The Effects of Multilateral IP Treaties on Patenting Behavior*

Elijah Coleman†
Vanderbilt University

January 4, 2022
Click here for the latest version

Abstract

I analyze the impact of several international treaties on global patenting behavior. Each treaty effect is estimated using a 3-way fixed effects Poisson pseudo-maximum likelihood estimator with the corresponding bias corrections thereof. Using the PATSTAT dataset from the EPO covering 211 countries/territories over the period 1980-2015, I construct bilateral aggregate patent flows using applicant addresses to determine the origin of each patent family. I find that trade liberalizing treaties as well as treaties strengthening intellectual property rights increase patent flows between countries. Joint membership in the Patent Cooperation Treaty has a robust positive effect in both developed and developing economies. Joint membership in other treaties (Paris, GATT, WTO, TRIPS) have strong impacts on patent flows between developing countries but limited effect between developed economies.

*I am grateful for the guidance of my advisor, Eric Bond, as well as for the insights from the members of my committee, Daniel Gervais, Tong Li, and Joel Rodrigue. I am also grateful to Vanderbilt University, and Scott Evans in particular, for providing a remote computer on which to host the large PATSTAT dataset and the ample assistance in creating and maintaining the server. This research is supported in part by the Kirk Dornbush Research Assistant grant.

†Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA. Email: elijah.s.coleman@vanderbilt.edu

1 Introduction

During the period from 1980 to 2015, there was substantial growth in world trade. Reductions in trade and communication costs, combined with reductions in tariff and non-tariff barriers through the WTO and the expansion of preferential trade agreements, resulted in an average growth in aggregate bilateral exports of 6.3% per year. Over the same period, cross border patenting activity grew even faster, with the number of patent applications by foreigners averaging a growth rate of 8.3% per year over the same period (Figure 1). The large growth in cross border patent applications is in part driven by reductions in both patenting and trade frictions between markets, largely as a result of bilateral or multilateral treaties. In the past few decades, the Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) and the Patent Cooperation Treaty (PCT) were developed. These treaties are focused on strengthening and standardizing intellectual property protections in member countries. The major trade agreement during the period is the World Trade Organisation (WTO) which subsumed and strengthened the trade liberalizations of the General Agreement on Tariffs and Trade (GATT). In addition to these large multilateral treaties, several regional or bilateral free-trade agreements were signed, many of which include explicit regulation regarding the issuance of patents. A cursory inspection of the growth of trade and patent flows (Figure 1) hints at the closely linked nature of patenting and trade and an acceleration of cross-border patenting circa 1995, concurrent with the introduction of the WTO/TRIPS.

Despite the TRIPS agreement being among the "most significant milestones in

[1]Trade flow data from the Atlas Dataset at Harvard University (2019)

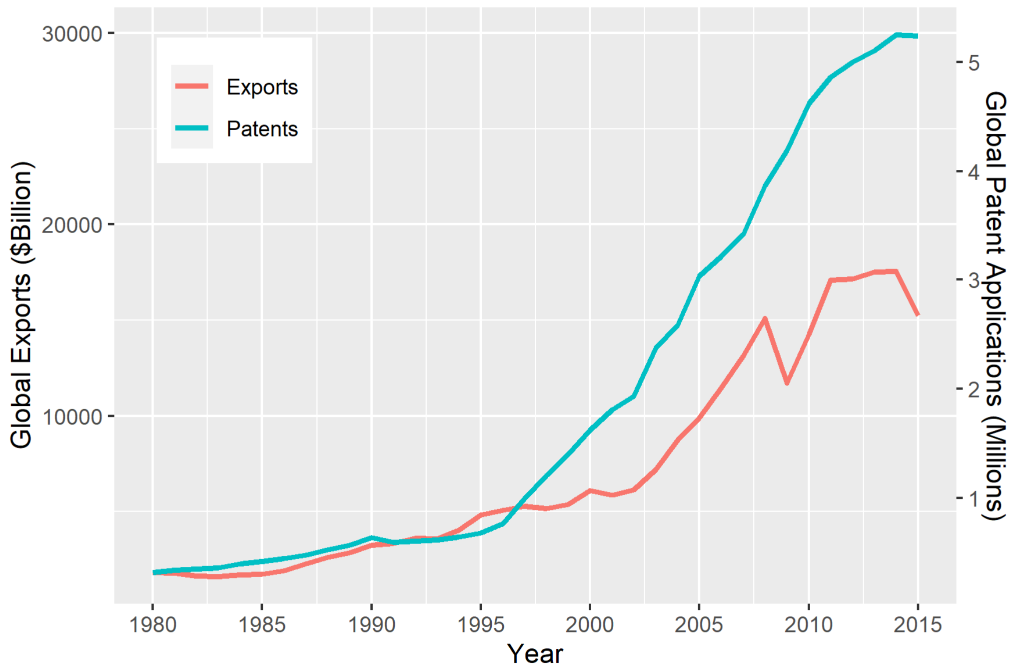


Figure 1: Total patenting vs. total exports[1]
"Domestic" patents have been removed to be comparable to exports which does not include "domestic" trade. The two measures are plotted on different scales to be visually comparable, trade flows follow the left axis, patent flows follow the right axis. Trade flows are measured in total USD value, patent flows are measured as aggregate patent applications.

the development of intellectual property"[2], the impact of TRIPS membership on international patenting behavior remains an open question. The goal of this paper is to evaluate the impact of WTO/TRIPS and other intellectual property treaties on international patent flows.

I employ estimation techniques developed in the gravity literature which are traditionally used to identify the determinants of international trade. I also take advantage of several new techniques to efficiently estimate a 3-way fixed effect Poisson

[2]Gervais (2012) pp. 3

Pseudo-maximum likelihood (PPML) and perform necessary bias corrections (Larch et al. (2019), Weidner and Zylkin (2021)). I find that the impacts of the treaties I consider are localized between developing countries with limited impact on developed economies. The exception to this is the Patent Cooperation Treaty which I find has a significant impact in both developing and developed countries, increasing patent flows by as much as 70%.

The remainder of this paper is organized as follows. Section 2 briefly outlines the related literature. Section 3 presents a model of patent flows that motivates the subsequent analysis. Section 4 reviews the construction of the data set. Sections 5 and 6 present the empirical strategies and results of the paper. Section 7 provides some closing remarks.

2 Related Literature

This paper dovetails into an existing literature studying the relationship between international trade, innovation, and intellectual property rights (IPR). Several papers have studied the impact of increased protections afforded to intellectual property. Maskus and Penubarti (1995) and Maskus and Yang (2018) provide evidence that stronger patent laws in a destination market attracts increased trade. Both papers employ qualitative survey data regarding the strength of IPR in each country like the Ginarte & Park index or the Fraser Freedom index. Brunel and Zylkin (2019) further examine this question but focuses on the impact of the number of patents in a market rather than indicators of IPR strength. They conclude that the number of patents in a destination market is positively correlated with its corresponding level of trade. This paper furthers this discussion by examining the impact of strengthening local IPR on the level of patenting.

Another branch of the literature takes a different perspective and analyzes the impact of increased profitability of trade on innovation and patenting. Aghion et al. (2018) study the impact of positive demand shocks in destination markets on firm level innovation decisions. Using data covering innovating French firms, they observe increased innovation and patenting in response to destination demand shocks. Bloom et al. (2016) and Coelli et al. (2016) ask a similar question but analyze it from the perspective of China as a destination, it further demonstrates the positive impact of increased demand in a destination on innovation and patenting activity elsewhere.

Apart from the impact of IPR on trade, and vice-versa, strengthening IPR has been shown to spur economic development. Branstetter (2006) demonstrates an increase in industrial activity in developing economies in response to IP reforms. The increase is driven largely by transfers of technology, and corresponding shifts in production, from parent companies to their affiliates in "southern" economies. Moser (2011) uses historic data from several world's fairs to demonstrate the diffusion of technology that patents provide, showing that stronger IPR led to greater geographic diffusion of technologies. While there is still a large debate regarding the welfare changes resulting from an IP reform, the increase in technology transfer seems unambiguous.

Empirically, this paper is related to the gravity literature that evaluates the impact of various policies on international trade. The structure of a gravity analysis was established in the seminal paper by Anderson and Van Wincoop (2003) which establishes a model of trade flows between countries based on characteristics of each market and the frictions between them. This analysis is well suited to analyzing patent flows because the decision to patent in a market is driven by many of the same factors as trade. More to the question at hand, a gravity analysis provides a useful framework by which to estimate the effect of bilateral treaties on patent flows. Silva and Tenreyro (2006) establish the standard practice of using the PPML estimator to identify various effects on bilateral flows, pointing out that using PPML bypasses many concerns regarding bias and consistency. Piermartini and Yotov (2016) emphasize this technique and further suggests using 3-way fixed effects to control for multilateral resistances a la Anderson and Van Wincoop (2003). Including 3-way fixed effects also alleviates concerns about the endogeneity of policy shocks. Larch et al. (2019) develop tools to overcome the computational intensity of deploying a 3-way FE PPML estimator in its analysis of the impact of currency unions on trade. Weidner and Zylkin (2021) demonstrate a bias in this estimator as a result of the incidental parameter problem. Despite the bias, they prove the estimator's consistency as well as providing a method to construct the necessary bias corrections. This paper implements the sum of these techniques to robustly estimate the effect of joint treaty membership on bilateral patent flows.

3 Motivating Model

3.1 Overview of Patents

A patent can be used by an innovating firm "to prevent third parties not having the owner's consent from the acts of: making, using, offering for sale, selling, or importing for these purposes that product."[3] This enables a firm to claim monopolistic profits over such a product within a given market. This protection begins in a limited capacity as soon as the patent application is submitted, and takes full effect upon the grant of a patent.[4] Note that the protections afforded by a patent in a given country exist only within that country, eg. a patent in the U.S. does not grant a monopoly in China. Therefore, innovating firms must decide in which countries to apply for a patent. This decision for each market is informed by a comparison of the market specific expected profits of applying for a patent versus the corresponding cost of the application.

The profitability of a patent is derived from sales of the patented product and thus is heavily dependent on many of the same factors that incentivize international trade: the size/wealth of a destination market, the iceberg trade costs, etc.

3.2 Model

Because of these similarities, I use a model developed in Bond & Coleman (2021)[5] which modifies the gravity equations described in Anderson and Van Wincoop (2003)

[3]TRIPS Article 28 1197 (1995)

[4]In the sample, about half of all patent applications are eventually granted. An application can be pending for several years before being either granted or rejected.

[5]Currently an early draft

to describe bilateral patent flows. The traditional gravity equation models the flow of trade from country to country based on several structural characteristics of the two markets:

$$X_{ijt} = \frac{Y_{it}E_{jt}}{Y_t}\left(\frac{t_{ijt}}{P_{jt}\Pi_{it}}\right)^{1-\sigma} \tag{1}$$

where X_{ijt} is the trade flow from country i to country j in year t. The gravity equation can be separated into three groups of effects. The first group is characteristics of the origin country: Y_{it}, the total production in the origin country; and Π_{it}, a measure of the profitability of all destination markets from the origin's perspective. The next group is characteristics of the destination country: E_{jt}, the total expenditure; and P_{jt}, the price index, together they form a measure of a country's tendency to import. The third effect comes from the features of pairs of countries, t_{ijt}, an iceberg trade cost that represents obstacles to trade. The iceberg trade cost can be composed of time invariant features of a pair of countries like the distance between them as well time varying features like bilateral membership in treaties governing trade.

In addition to trade frictions present in the traditional gravity equation, a patent is affected by the strength and structure of the intellectual property protections available between markets. In a "weak" patent regime, the monopoly afforded by a patent may provide limited protection from imitation (loss of the monopoly). To this end, the model incorporates the strength[6] and length of a patent in the expected profits of an application. This structure results in a cutoff rule for each country pair

[6]Strength of a patent is modeled as a hazard rate of the resulting monopoly. A "perfectly strong" patent carries no risk of infringement, the monopoly it grants lasts the entire length of the patent. A "perfectly weak" patent provides no protection against infringement, the monopoly is immediately lost.

based on the quality of an innovation. Aggregating the decisions of each firm over the distribution of innovation quality provides an equation describing aggregate bilateral patent flows similar to equation 1:

$$N_{ijt} = \overbrace{\left(\frac{P_{jt}}{\mu c_{it}\tau_{ijt}}\right)^{\theta}\left(\frac{E_{jt}}{\sigma}\right)^{\frac{\theta}{\sigma-1}}}^{\text{Traditional Gravity}} \overbrace{\iota_{it} b_{it}^{\theta}\left(\frac{\Psi_{ijt}}{F_{ijt}}\right)^{\frac{\theta}{\sigma-1}}}^{\text{Patenting}} \tag{2}$$

Where N_{ijt} is the total number of patent applications stemming from a firm in country i to country j in year t.[7] A full description and derivation of equation 2 is found in Appendix A. Here I merely point out the differences from the traditional gravity model. The innovative capacity of each origin country is captured jointly by ι, the total number of innovations in a source country, and b, the minimum productivity of those innovations.[8] The strength of an IP system is represented by Ψ which aggregates the impact of both patent strength and length[9]. Finally, F represents the cost of applying for a patent which is allowed to vary between pairs of countries. These costs can vary for a number of reasons: translation fees, legal representation costs, formal application fees, etc.

This model provides some insight into why different treaties could impact patent flows. Treaties dealing with intellectual property rights will affect both the overall value of a patent, Ψ, and the cost of a patent application, F. I discuss each treaty of

[7]For expediency, I frequently refer to "patent application flows" as just "patent flows." Throughout the paper, I am dealing purely with patent applications.

[8]Productivity is distributed according to a Pareto distribution with minimum b and dispersion θ.

[9]If national treatment is extended to all origin countries, Ψ will not vary across origins. However, at the beginning of the sample, there are still a number of countries not a party to any treaty requiring national treatment.

interest in more detail in section 4. Treaties involving trade liberalizations will also have an effect through a reduction of trade frictions, τ.

Note that this produces a model of applications as opposed to granted patents. The reason for this distinction lies in the information revealed in a firm's decision to submit an application. A firm forms expectations of the profits it could gain from having a patent in each country and compares it to the costs of acquiring such a patent. An application is therefore revelatory of the firm's view of the strength of a patent. Additionally, firms do not have the opportunity to wait for a patent to be granted in one country before applying in other locations.[10] This forces to firms to make expectations regarding profits in each location in the same way.

To identify the impact of bilateral and multilateral treaties, I translate equation 2 into an estimable form with 3-way fixed effect as follows:

$$N_{ijt} = \exp[o_{it} + \delta_{jt} + d_{ij} + \beta_1 \textbf{IPTreaties}_{ijt} + \beta_2 \textbf{TradeTreaties}_{ijt}]\epsilon_{ijt} \tag{3}$$

The set of origin-time (o_{it}), destination-time (δ_{jt}), and origin-destination (d_{ij}) fixed effects control for the variation in patent flows not related to treaty membership. The remaining variation is explained by features specific to each pair of countries that vary over time. This variation includes bilateral treaty membership. Joint membership in each treaty is indicated by a dummy variable (contained in

[10] Firms have 12 months, extendable to 30 months, to decide in which countries to submit a patent application. The decision of whether or not a patent is granted is made and average of 36 months after submission, too late to base further applications on it.

IPTreaties and **TradeTreaties**). For example:

$$WTO_{ijt} = \begin{cases} 1 & ,i \text{ and } j \text{ both members of the WTO in year } t \\ 0 & \text{o.w.} \end{cases} \tag{4}$$

An important implication of this analysis is that the estimated effect of each treaty will be its *discriminatory* effect, the estimated impact of a treaty on the patent flow between two countries conditional on *both* being members. While the text of each treaty is written to be applied in this fashion, it seems more than reasonable that part of the effect of each treaty may be purely non-discriminatory, affecting an origin or destination country uniformly regardless of its partner. A common technique used to identify such non-discriminatory effects on trade flows is to leverage the differential application of trade policy to inter- vs. intra-national trade. A change in tariffs impacts international trade flows but has no effect on domestic trade. This provides a baseline by which to measure the otherwise non-discriminatory effect of trade policies. The same solution is not applicable here because the policies regarding intellectual property rights in a country are applied equally to both domestic and international patent flows. For this reason, my analysis focuses purely on the discriminatory effect of each treaty.

4 Data

To implement equation 3, I need measures of patent flows and treaty membership. To construct bilateral patent flows, I employ the PATSTAT database from the European Patent Office (EPO). It is collected from national patent offices, unified, and distributed biannually. I use the Spring 2018 release. Each primary entry in the database is a patent application accompanied by many of its characteristics, e.g. date and type of application. I limit the sample period to the years post 1980 due to the relative sparsity of data prior to 1980. I also truncate the data post 2015 to avoid potential issues related to the data access for recent periods.[11]

Every patent application is classified by type. The most prominent patent type (and the type I focus on) is "patents of invention." A patent of invention typically represents a new innovation. This is in contrast to other classifications of patents which typically indicate smaller changes to existing technologies. The sample is limited to only patents of invention[12] to capture innovation flows rather than small improvements. These restrictions leave over 50 million applications in the sample (Table 1).

Restriction →	Full Sample	Years 1980-2015	Patents of Invention	Patents of Invention 1980-2015
# Patent Applications	99,298,592	62,926,896	82,147,980	52,031,026

Table 1: Sample Restrictions

Constructing the bilateral patent flows, N_{ijt}, for equation 3 requires knowing the origin and destination of each application. The origin of an application is concep-

[11]Patent applications are not published for a period a time following submission. The delay of publication varies by issuing authority and ranges from a few months to a few years. Truncating 3 years from the tail end of the data set removes concerns about the data in these periods.

[12]The U.S. classification terminology differs from elsewhere, a "utility patent" in the U.S. is akin to a patent of invention for the EPO

tually the origin of its innovation. Thus the constructed origin should be uniform across all applications within a *patent family*. Such a family is the collection of all applications (anywhere in the world) stemming from a single innovation. The first application submitted in each patent family is known as the *priority* application. Priority is a notion codified by the Paris Convention for the Protection of Industrial Property (1883) which grants an innovating firm 12 months following the "priority application" during which to submit patent applications in other Paris Treaty member countries. During this "priority period", no other actors can attempt to acquire IP protections over that innovation. I use the fact that every application in a patent family refers to the same priority application to link applications into these simple patent families.

Determining the destination of an application is straightforward in most cases. For most applications, the destination is merely the country in which patent protection is sought and is listed as the "country of authority." Some applications, however, are submitted as "European Union" patents. A firm can submit an application directly to the EPO for an EU patent and designate in which countries it is seeking patent protection[13]. Designating additional countries incurs additional fees (formal fees, translation costs, etc.) so firms make designation decisions in much the same way as other applications. Previous work has ignored this variation, assuming that each EU application was submitted in every EU country. I find that this assumption holds for less than 25% of EU applications with over 15% of applications designating

[13] An EU patent is merely an additional option available to innovating firms, it does not preclude national applications directly to EU countries

fewer than 5 countries. To capture this variation, I separate each EU application into distinct applications in each of its designated states.

Determining the origin of a patent is not quite as straightforward. Previous work constructing bilateral patent flows has used the home country of the applicants for an application as the origin. It is common for different applications in the same patent family to have different applicants, potentially from different countries. This leads to assigning multiple origins to a single innovation. To resolve this issue, I develop a method to determine the origin of an innovation using the addresses of the applicants of the priority patent. This method produces higher coverage over the data, allowing origin assignment for 13% more of the applications in the sample. It also constrains the origin to be uniform within a patent family which is more consistent with the conceptual definition of the origin of an innovation.[14]

A potential complication in determining the origin of a patent family lies in the fact that applicants can be from different countries. To resolve this discrepancy, I determine which applicant best represents the origin of the innovation. I use the "type" of each applicant (government, company, individual, etc.) with the assumption that the "economically larger" applicant best represents the origin of the innovation[15]. The specific ranking of applicant types along with other details of the origin assignment algorithm and considerations regarding its implementation are included in Appendix B.

[14]Without this method, 30% of all applications that are part of non-singleton patent families are assigned an origin that is different from the origin of the priority application.

[15]For example, an innovation with applicants 'General Electric from the U.S.' and 'John Doe from Canada' is assumed to have originated in the U.S.

Many of the patent families in the sample are singletons, i.e. patent protection for the given innovation is only ever sought in one country. Over the course of the sample, however, the average family size grows. In other words, firms are finding it on average more profitable to patent an innovation in multiple markets. The full distribution of family size is provided in Figure 2. Each bar in the figure gives the percentage of families of the given size in each sub-period. The bars advance in time from left to right, with the most recent distribution having the lowest proportion of singleton patent families.

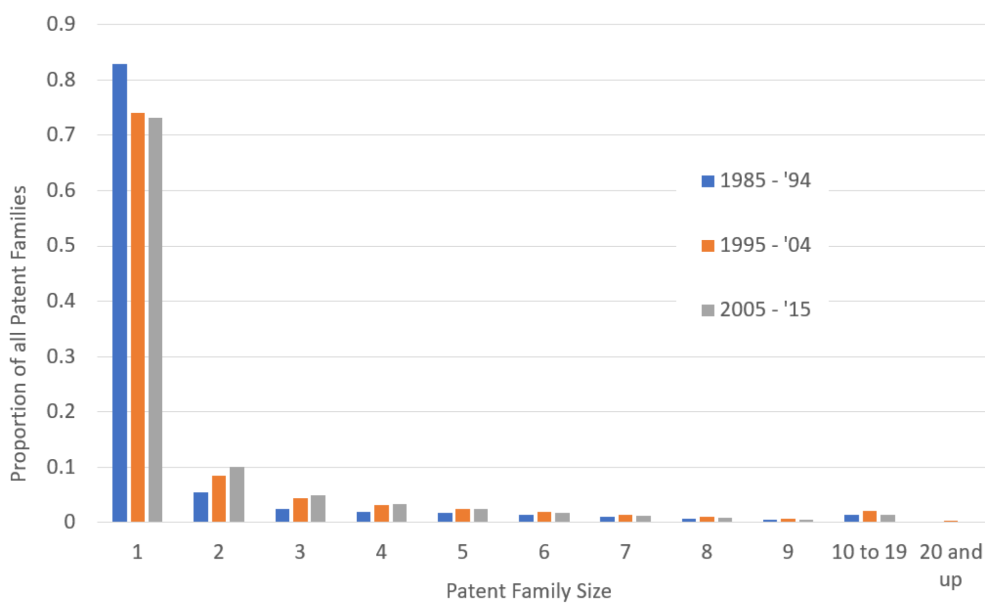
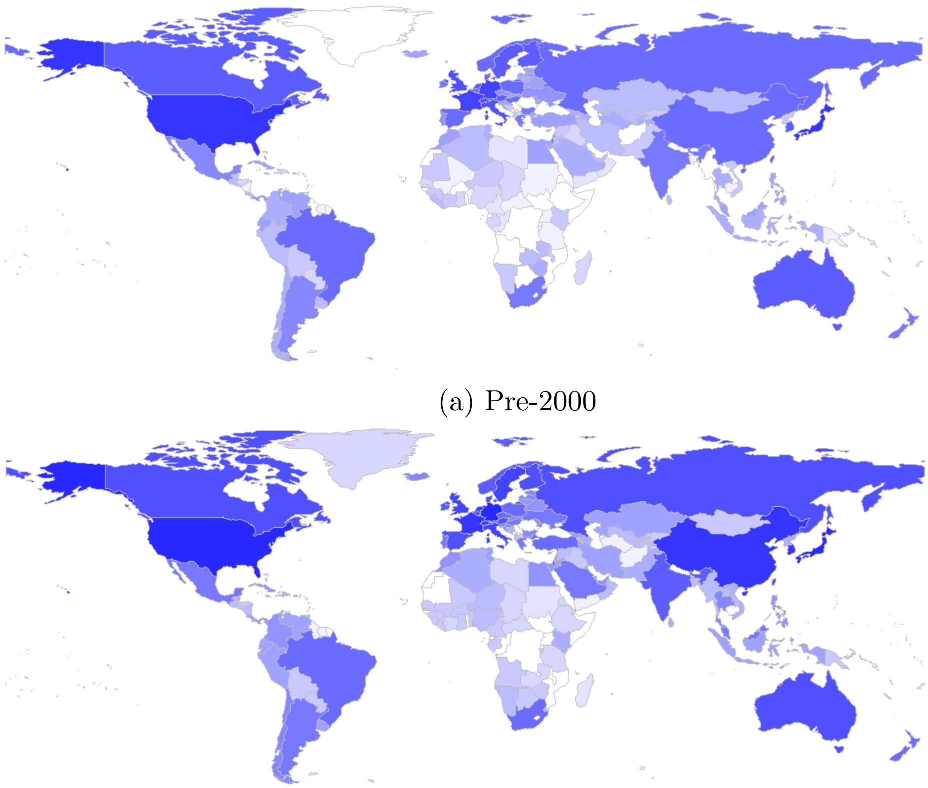


Figure 2: Distribution of patent family sizes over three sub-periods

A cursory examination of the constructed bilateral patent flows illustrates several notable features of patenting behavior. Figure 3a shows the global distribution of patent origins[16] before year 2000, with darker colors indicating higher patenting activity. Both frames show the concentration of patent origins in more developed countries. Figure 3b focuses on applications after 2000 and demonstrates an increasing prevalence patent originating in smaller or developing countries. Across the whole period, this is indicative of a large degree of heteroskedasticity in patent flows. Heteroskadasticity is also a prominent feature of trade flows. I address this using established techniques in the trade literature.

Another similarity of patent flows to trade flows lies in the prevalence of zeros in the panel of data. Particularly amongst smaller economies, it is relatively frequent that there are no patent applications between two countries in a given year. The simplest method of estimating the multiplicative models for both trade and patent flows is to log-linearize the equations and use OLS. The presence of zeroes in the dependent variable eliminates the possibility of log-linearizing without some modification. Silva and Tenreyro (2006) demonstrate that the common methods of circumventing the problem of zeros in log-linearizing all result in biased estimates. They propose the use of the PPML estimator which assumes the distribution of the dependent variable is Poisson. In addition to resolving the zeros issue, it is robust to misspecification of the conditional variance and is not biased in the presence of heteroskedasticity. For these reasons, I employ the PPML estimator to study bilateral patent flows.

[16]Constructing the same map using patent destinations rather than origins results in an almost identical map.



(a) Pre-2000

(b) Post-2000

Figure 3: Heat map of patent origins, darker colors imply more patents originating there. Aggregated within period.

4.1 Treaties

I now turn to a more detailed discussion of each treaty I include in my empirical analysis. Table 2 details the number of countries joining each treaty in the sample. There is a fair amount of variation in membership for each treaty.

Joined	Paris	GATT	PCT	WTO/TRIPS
Before	84	83	24	0
During	94	44	124	164

Table 2: Number of countries entering each treaty before and during 1980-2015

Paris Treaty: The Paris Convention for the Protection of Industrial Property (1883) sets a baseline for international cooperation in patent protection. It establishes two important features of international patenting law: 'national treatment' and 'priority.' National treatment stipulates that citizens from any member country be afforded the same availability of IP protections along with equal protection therefrom. This will affect patent flows through both increased access to patents (reduction in F_{ijt}) and increased protection (increased Ψ_{ijt}). Priority grants an innovator 12 months following the first patent application[17] of an innovation in which to submit further applications elsewhere. By allowing firms to delay patent applications, priority can alleviate financial frictions in paying the application costs in multiple markets, F_{ijt}. It also allows firms the chance to gather more information both on the economic and technological feasibility of a patent.

GATT: The general agreement on tariffs and trade (GATT) is designed to liberalize trade amongst its members. There is a small paragraph in the GATT text that addresses the issuance and protection of patents, but this section was rarely invoked and to little effect. Therefore, the channel through which GATT impacts patent flows is limited to the reduction in trade frictions. As such, I expect to find a positive effect on patent flows arising from a decrease in τ_{ijt}. The limitation of the sample to post 1980 has both pros and cons for this evaluation. A major drawback is that I do not observe many of the larger economies as they join GATT[18], this trun-

[17] Priority is based on the first application submitted to any country, all subsequent applications anywhere in the world refer back to the priority application.

[18] At the beginning of the sample, 83 countries were already GATT members

cation is necessary due to data coverage issues in earlier years. The effect of joining GATT is identified based on 44 countries joining between 1980 and 1995 (when the WTO subsumes GATT). Note that though a large economy like the U.S. does not join during the sample period, it is still part of many country pairs that become joint members (requiring both parties to be members). One of the benefits of the sample restriction is reduced concern that the effect of GATT varied over time, a possible result of the several negotiation rounds. The last round of GATT negotiations prior to its replacement by the WTO was the Tokyo round in 1979, conveniently. Because the sample period is lies between the Tokyo round and the WTO, the estimates can be viewed as the impact of the Tokyo round of GATT. For simplicity I continue to refer to GATT membership without the Tokyo qualifier.

PCT: The Patent Cooperation Treaty was developed in 1970 and "went into action" in 1978. One of the stated goals of the treaty was "to simplify and render more economical the obtaining of protection for inventions where protection is sought in several countries[19]." To that end, it introduced the 'international patent application' by which firms can apply for a patent in multiple member countries much more easily. An international application does not replace the need for national applications but it does provide a unified portal through which to apply for national patents. A PCT application begins with an international search to reveal any relevant, existing work followed by an examination of the application to determine the patentablity of an innovation. The applicant can then direct the application to any member country's

[19] (1970)

patent office for a national patent. While there are still fees associated with each of these national filings, several Offices' national filing fees are lower for international patent applications than they are for direct national applications in recognition of the work already done during the international phase. In addition, a PCT application extends the priority period granted by the Paris treaty to 30 months[20]. The combined effect of a PCT application is to reduce the cost of acquiring a patent amongst its member countries and provide additional information to firms as they determine in which countries to apply for a patent.

WTO/TRIPS: The WTO subsumed GATT in 1995 following several rounds of negotiations. Much like GATT, I expect the WTO to reduce the trade frictions between member countries resulting in increased bilateral patent flows. Beyond its role in liberalizing trade, the WTO introduced the agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS). The TRIPS agreement is groundbreaking for international intellectual property insofar as it requires a 'minimum standard level of protection' for IP available to citizens of any member country. This is a strengthening of the national treatment stipulation in the Paris Treaty which only requires equal treatment of all citizens of member countries, allowing for little to no available protection. The full impact of this requirement is likely not limited to discriminatory effects, a reformed IP system would likely be applied evenly regardless of applicant

[20]Priority is always measured from the priority application regardless of the order of application. While a PCT application can be used to extend the priority period from the standard 12 months to 30, it does not reset the clock. If a firm waits 11 out of the 12 months afforded by the Paris Treaty then submits a PCT application, the 30 months is measured from the beginning of the initial 12 month period.

nationality[21]. Because the WTO and TRIPS are so tightly linked, the opportunities to separately estimate the effect of each are limited. In an effort to separate the two effects, I leverage the transition period (5 years) given to developing countries to become TRIPS compliant. Note that this transition period was only granted to developing countries[22] that were original members of the WTO and is not available to countries joining subsequently. This period creates some variation in the timing between joining the WTO and being required to be TRIPS compliant.

Free Trade Agreements (FTAs): A number of regional and bilateral free trade agreements were formed over the sample. These agreements are typically viewed as trade liberalizations amongst the member countries. In that capacity, they will affect patent flows in the same way as GATT, reducing trade frictions (τ_{ijt}). I follow Santacreu (2021) and separate FTA's into those that deal strongly with intellectual property and those that do not. The data by which this determination is made is the DESTA dataset developed in Dür et al. (2014).

5 Empirical Strategy

Recall the estimation equation presented in section 3:

$$N_{ijt} = \exp[o_{it} + \delta_{jt} + d_{ij} + \beta \mathbf{Treaties}_{ijt}]\epsilon_{ijt} \tag{5}$$

[21] This also precludes the possibility of identifying the non-discriminatory effect using comparisons between inter- and intra-national flows.

[22] A transition period was also granted to "least-developed" countries. The period was originally 10 years but has been been extended several times, most recently set to expire in 2021.

To identify the effects of each treaty, I take guidance from Piermartini and Yotov (2016) which provides an assortment of pitfalls when dealing with trade data in a gravity setting. Fortunately, they also provide solutions to these problems. The issues are raised in the context of trade flows, but they are equally applicable to patent application flows. The most prominent feature of the data is the prevalence of zeros in aggregate patent application flows, particularly between smaller countries. This poses a problem for any method that requires log-linearization. Potential solutions to this issue include censoring the data to include only larger countries that do not have zeroes in the bilateral patent flows or removing individual zero observations. Both solutions remove a significant amount of information and ignore the meaning of a zero in the patent flows. Another method to eliminate zeros is to add 1 to all observations, this understandably produces a significant bias in estimates.

The second complicating facet of trade and patent data is its heteroskedasticity. Specifically, the variance in patent flows between large countries is correspondingly larger than that between smaller countries. If unaddressed, heteroskedasticity has the potential to induce a large bias in estimates. To address both of these issues, I use a Poisson pseudo-maximum likelihood estimator as presented in Silva and Tenreyro (2006). The PPML does not require log-linearization, thus skirting the problem of zero patent flows. Additionally, the PPML estimator is robust to heteroskedasticity. The only requirement for the PPML to provide consistent estimates is that the conditional mean is correctly specified. The traditional Poisson assumption is that the conditional variance is equal or directly proportional to the conditional mean. In practice, the conditional variance of patent data seems to be a higher order of the conditional mean. The variance of patent flows between large countries is disproportionately larger than that of smaller country pairs. Econometrically, this suggests an estimator that gives less weight to the larger observations due to their higher variance. Another consideration, however, is the quality of the data. I expect that the data from the larger countries is more reliable. A happy medium between the econometric and data concerns is to accept the traditional assumption of direct proportionality.

To control for the unobservable "multilateral resistances" like MFN[23] tariffs, average legal costs in a country, formal application costs, etc. I employ origin-year and destination-year fixed effects. This unfortunately eliminates our ability to estimate the effects of other interesting covariates like distance, GDP, GDP/Cap, and the Fraser freedom index, along with anything else that does not vary within a year-country pair. Fortunately, this does not disqualify the inclusion of trade policies, specifically joint membership in international treaties. Another potential problem is the endogeneity of those policy decisions based on some preexisting relationship between two countries that drives both patent flows and trade policies. I include origin-destination fixed effects to alleviate this concern. This inclusion unfortunately disqualifies the estimation of other bilateral characteristic effects, i.e. distance, common language, etc. Including all three types of fixed effects significantly limits what can be included in each regression, but makes the estimates of the trade policy effects more robust.

Estimation using this technique has long been prohibitively computationally in-

[23]Most Favored Nation

tensive. Blundell et al. (1995) suggest an alternate method by which to control for the possible endogeneity of bilateral trade agreements. Rather than including a full set of origin-time fixed effects, they control for persistent dyad effects by constructing a pre-sample mean[24]. Effectively, this implies using a few years of data at the beginning of the sample to construct the pre-sample, leaving less data for analysis. This is certainly effective at reducing computational demands and does not require too much data on the front end. It does, however, require some regularity on the nature of country relationships over the whole sample. Over the thirty-six year period in question, it is not unreasonable to think that the nature of relationships within each country dyad have changed in ways not fully reflected in the treaties being considered. While the set of origin-destination FEs does not capture such changing dynamics either, it does treat the whole period uniformly. I report results using the pre-sample mean for a comparison of the two methods in Table 5.

To implement the three-way FE PPML estimator, I use the technique described in Larch et al. (2019) which utilizes an iterative estimation approach. Weidner and Zylkin (2021) further refine this method by pointing out an asymptotic bias when using a 3-way FE PPML estimator. They demonstrate a bias in both the point estimates and standard errors[25] to a degree that can impact inference. For illustrative purposes, I report results both with and without this bias correction in Table 4.

[24]In addition to the pre-sample mean between a pair of countries, an indicator of whether or not any patenting activity occurs between them. This is included to deal with the zero lower bound on patent activity, zero activity between two countries indicates a different type of relationship altogether

[25]Weidner and Zylkin (2021) further demonstrate a bias in the standard errors of a 2-way FE PPML

5.1 Treaty Dummies

Since the WTO supplants GATT during the sample, the construction of the treaty indicators merits some attention. The GATT indicator represents joint membership in GATT up to 1994. In 1995 and following, GATT is set to 0 to reflect the fact that GATT as a standalone treaty no longer exists. One result of this construction is that the effect of joint GATT membership is estimated based on the variation provided by countries that enter GATT between 1980 and 1994. These countries are primarily developing countries.

The WTO indicator reflects joint membership in the WTO in years following 1995. Since the beginning of the WTO occurs in the sample, the estimated effect of joint membership is identified based on all members. Notably this includes several developed country pairs that are excluded from the effects of other treaties due to joint membership before the beginning of the sample.

In column 3 of Table 4, I report results with the effects of the WTO and TRIPS identified separately. The variation in timing between WTO and TRIPS "membership" stems from the transition period granted to developing countries before being required to be TRIPS compliant. The TRIPS indicator is set to 1 if the origin is a WTO member and the destination is required to be TRIPS compliant.

Following Santacreu (2021), I include an indicator of the presence of intellectual property provisions in each free trade agreement. I take these indicators from the Design of Trade Agreements dataset (DESTA[26]). The interpretation of the estimated effect of such provisions is an incremental effect relative to the overall impact of FTAs.

[26] Dür et al. (2014)

6 Results

Table 4 reports estimates of the effect of each treaty using the baseline specification. The first two columns present the results of 3-way FE PPML estimation with and without the bias correction[27]. While the bias correction has relatively limited effect on the point estimates, the standard errors are significantly biased downwards leading to overstated significance of the impact of each treaty. Column (3) performs the same estimation with the inclusion of the TRIPS indicator, leveraging the variation arising from the transition period.

The baseline specification is reported in column 2. I find that both trade and IP treaties have significant effects on patent flows. These estimates represent the percent change in bilateral patent flows stemming from joint treaty membership, i.e. the Patent Cooperation Treaty and the Paris Treaty have a strong positive effect, increasing patent flows by 71% and 107% respectively. The estimated effect of the other treaties are puzzling.

An FTA between countries stimulates 20% higher patent flows. One might expect this effect to be magnified by provisions in the FTA regarding intellectual property regulations. However, I find that there is no additional impact arising from such provisions based on the insignificant estimate for $FTA\ IP$.

The differing effects of GATT and the WTO present another puzzle. The impact of joint GATT membership is both strong and significant, increasing patent flows by 105%. In contrast to this, the estimated effect of joint membership in the WTO

[27] I include column (1) merely to highlight the importance of using the bias correction developed in Weidner and Zylkin (2021). All other estimations are bias-corrected using this method.

Dep. Var.: N_{ijt}	Biased	Bias Corrected	w/ TRIPS
GATT	1.166***	1.054**	1.054**
	(0.244)	(0.524)	(0.524)
WTO	0.500**	0.410	0.388
	(0.201)	(0.412)	(0.370)
PCT	0.671***	0.707***	0.707***
	(0.106)	(0.125)	(0.125)
Paris	0.930***	1.070*	1.070*
	(0.345)	(0.647)	(0.648)
FTA	0.190***	0.195**	0.195**
	(0.070)	(0.084)	(0.084)
FTA IP	-0.002	-0.011	-0.011
	(0.067)	(0.078)	(0.078)
TRIPS			0.023
			(0.193)
Observations	158,424	158,424	158,424
Origin-Time FE	Yes	Yes	Yes
Destination-Time FE	Yes	Yes	Yes
Origin-Destination FE	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Effect of various treaties, estimated using 3-way FE PPML with several specifications. Each estimate indicates a percent impact on patent flows.

is both smaller and statistically insignificant. This contrast is a bit surprising given the role of the WTO in both strengthening the trade liberalizations of GATT and introducing TRIPS. Column 3 reports estimates of the effects of the WTO and TRIPS identified separately, but I find no significant difference in the effect of each treaty based on this additional variation.

6.1 North and South

The estimated effects of each treaty are determined based on variation in joint treaty membership during the period. This variation is quite different for each treaty (Table 2). For GATT and the Paris Treaty in particular, the countries joining each treaty during the sample are primarily developing countries. This means that the estimated effect is heavily dependent on its effect in the economic 'South.' To demonstrate this, I split the sample into 4 groups: North to North, North to South, South to North, and South to South[28].

	North2North	North2South	South2North	South2South
GATT	0.169	-0.085	1.549***	1.630***
	(0.958)	(0.264)	(0.591)	(0.201)
WTO	-0.107	-0.244	0.402	0.635***
	(0.587)	(0.377)	(0.401)	(0.200)
PCT	0.745***	0.058	0.057	0.288**
	(0.157)	(0.093)	(0.145)	(0.145)
Paris		-0.515	-0.384	0.506**
		(0.490)	(0.501)	(0.221)
FTA	0.250*	0.086*	-0.073	0.799***
	(0.142)	(0.045)	(0.051)	(0.261)
FTA IP	-0.055	-0.114**	0.227***	-0.374
	(0.095)	(0.049)	(0.066)	(0.308)
Observations	59,199	26,934	43,859	15,469
Origin-Time FE	Yes	Yes	Yes	Yes
Destination-Time FE	Yes	Yes	Yes	Yes
Origin-Destination FE	Yes	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Effect of various treaties, estimated using 3-way FE PPML. The Paris Treaty is excluded from the North to North estimation due to a lack of variation in northern membership over the period.

Sub-setting the sample in this fashion reveals disparate effects of each treaty in the North vs. the South. The large impact of GATT and the Paris treaty in Table

[28]I define the economic North and South to be developed and developing countries respectively. This status is based on their WTO designation in 1995.

4 seems to be based on their effect in the 'South.' The significance of the WTO estimate in the South does not pass through to the aggregate estimate in Table 4 because a higher proportion of the variation in membership comes from developed countries. Overall, this analysis demonstrates a heavy impact of each treaty in South to South patent flows. This makes sense because the regulatory changes required in each treaty most heavily impact developing countries. Notably, the only multilateral treaty that has a significant effect in both the North and the South is the Patent Cooperation Treaty which is aimed at reducing the cost of patent applications in multiple countries.

6.2 Dyad FE vs. Pre-sample Mean

In Table 5, I report estimates using the pre-sample mean method described in Blundell et al. (1995) to control for country pair effects. Specifically I construct the following:

$$PreSampleMean_{ij} = \frac{1}{6} \sum_{t=1980}^{1985} N_{ijt} \tag{6}$$

$$PreSampleActive_{ij} = \begin{cases} 1 & PreSampleMean > 0 \\ 0 & PreSampleMean = 0 \end{cases}$$

Pre-Sample Mean controls for persistent dyadic frictions. Pre-Sample Active is included to differentiate between active dyads and inactive dyads which can have different underlying relationships. The benefits of this technique lie in its computational simplicity (relative to dyadic fixed effects). The cost of this method is the

	Baseline	Pre-sample Mean
GATT	1.054**	1.611***
	(0.524)	(0.490)
WTO	0.410	2.293***
	(0.412)	(0.753)
PCT	0.707***	0.865***
	(0.125)	(0.244)
Paris	1.070*	0.703*
	(0.647)	(0.401)
FTA	0.195***	0.295***
	(0.084)	(0.068)
FTA IP	-0.011	0.555***
	(0.078)	(0.062)
log(Presample Mean)		2.701***
		(0.043)
Presample Active		-1.183***
		(0.268)
Observations	146,702	
Origin-Time FE	Yes	
Destination-Time FE	Yes	
Origin-Destination FE	No	

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Effect of various treaties. Column 1 estimated using 3-way FE PPML. Column 2 estimated using 2-way FE PPML alongside the Pre-sample mean method from Blundell et al. (1995).

assumption that country pair relationships remain constant over the sample period. While this is a reasonable assumption when considering the effect of geographic remoteness or a shared language, it is less plausible that the underlying political and economic relationships between countries remain unchanged. Consider a scenario in which the idiosyncratic relationship between countries is increasing over time. The pre-sample mean is formed based on smaller interactions in early periods. The increasing dyadic effect is then incorrectly attributed to treaties that are joined later in the sample.

Column 2 reports estimates using this method in comparison to the baseline specification in column 1. As expected, the estimated effect of each treaty is significantly higher. The most significant increase is for the WTO estimate which rises from 40% to 230%. This is consistent with the expectation that the effect of treaties joined later in the sample (most true of the WTO) would be overstated using this method. To avoid this bias, I employ origin-destination fixed effects rather than a pre-sample mean to control for dyadic frictions.

7 Conclusion

This paper provides a quantitative evaluation of the impact of several multilateral treaties on international patent flows. To accomplish this, I develop a robust method by which an origin and destination is assigned to each patent application. A key feature of this method is the constraint that origin assignment must be consistent within a patent family. This requirement not only more closely matches the intuitive definition of an innovation's origin, but also provides higher coverage in origin assignment. On this basis, I construct patent flows between countries in each year. To estimate the discriminatory effect of treaty membership on bilateral patent application flows, I adapt techniques developed in the empirical gravity literature to patent flows. I use a PPML estimation alongside a full complement of fixed effects. I use a recently developed method to correct for the bias produced by a 3-way fixed effect PPML specification. I find that both trade and intellectual property treaties have significant impacts on patent flows between developing countries but have little effect in the developed world. A notable exception to this is the Patent Cooperation Treaty. In addition to its impact in developing countries, the PCT increases patent flows between developed economies by about 70%. This highlights the importance of the cost of patent applications in firms' decisions regarding destination markets for new innovations.

References

1197, 1869 U.N.T.S. 299; 33 I.L.M., "Agreement on Trade Related Aspects of Intellectual Property," 1995. Annex 1C to the Marrakesh Agreement.

(1970), 28 U.S.T. 7645; 1160 U.N.T.S. 231; 9 I.L.M. 978, "Patent Cooperation Treaty (PCT)," 1970. amended on September 28, 1979, modified on February 3, 1984, and on October 3, 2001.

305, 21 U.S.T. 1583; 828 U.N.T.S., "Paris Convention for the Protection of Industrial Property," 1883. as last revised at the Stockholm Revision Conference, September 28, 1979.

Aghion, Philippe, Antonin Bergeaud, Matthieu Lequien, and Marc J Melitz, "The heterogeneous impact of market size on innovation: evidence from

French firm-level exports," Technical Report, National Bureau of Economic Research 2018.

Anderson, James E and Eric Van Wincoop, "Gravity with gravitas: A solution to the border puzzle," *American economic review*, 2003, *93* (1), 170–192.

at Harvard University, The Growth Lab, "International Trade Data (SITC, Rev. 2)," 2019.

Bloom, Nicholas, Mirko Draca, and John Van Reenen, "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity," *The review of economic studies*, 2016, *83* (1), 87–117.

Blundell, Richard, Rachel Griffith, and John Van Reenen, "Dynamic count data models of technological innovation," *The Economic Journal*, 1995, *105* (429), 333–344.

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

_ , Ray Fisman, C Fritz Foley, and Kamal Saggi, "Does intellectual property rights reform spur industrial development?," *Journal of International Economics*, 2011, *83* (1), 27–36.

Brunel, Claire and Thomas Zylkin, "Do Cross-border Patents Promote Trade?," 2019.

Coelli, Federica, Andreas Moxnes, and Karen Helene Ulltveit-Moe, "Better, faster, stronger: Global innovation and trade liberalization," *The Review of Economics and Statistics*, 2016, pp. 1–42.

Dür, Andreas, Leonardo Baccini, and Manfred Elsig, "The design of international trade agreements: Introducing a new dataset," *The Review of International Organizations*, 2014, *9* (3), 353–375.

European, Patent Office, "PATSTAT 2018 Spring Edition," https://www.epo.org/searching-for-patents/business/patstat.html.

Figueiredo, Octávio, Paulo Guimarães, and Douglas Woodward, "Industry localization, distance decay, and knowledge spillovers: Following the patent paper trail," *Journal of Urban Economics*, 2015, *89*, 21–31.

Fisman, Raymond, Lee G Branstetter, and C Fritz Foley, *Do stronger intellectual property rights increase international technology transfer? Empirical evidence from US firm-level panel data*, The World Bank, 2004.

Gervais, Daniel J, *The TRIPS agreement: drafting history and analysis*, Sweet & Maxwell/Thomson Reuters, 2012.

Glick, Reuven and Andrew K Rose, "Currency unions and trade: A post-EMU reassessment," *European Economic Review*, 2016, *87*, 78–91.

Hall, Bronwyn H and Dietmar Harhoff, "Recent research on the economics of patents," *Annu. Rev. Econ.*, 2012, *4* (1), 541–565.

Head, Keith and John Ries, "Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade," *American Economic Review*, 2001, *91* (4), 858–876.

— **and Thierry Mayer**, "Gravity equations: Workhorse, toolkit, and cookbook," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 131–195.

Institute, Fraser, "Fraser Freedom Index," https://www.fraserinstitute.org/economic-freedom/dataset. accessed on March 20th, 2021.

Larch, Mario, Joschka Wanner, Yoto V Yotov, and Thomas Zylkin, "Currency unions and trade: A PPML re-assessment with high-dimensional fixed effects," *Oxford Bulletin of Economics and Statistics*, 2019, *81* (3), 487–510.

Maskus, Keith E and Lei Yang, "Domestic patent rights, access to technologies and the structure of exports," *Canadian Journal of Economics/Revue canadienne d'économique*, 2018, *51* (2), 483–509.

— **and Mohan Penubarti**, "How trade-related are intellectual property rights?," *Journal of International economics*, 1995, *39* (3-4), 227–248.

Melitz, Marc J, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *econometrica*, 2003, *71* (6), 1695–1725.

Moser, Petra, "How do patent laws influence innovation? Evidence from nineteenth-century world's fairs," *American economic review*, 2005, *95* (4), 1214–1236.

_ , "Do patents weaken the localization of innovations? Evidence from world's fairs," *The Journal of Economic History*, 2011, *71* (2), 363–382.

Piermartini, Roberta and Yoto Yotov, "Estimating trade policy effects with structural gravity," 2016.

Rose, Andrew K, "Do we really know that the WTO increases trade?," *American economic review*, 2004, *94* (1), 98–114.

Santacreu, Ana Maria, "Intellectual Property Rights, Technology Transfer and International Trade," *Technology Transfer and International Trade (July 12, 2021)*, 2021.

Silva, JMC Santos and Silvana Tenreyro, "The log of gravity," *The Review of Economics and statistics*, 2006, *88* (4), 641–658.

Weidner, Martin and Thomas Zylkin, "Bias and consistency in three-way gravity models," *Journal of International Economics*, 2021, p. 103513.

WIPO, "PCT Membership," 2021. data retrieved from WIPO website, `https://www.wipo.int/pct/en/pct_contracting_states.html`.

A Model

I consider the decision problem for a firm in a monopolistically competitive industry in country i that has developed an idea for producing a new product.

I assume there are N potential national markets in which the product can be sold, with CES preferences in market j given by

$$U_j = \left(\int_{\Omega_j} d_j(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \tag{7}$$

where Ω_j is the set of varieties sold in market j and ω indexes the varieties in that market. The differences in the value of innovations across firms gives rise to a structure of heterogeneous firm monopolistic competition as in Melitz (2003).

If the firm has exclusive right to sell the product in market j, it will earn revenues at time t given by

$$r_{ijt}(p(\omega)) = p(\omega) d_j(p(\omega)) = \left(\frac{p(\omega)}{P_{jt}} \right)^{-\sigma} E_{jt} \tag{8}$$

where E_{jt} is expenditure and $P_{jt} = \left(\int_{\Omega_j} p(\omega_{jt})^{1-\sigma} d\omega_j \right)^{1/(1-\sigma)}$ the price index in market j. Revenues in j will depend on the number and price of competing varieties in the market, as well as on the overall expenditure level of the destination market.

I assume that output is produced with a constant marginal cost of $\frac{c_{it}}{z(\omega)}$ per unit, where c_{it} is the cost of a unit of factor input and $z(\omega)$ the units of output per unit of input for for variety ω. The quality of the innovation is reflected in the value of z, which can be interpreted as the efficiency of the process for producing the product or as the quality of the new product. I simplify by assuming that the firm knows the value of z at the time it makes the patenting decision in market j. Letting $\tau_{ij} \geq 1$ be the iceberg transport cost of getting the product from i to j, the flow of firm profit will be $(p - c_i \tau_{ij}) d_{ijt}$. Profit is maximized by choosing a price $p_{ijt} = \frac{\sigma c_{it}}{1-\sigma}$, which yields profits of

$$\pi_{ij}(z) = \left(\frac{c_i \mu \tau_{ij}}{P_j}\right)^{1-\sigma} \frac{E_j}{\sigma} = z^{\sigma-1}\overline{\pi}_{ij} \tag{9}$$

where $\mu = \frac{\sigma}{\sigma-1}$ is the markup and $\overline{\pi}_{ijt}$ captures the profitability of market j for a representative firm from i at time t

The existence of a patent gives the firm the right to exclude firms from using the product or process innovation in the market where it is patented. This protection would apply to both domestic and foreign firms that might attempt to infringe on the patent, since foreign firms would be excluded from exporting goods that infringed on the patent. If the firm enters market j without a patent, it has a probability α_j per unit time that it's product will be copied and the monopoly lost. If the firm obtains a patent, the probability that the patent will be infringed on and copied (without compensation) is $\alpha^P < \alpha$. Letting T_j denote the length of patent protection and assuming an expected growth rate of profits γ_j, the gain from obtaining a patent in market j will be

$$\Delta^P_{ijt}(z) = \int_0^{T_j} \pi_{ij}(z)\left(e^{-(r+\alpha^P_j+\gamma_j)t} - e^{-(r+\alpha_j+\gamma_j)t}\right) dt = z^{\sigma-1}\overline{\pi}_{ij}\Psi_j \tag{10}$$

where $\Psi_j = \frac{(r+\alpha_j+\gamma_j)(1-e^{-(r+\alpha^P_j+\gamma_j)T_j}) - (r+\alpha^P_j+\gamma_j)(1-e^{-(r+\alpha_j)T_j})}{(r+\alpha_j+\gamma_j)(r+\alpha^P_j+\gamma_j)} > 0$ captures the expected gain in profits with patent protection in market j. The expected growth rate depends on both the growth rate of the market as well as the strength of patent protection. The strength of patent protection will depend on both the duration of protection, T_j, and the strictness of enforcement, which determines α^P_j.

A firm will patent in market j if the expected gain from patenting exceeds the expected cost. I denote the cost of obtaining a patent in market j by F_{ij}. This cost will consist both of fees for obtaining and maintaining the patent and the transactions costs of filing the patent. Since transactions costs may depend on the similarity of the origin and destination markets in factor such as language and institutions, I treat these the cost of a patent as pair specific.

A firm will choose to patent if $\Delta^P_{ijt}(z) \geq F_{ij}$, patenting of a new product from i will be patented in j if

$$z \geq \underline{z}_{ij} \equiv \left(\frac{F_{ij}}{\Psi_j \overline{\pi}_{ij}}\right)^{\frac{1}{\sigma-1}} = \frac{c_i \mu \tau_{ij}}{P}\left(\frac{E_j \Psi_j}{\sigma F_{ij}}\right)^{\frac{1}{1-\sigma}} \tag{11}$$

The critical value for patenting is increasing in the level of source country costs, c_{it} increasing in the bilateral trade and patenting frictions, $\tau_{ij} F_{ij}^{1/(\sigma-1)}$, and decreasing in the profitability of market j, $\Psi_j E_{jt} P_{jt}^{\sigma-1}$.

A.1 Bilateral Patent Flows

I obtain the bilateral patent flows by aggregating over the firms decisions. Firms in each country engage in R&D that generates ideas for new products. The quality of the new product ideas is assumed to be described by a Pareto distribution, $G_i(z) = 1 - \left(\frac{z}{b_i}\right)^{-\theta}$ for $z \geq b_i$. [29]

[29] The value of patents has been measured based on surveys, citations, and licensing revenue. As surveyed in Hall and Harhoff (2012) Hall and Harhoff (2012), These measures all suggest that the value of patents is highly skewed, so their value is typically modeled using a Pareto or lognormal distribution.

The probability that a firm from i patents a new idea in country j is

$$1 - G(\underline{z}_{ij}) = \left(\frac{\underline{z}_{ij}}{b_i}\right)^{-\theta} \tag{12}$$

The parameter b_i is the lower bound of the distribution, so a higher productivity for country i innovators is reflected in a larger value of b_i. The parameter θ captures the shape of the distribution - a higher value is associated with a lower spread of the distribution. In order for the expected profits from a market to be bounded, we must have $\theta > \sigma - 1$.

Assuming that the output of new products in country i is ι_i, the number of patents from i to j, N_{ij} will be

$$N_{ij} = \iota_i \left(\frac{\underline{z}_{ij}}{b_i}\right)^{-\theta} = \iota_i \left(\frac{b_i P_j}{\mu c_i \tau_{ij}}\right)^{\theta} \left(\frac{E_j \Psi_j}{\sigma F_{ij}}\right)^{\frac{\theta}{\sigma-1}} \tag{13}$$

The number of patents will be an increasing function of the destination market size, $\frac{E_j \Psi_j}{P^{1-\sigma}}$, an increasing function of the innovative capacity and cost in the source country, $\iota c_i^{-\theta}$, and decreasing in the level of bilateral trade and patenting frictions, $\tau_{ij} F_{ij}^{1/(\sigma-1)}$.

B Data Assignment

For 93% of application families, the applicants for the priority application are from a single location. In these cases, I merely take the applicants' address country to be the family origin. In the cases where there is discord amongst the applicants' addresses, I sort applicants by "type", e.g. company vs. individual. Specifically, I regard the type of applicant to indicate the origin in the following order:

- Company, Government
- Hospital, University, Non-profit
- Individual

The rationale being that an application for which both a company and an individual are listed as applicants, the company represents a more clear indication of where the innovation originated, if not in the literal sense maybe in the financial sense.

This process requires very complete data, as such it reduces the available sample yet further. There are many families that have applicant addresses not for the priority application but for other subsequent applications. I have decided not to try to reclaim those entries to avoid the need to reconcile the idiosyncratic ways each country has to record the ”type” of person. Importantly, Japan (a very active patenting country) almost exclusively records persons as individuals. If I were to determine the origin using the entire family, I would run the risk of underrepresenting Japan as an innovator merely due to its data recording practice (e.g. between a German “company” and a Japanese “individual”, Germany would be assigned as the origin). By limiting the origin assignment to within a single application (the priority application) I bypass this potential for different data recording methods to produce unrepresentative origin assignments.

As a means of alleviating the significant requirements of having applicant addresses, I allow the destination of the priority application to serve as the family origin for singleton families. The assumption here being that an innovation only ever patented in a single country is likely to have originated in that country.

B.1 Destination Assignment

While most of the assignments of destination are straightforward. A significant complication lies in the EU. A single EU application can be submitted to get protection in several "designated states" within the EU. The application fee is higher depending on how many designated states are indicated (though supposedly more efficient than applying in each country separately). For applications submitted as "European patent applications," I treat it as an individual application in each of the designated states.