(615):2895–2932, 2018.

J. Singh and M. Marx. Geographic constraints on knowledge spillovers: Political borders vs. spatial proximity. *Management Science*, 59(9):2056–2078, 2013.

T. E. Stuart and J. M. Podolny. Local search and the evolution of technological capabilities. *Strategic Management Journal*, 17(S1):21–38, 1996.

G. Subramani, A. Aneja, O. Reshef, and W. S. Louis. Persistence and the gender innovation gap, 2021.

C. Tardy. The role of English in scientific communication: Lingua franca or Tyrannosaurus rex? *Journal of English for Academic Purposes*, 3(3):247–269, 2004.

P. Thompson. Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations. *The Review of Economics and Statistics*, 88 (2):383–388, 2006.

M. Trajtenberg, R. Henderson, and A. Jaffe. University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, 5(1):19–50, 1997.

D. Verhoeven, J. Bakker, and R. Veugelers. Measuring technological novelty with patent-based indicators. *Research Policy*, 45(3):707–723, 2016.

G. Von Graevenitz, S. Wagner, and D. Harhoff. How to measure patent thickets—a novel approach. *Economics Letters*, 111(1):6–9, 2011.

M. L. Weitzman. Recombinant growth. *The Quarterly Journal of Economics*, 113(2):

331–360, 1998.
WIPO. World intellectual property indicators 2021. Technical report, 2021.

# Appendices

## A Selection of backward citation threshold

For the purposes of the current work, super-citing patents that cite our cohorts present a potential problem for several reasons. First, the initial citations received by given patent are more likely to come from these highly-citing patents by virtue of the sheer volume of citations they make. Second, these highly-citing patents are growing in number over time. Third, and finally, and this practice is more common in the United States than in Japan. The net effect is that citations (and sometimes the first citation) from US patents are more common than expected at later lags. In the DID context, we care about whether this magnitude of this problem is the same for patents filed before and after the policy change. If the frequency of citations to our focal cohort from super-citing patents increases with time, their effect on citation lags is likely to be slightly more pronounced for more recently filed patents (i.e., those filed after the policy change). Further, because we are interested in the time to first citation, we cannot directly control for such citation inflation, as would be the preferred method in the case of accumulated citation counts (Hall et al., 2001; Higham et al., 2017). Instead, we are interested in extracting a signal of the timing of 'real' knowledge diffusion using applicant citations. As such, we follow an assumption of Kuhn et al. (2020): that the likelihood of a particular citation indicating real knowledge transfer decreases with the number of citations made by the citing patent. That is, by excluding patents that make more than some threshold number of citations, we may be able to get a more accurate view of the effect of the AIPA on knowledge diffusion across a language barrier.

With all this in mind, we present four initial DID regressions in Table A1 that inform the empirics conducted in this work. The first of these analyses, reported in Column (1), places no restrictions on citing patents in terms of the number of backward citations they make. The remaining three analyses conduct the same DID regression, but with citing patents making more than a certain number of citations (to any patent) excluded from the analysis (with thresholds of 100, 50 and 20). We note that while the raw number of citations being excluded after each cut is substantial, our focus on first citations means that the sample size isn't as affected as one may expect because there is often a another, non-excluded, citing patent that becomes the 'first citation'.

Comparing results for each threshold, we find that excluding citing patents making more than 100 citations yields a significant increase in the magnitude of the treatment effect over the unrestricted sample, and that decreasing the threshold further makes little difference thereafter. Further, thresholds of 20 and 50 lead to the parallel trend hypothesis being rejected at the 10% and 5% significance levels, respectively. As such, in the interest of keeping one threshold for which the parallel trend assumption is likely to hold for all analyses, we set the limit of total backward citation counts for citing patents to 100.

Table A1: Baseline DID results: Threshold discovery

| | (1) All | (2) ≤ 100 | (3) ≤ 50 | (4) ≤ 20 |
|---|---|---|---|---|
| US × post-AIPA | -0.066** | -0.129** | -0.122** | -0.136* |
| | (0.033) | (0.013) | (0.044) | (0.083) |
| US assignee (citing) | 0.190*** | 0.277*** | 0.347** | 0.458** |
| | (0.004) | (0.006) | (0.030) | (0.033) |
| Science dependence | -0.012 | 0.006 | 0.016* | 0.036 |
| | (0.856) | (0.914) | (0.065) | (0.311) |
| Originality | -0.019 | -0.013 | -0.018 | -0.020 |
| | (0.362) | (0.681) | (0.759) | (0.826) |
| Citing inventors | -0.048 | -0.057 | -0.055 | -0.066 |
| | (0.589) | (0.564) | (0.600) | (0.555) |
| Constant | 6.776*** | 6.790*** | 6.752*** | 6.666*** |
| | (0.001) | (0.007) | (0.005) | (0.002) |
| Observations | 3054 | 2770 | 2444 | 1874 |
| Parallel trends p-value | 0.589 | 0.341 | 0.046** | 0.052* |

Notes: Columns are different thresholds on the number of backward citations made by citing patents. p-values displayed in round brackets. *p<0.1; **p<0.05; ***p<0.01

# B Graphical explorations of average citation lags

Here, we explain the construction of Figures 1–4. These figures show the weekly averages of the first citation lags to our focal group (after the logarithmic transformation) for 26 cohorts (weeks) either side of the implementation of the AIPA pre-grant publication policy; these averages are shown separately for applicant citations to the focal cohort and for the alternative citation sets (applicant citations to the triadic cohort and examiner citations to the focal cohort). Those patents which were filed at the USPTO before the AIPA implementation date were only published in English at grant, while those filed after the implementation date were published 18 months after their priority date (which is generally not the US application date for this particular cohort). Citations over which the averages are calculated are split by geographic source (though note that all citations are from USPTO-granted patents), and the average for each source-disclosure status set is displayed. For the reasons discussed in Section A above, citing patents are restricted to those making 100 citations or less to match our baseline results.

# C Applicant and examiner shares of first citations

Table A2 shows the raw percentage of first citations to our control and treatment groups that come from each citing party when we pool the two citation types, as done by Hegde et al. (2022). While 52.1% of all citations to our cohort are applicant citations (unreported), the probability that the *first* citation received by a patent (that is not a self-citation) was made by an applicant is much lower (≤40%). In addition, the share of the first citations made by examiners declined after the implementation of the AIPA, since more prior art became available early.

Table A2: First citation by citing party and geographical source.

| | US inventors | | JP Inventors | |
|---|---|---|---|---|
| Citing party | Pre | Post | Pre | Post |
| Applicant | 34.8% | 40.0% | 33.8% | 36.0% |
| Examiner | 65.2% | 60.0% | 66.2% | 64.0% |

Notes: Percentages of all first citations (without conditioning on source) that were added by applicants and examiner, split by geographical source and by whether the cited patent was filed pre- or post-AIPA. Threshold on the number of backward citations made by citing patents is set to 100.

# D JPO Patent Grant

One factor that could mediate the rate of flow of knowledge across a language barrier is the legal quality of the invention relative to the prior art. To this end, we consider whether the first-filed JPO equivalent within each cited family was eventually granted — while we condition all cited families on the grant of their US equivalent, we do not do so for Japanese equivalents as we only care that the patent applications were published in Japanese for all families in our cited cohort. Observing whether or not a patent is granted provides us with a threshold measure of novelty and inventive-step size or, at the very least, firm persistence (Lemley and Sampat, 2008; Subramani et al., 2021) that may capture aspects of the application's relevance to the US market. However, at the time that the AIPA was implemented, firms could choose to delay substantive examination for up to seven years. That is, while the applications were published after 18 months as usual, the grant decision may be influenced by the USPTO patent prosecution process and any subsequent events, at least for those assignees who chose to delay examination at the JPO. For this reason, this mediator may be partially endogeneous.

At the JPO, the patent prosecution process starts with an initial examination to confirm that the application meets the formal requirements for a patent.[20] Following this, an applicant files a request for examination if they would like to continue the patent application process. If an applicant does not do so within three years from the filing date, the application is deemed 'abandoned'. This is presumably a strategic move on the part of the applicant, particularly for the patents we consider in this work (that is, those with a granted US equivalent). Because applications are abandoned for a number of

[20]Note that this process does not attempt to make a judgement on novelty or non-obviousness; these properties are assessed in the 'substantive' examination that follows a request for examination from the applicant.

reasons, many of which are not directly related to quality (Jensen et al., 2005; Kishi and Takahashi, 2010), we exclude cited families with abandoned applications (about 15% of the sample) from the analyses for which JPO grant is the mediator. Grants and rejections (about 55% and 30% of the sample, respectively), on the other hand, are decisions made by the same actor (the JPO) and thus provide a reasonable division of the sample that results in a binary variable indicating whether the family's first JPO application was granted.

Incorporating this binary variable into the triple difference framework, we find that families associated with a granted JPO equivalent see a significant acceleration to citations received from US-based inventors (relative to those based in Japan), as shown in Column (1) of Table 6. However, within this specification, we also see a slight relative increase in citation lag from treated inventors to applications that were not granted at the JPO. While the interplay between grant decisions in one jurisdiction and the timing forward citations in another is undoubtedly complex, we interpret this result as the AIPA providing an earlier signal about the ex-ante quality of foreign knowledge through publication.

One explanation for this apparent deceleration could be a strategic change by Japan-based applicant with respect to USPTO filings; in particular, the AIPA may have dissuaded Japanese firms from filing weak patent applications for inventions which are more likely to be useful to American competitors. Previously, the technical content of these applications, if rejected, would never be disclosed in English. Conditional on a patent being of low-quality, Japanese firms may not want to take the unnecessary risk of disclosing these ideas to competitors without the guarantee of patent protection. This would lead to a decrease in the relevance of Japan-based firms' lowest-quality applications to US inventors. While we believe this explanation to be reasonable, we emphasise that this reasoning and others like it are purely speculative at this stage and leave the examination of such strategic responses for future research.

Upon splitting the cited families along JPO-granted/JPO-rejected lines we confirm that an acceleration is only observed for those families with eventual JPO grants. However, we also note that we do not see a deceleration for the rejected subset, as indicated in the triple-difference specification; instead, we don't observe a treatment effect at all. It may be the case that unobserved changes in patenting strategy of the kind described about lead to complex interactions exist between the three components of the triple difference estimation that lead to an apparent deceleration in that specification.

# E Joint analysis of patent quality and R&D scale

When justifying our choice of mediators, we make the argument that patent quality (e.g., technological importance, relevance, or timeliness) is at least partially unobserved. In the most extreme case where quality is completely unobserved, pre-AIPA translation incentives are independent of quality. However, we do not believe that this is entirely true; instead, we suggest that quality is probably observed, but extremely noisily. This assertion is consistent with the results in the main body—after all, universal translation plausibly increases firms' ability to judge the potential importance and relevance of a particular invention described in this foreign prior art.

At the same time, the impact of this quality effect may be itself mediated by how much access a firm likely had to the prior art before the AIPA was implemented, which we proxy with a measure of R&D scale. To investigate this, we further split our sample: first we split it at the median by the R&D scale of the citing firm in the year leading up to the policy change, then we split each of these subsamples in half again at the median number of in-text citations the cited patent received from purely-domestic JPO applications (our proxy for patent impact in the Japanese market). For convenience, we will refer to this latter metric as patent 'quality'. We then run a simple difference-in-differences regression in the same manner as the split-sample analyses in the main body.

The results of this analysis are displayed in Table A4. For small firms, we observe a large acceleration of knowledge diffusion for both low and high quality patents. While the average acceleration appears to be larger for low quality patents, the two coefficients are not significantly different.

Large firms, on the other hand, saw no overall treatment effect when only firm size was considered as a mediator. Upon splitting the sample further, more complex behaviour emerges. We see that while citations to low-quality patents decelerates, citations to high-quality patents accelerates with about the same magnitude. This suggests that larger firms were likely translating foreign documents much more indiscriminately than small firms and with less concern for the usefulness of the knowledge therein (perhaps, for example, as part of a prior art search rather than for technological exploration). After the removal of the language barrier, these firms appear to have substituted citations to low quality patents for citations to higher quality patents, indicating that a more accurate observation of quality was made possible by the AIPA.

Table A3: Auxilliary mediator: JPO grant

| | JPO grant |
|---|---|
| $M$ × US × post-AIPA | -0.132** |
| | (0.013) |
| US × post-AIPA | 0.053*** |
| | (0.010) |
| $M$ × post-AIPA | 0.071** |
| | (0.041) |
| $M$ × US | 0.131*** |
| | (0.007) |
| US assignee (citing) | 0.219** |
| | (0.034) |
| JPO grant | -0.082* |
| | (0.062) |
| Observations | 2638 |
| | ***JPO Reject*** |
| US × post-AIPA | -0.021 |
| | (0.315) |
| US assignee (citing) | 0.334** |
| | (0.029) |
| Observations | 800 |
| | ***JPO Grant*** |
| US × post-AIPA | -0.094** |
| | (0.022) |
| US assignee (citing) | 0.408* |
| | (0.058) |
| Observations | 1590 |

Notes: $M$ corresponds to a binary variable indicating whether a patent was granted by the JPO. Threshold on the number of backward citations made by citing patents is set to 100. p-values displayed in round brackets. *p<0.1; **p<0.05; ***p<0.01

Table A4: Joint analysis of R&D scale and patent quality

| R&D scale:<br>Patent quality: | **Low**<br>**Low** | **Low**<br>**High** | **High**<br>**Low** | **High**<br>**High** |
|---|---|---|---|---|
| US × post-AIPA | -0.356* | -0.285* | 0.075** | -0.116** |
| | (0.067) | (0.083) | (0.048) | (0.047) |
| US assignee (citing) | 0.200 | 0.218** | 0.205 | 0.474** |
| | (0.183) | (0.013) | (0.123) | (0.027) |
| Observations | 432 | 588 | 500 | 602 |

Notes: Threshold on the number of backward citations made by citing patents is set to 100. p-values displayed in round brackets. *p<0.1; **p<0.05; ***p<0.01