Cross-border Patenting, Globalization, and Development*†

Jesse LaBelle
FRB of St. Louis

Inmaculada Martinez-Zarzoso
University of Göttingen
Universitat Jaume I

Ana Maria Santacreu
FRB of St. Louis

Yoto V. Yotov
Drexel University

January 2024

Abstract

We build a stylized model that captures the relationships between cross-border patenting, globalization, and development. Our theory delivers a gravity equation for cross-border patents. To test the model's predictions, we compile a new dataset that tracks patents within and between countries and industries, for 1980-2019. The econometric analysis reveals a strong, positive impact of policy and globalization on cross-border patent flows, especially from North to South. A counterfactual welfare analysis suggests that the increase in patent flows from North to South has benefited both regions, with South gaining more than North post-2000, thus lowering real income inequality in the world.

JEL classification: F63, O14, O33, O34.

Keywords: Cross-border Patents, Gravity, Policy, Globalization, Development.

*The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†We thank Eric Bond, Elijah Coleman, , Beata Javorcik, Kristina Manysheva, and seminar participants at Vanderbilt University, the Gravity Workshop in The Hague, the Applied Economics Meeting and the Goettingen School of Development Economics. We thank Ashley Stewart for excellent research assistance. Contact information: LaBelle: Northwestern University, E-mail: jesselabelle2029@u.northwestern.edu; Martinez-Zarzoso: University of Göttingen and Universitat Jaume I, E-mail: martinei@uji.es; Santacreu: FRB of St. Louis, E-mail: am.santacreu@gmail.com; Yotov: School of Economics, Drexel University; ifo Institute; CESifo, E-mail: yotov@drexel.edu.

1 Introduction

Cross-border patenting plays a crucial role in fostering development globally. It involves seeking patent protection for innovations in multiple countries, enabling inventors and companies to protect their intellectual property rights (IPR) beyond national borders. The connections between cross-border patenting and development can manifest through various channels.

First, cross-border patenting facilitates the exchange of technology and knowledge across countries. When innovators protect patents in multiple jurisdictions, they are compelled to disclose detailed information about their inventions. This disclosure not only ensures transparency but also facilitates knowledge dissemination and contributes to the spread of advanced technologies, helping the development process of countries that might not have had access to such innovations otherwise. According to the UNCTAD, technology transfers are crucial to promote structural transformation and productive capacity in the global south and to facilitate sustainable development within the framework of the Sustainable Development Goals (SDG), more so in the XXI century, with technological change taking place at higher speed and mostly within global networks (UNCTAD, 2023).

Second, countries that prioritize cross-border patenting create a favorable environment to attract foreign direct investment (FDI). Companies and investors seek regions where IPR are well-protected, encouraging them to invest in research and development (R&D). FDI brings advanced technologies into developing countries, stimulating economic growth. Indeed, according to the World Bank:

> "Today, FDI is not only about capital, but also –and more important– about technology and know-how, [...] International patterns of production are leading to new forms of cross-border investment, in which foreign investors share their intangible assets such as know-how or brands in conjunction with local capital or tangible assets of domestic investors." (The World Bank, 2015)

In this context, there has been a growing interest in understanding the determinants of cross-border patenting and the role that IPR play in shaping these decisions.

Against this backdrop, we make the following contributions. First, we build a new comprehensive database that tracks cross-border patenting flows and citations across and within countries and industries from 1980 to 2019. Second, from a methodological perspective, we develop a stylized model that (i) captures the importance of patent transfers from developed (North) to developing (South) countries and (ii) delivers a structural gravity equation for cross-border patents, which resembles familiar and intuitive gravity models from physics and trade. Third, on the estimation front, we translate our structural model into an estimating equation for cross-border patents by capitalizing on the latest developments in the empirical gravity literature on trade, migration, and FDI. Fourth, from a policy perspective, we offer a series of estimates of the effects of various policy determinants on the cross-border patent flows, as well as estimates of the effects of globalization, which we define as trends that go beyond observable policies.[1] Finally, we use our theory and partial estimates to show that the exchange of cross-border patents has been mutually beneficial to developed and developing countries but has benefited developing countries disproportionately more after 2000, thus decreasing real income inequality in the world.

Our novel *International Patent and Citations across Sectors* (INPACT-S) database tracks cross-border and domestic patent flows across industries over four decades.[2] INPACT-S is more comprehensive than other publicly available datasets along five key dimensions: (i) It encompasses a wider array of patent authorities, offering a full view of global patent activity; (ii) it provides industry-specific bilateral data, allowing to do sectoral analysis; (iii) it captures a greater number of patent applications through imputation methods; (iv) it includes comprehensive data on cross-country and cross-sector citation data; and (v) it includes consistently constructed data on cross-border and *domestic* patents. The domestic dimension of INPACT-S is crucial for our analysis, as it enables us to obtain estimates of the effects of globalization, which go beyond policy and which cannot be

[1]On the policy front, we examine the role of trade agreements that require increasing protection of IPR, which has been a key topic of discussion in multilateral trade negotiations since the World Trade Organization (WTO) Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS).

[2]The INPACT-S dataset is available upon request by filling this questionnaire.

identified with data on cross-border patents only.[3] We describe in detail the methods that we use to construct INPACT-S, and we highlight the main features of our dataset in Section 2, where we also devote a subsection to comparing INPACT-S with related datasets.

We then build a stylized model of cross-border patenting in a global economy that encompasses two countries: North and South. Innovators in North invest resources to create new ideas. Ideas serve as blueprints to produce a new intermediate good. Innovators license their ideas to producers in both North and South in exchange for royalty payments. Due to imperfect enforcement of IPR in South, innovators apply for patent protection there before licensing. Patenting is costly but offers protection to North innovators from imitation in South. The number of patented technologies depends on the value of an innovation, the probability of imitation, and the cost of patenting. The value of innovation, in turn, depends on how profitable South is at commercializing products produced with North technology (i.e., their size, productivity, etc.). The focus on technology licensing in patenting is motivated by its role in cross-border knowledge transfer. Innovators in North, through licensing, enable Southern firms to access and implement new technologies. Patent protection and technology licensing are interconnected, providing legal incentives for innovators to disclose inventions. While our analysis centers on technology licensing, the framework extends to other motivations such as trade and foreign direct investment (FDI), offering insights applicable to the interplay between patents, trade, and FDI. Our model yields a structural gravity equation for international patenting, which guides our empirical analysis.

We use our model to investigate the impact of changes in globalization or trade policy on cross-border patenting through their impact on the probability of imitation. Guided by our theory and capitalizing on the rich dimensionality of INPACT-S, we employ established developments from the gravity literature on trade, migration, and FDI to specify an estimating gravity equation for cross-border patents. We estimate our model with

[3] In combination with cross-border patents, the use of domestic patents offers a series of additional benefits for identification, e.g., of the effects of non-discriminatory policies that may target international patenting or, more broadly, any country-specific policy or characteristic that may impact cross-border patents and domestic patents differentially.

the Poisson Pseudo Maximum Likelihood (PPML) estimator, which takes into account zero patent flows and potential heteroskedasticity of our patent data, which may render OLS estimates inconsistent. We also employ a rich set of fixed effects, including source-time and destination-time fixed effects, which absorb all possible country determinants of patent flows, as well as pair fixed effects, which control for all time-invariant determinants of cross-border patents.

To highlight several important aspects of our data and identification strategy, we develop the estimation analysis in three nested specifications. First, we impose common (across countries) globalization effects without accounting for any policy effects. The corresponding estimates reveal that, during the sample period, globalization forces have led to an increase in the number of cross-border patent flows in our sample by about 84%. Second, we allow for heterogeneous effects depending on the level of development. We find that the impact of globalization is the strongest for patent flows from North to South – our "North to South" globalization estimate is seven times larger than the corresponding average effect. The second largest impact of globalization on cross-border patent flows is for the group "South to South", followed by "North to North", and we do not find evidence for significant globalization effects for "South to North". The latter is consistent with and reinforces our theoretical assumptions.

Third, in addition to allowing for heterogeneous globalization effects, our main specification introduces a series of policy variables, including regional trade agreements (RTAs), which may or may not include technology provisions, the TRIPS agreement, and the Patent Cooperation Treaty (PCT). Similar to the analysis of the effects of globalization, we also allow for heterogeneous effects of each of the policy variables across the four bilateral income groups. We draw three main conclusions based on this analysis: (i) Policy efforts have been effective to promote cross-border patent flows; (ii) the policy effects have been heterogeneous across policies (e.g., for RTAs vs. PCT) and depending on the direction of patent flows (e.g., for "North" to "North" vs. "South to North"); and (iii) the policy covariates in our econometric model fully account for the change in cross-border patent flows across all groups, except for "North" to "South".

We conclude the estimation analysis with a battery of robustness experiments to test the sensitivity of our main findings and to highlight some additional dimensions of our new database. Three main findings stand out from this analysis. First, overall, our main conclusions regarding the impact of policy and globalization on cross-border patent flows are reinforced by the additional robustness experiments. Second, we find that some of the effects of the "standard" gravity variables, e.g., distance and common official language, are similar for trade flows and cross-border patents. However, we also find opposing effects for other gravity variables, e.g., contiguity and colonial relationships. While not important for our current purposes, we found these results interesting. Finally, our sectoral analysis (i) revealed heterogeneous effects, implying that sound policy analysis of the determinants of cross-border patents should be performed at a disaggregated level, and (ii) reinforced the message that, for RTAs to facilitate cross-patenting between rich and poor countries, the agreements must contain specific chapters on IPR and innovation.

Armed with the partial estimates, we use our theoretical model to answer two questions: "What would have been the trajectory of cross-border patenting from North to South between 1995 and 2018 if the globalization trends that we estimated had remained at their 1995 levels?" and "What are the welfare implications of the increase in patent flows from North to South in our model?" For simplicity, and consistent with our theory and empirical results, we focus on the impact of globalization on patent flows from North to South. To answer these questions, we start by calibrating the model using data on cross-border patenting flows, R&D intensity, and bilateral trade flows. Then, we study the effect of globalization on cross-border patenting, innovation and welfare.

We draw the following main conclusions based on our counterfactual analysis. First, in the absence of the globalization effects that we estimated, cross-border patenting would have been 43% lower in 2018. Second, both North and South have gained from the transfer of patents across international borders. However, the gains for South were larger after the 2000s, implying that cross-border patenting has led to a decrease in the real income gap between the poor and the rich countries in the world.

Related Literature. This paper is related to several strands of literature. First, it is related to studies on the connection between IPR, patents, and development (Helpman, 1993; Lai, 1998; Lai and Qiu, 2003; Kwan and Lai, 2003; Yang and Maskus, 2001; Branstetter et al., 2007, 2011; Tanaka and Iwaisako, 2014; Diwan and Rodrik, 1991). While some research finds stronger patent protection boosts innovation in developed nations at the expense of developing ones (Helpman, 1993; Grossman and Lai, 2004), others find that strong IP protection in developing countries can increase growth and development (Kwan and Lai, 2003). Hoekman and Saggi (2007) find in a theoretical framework that North and South trade agreements with technology provisions can be beneficial to the South if it has reached a certain level of IP protection. Bond and Saggi (2020) develop a North and South model to study the South's incentive for patent protection. Santacreu (2022) finds that improvements of IP that are associated with trade agreements have a positive impact on technology transfers from North to South through licensing. Our paper focus on the impact of both globalization forces and trade and IP policy on cross-border patenting.

Second, our paper is related to a strand of literature studying the connections between IPR and technology transfer. Maskus (2000) studies the connections between IPR and international trade, innovation, and growth. Keller (2004) studies the impact of international technology diffusion through various channels on innovation, growth, and development. Glass and Saggi (1998) study how technology transfer helps close the technology gap. We contribute to these two strands of literature by performing an empirical exploration of the impact of globalization and IPR reforms that are part of deep trade agreements on cross-border patenting and, hence, knowledge transfer.

Third, it relates to a recent literature studying the role of RTAs with IP provisions on bilateral flows. Martínez-Zarzoso and Chelala (2021) and Arregui and Martínez-Zarzoso (2022) find that better IPRs increase trade in goods, especially high-tech exports from developed to developing countries and international patenting. Santacreu (2022) finds that regional trade agreements with IP provisions have a positive effect on international technology licensing, especially from developed to developing countries. Hémous et al.

(2023) study optimal patenting policy in an open economy. More closely related to our work, Coleman (2022) and Howard, Maskus, and Ridley (2023) explore the impact of trade liberalizing treaties and treaties strengthening IPR on cross-border patent flows. We contribute to the existing literature by investigating the distributional impacts of globalization forces and RTAs with IP provisions on cross-border patenting. Our approach involves constructing a comprehensive dataset. Additionally, we employ a stylized model as a guide for our empirical analysis. By exploring these effects across various levels of development, our study provides insights into the impacts on cross-country inequality. Finally, more broadly, our paper is related to the recent literature that studies the impact of deep trade agreements on various economic outcomes.[4]

While the channel emphasized in our model is international technology licensing, there are other channels of technology transfer that have been studied in the literature, such as FDI, or trade. Our paper relates to recent studies on the connections between cross-border patenting and trade. Brunel and Zylkin (2022) find that cross-border patenting has a positive effect on exports of the protection-seeking country to a destination country, suggesting that innovators patent in countries where they anticipate more trade. Using disaggregated French firm-level data, De Rassenfosse et al. (2022) find that patent protection at the product-destination level increases exports on that product and at that destination country. The paper is also related to recent work documenting the impact of trade liberalization on innovation. Using firm-level data on patent applications from PATSTAT, Coelli, Moxnes, and Ulltveit-Moe (2022) find that tariff cuts increase patenting at the country level. Different from their approach, we investigate the effect of globalization on cross-border patenting, focusing both on the origin and destination of patents.

The rest of the paper is organized as follows. Section 2 describes the INPACT-S dataset. In Section 3, we develop our theoretical model (in Subsection 3.1), and we translate it into an estimating equation (in Subsection 3.2). Section 4 presents our main estimation findings (in Subsection 4.1) and offers counterfactual analysis for the impact of patents on welfare and income inequality (in Subsection 4.2). Section 5 concludes with

[4]See `https://datatopics.worldbank.org/dta/table.html`.

directions for future work. A Supplementary Appendix includes results and discussion from a series of robustness experiments and additional specifications.

2 The INPACT-S Database

Our new *International Patent and Citations across Sectors* (INPACT-S) database tracks international and domestic patent flows and citations across countries and industries over the period 1980-2019.[5] In this section, we describe the methods that we used to construct INPACT-S (in Subsection 2.1); we highlight some of its key features by documenting several patterns of international patenting across industries and over time (in Subsection 2.2); and we show that INPACT-S is more comprehensive and has several key advantages over existing related datasets (in Subsection 2.3).

2.1 Constructing the INPACT-S Database

To construct INPACT-S, we rely primarily on the PATSTAT Global Autumn 2021. Using patent-level data from PATSTAT, we compute the number of patent applications from a country of origin (i.e., the residence of the inventor or the owner of the technology) to an application authority at the International Patent Classification (IPC) level — 4-digit IPC codes — for the period 1980-2019. We account for both the applicant and the inventor, respectively.[6] We then use concordance tables developed by Lybbert and Zolas (2022) to transform IPC codes into industry codes—ISIC Rev 3 2-digit. The result is a dataset that contains 91 patent authorities, 213 countries of origin, 40 years, and 31 ISIC Rev 3 2-digit codes. We describe in detail how we construct the dataset next.

We proceed in several steps. From the raw PATSTAT data, we use Structured Query Language (SQL) to pull appln_id, person_id, earliest_pat_publn_id, appln_auth,

[5]We also construct a dataset of citations across country-sector pairs, which can be used to study knowledge flows across countries, as in Cai, Li, and Santacreu (2022). The details are relegated to an online appendix.

[6]The inventor country of residence reflects the country of origin of the innovation, whereas the applicant country of residence reflects the ownership of the intangible. Not all applicants are necessarily inventors, as the inventor may simply develop a new technology while ownership resides with the firm that employs or funds her. For the same reason, being an inventor does not automatically make one an applicant. Importantly, in PATSTAT, firms can be applicants but cannot be inventors.

person_ctry_code, appln_filing_year, publn_nr, publn_nr_original, publn_auth, publn_kind, ipc_class_symbol from tables tls201_appln (table containing the bibliographical data elements of the application), tls207_pers_appln (table linking the applicants/inventors of the most recent publication to an application), tls206_person (table with identifying information on the applicants/inventors), tls209_appln_ipc (table containing the IPC classifications of an application), and tls211_pat_publn (table containing information about patent publications). These variables give us a raw dataset that reports, for each patent, the jurisdiction where the application was filed, the country of the applicant(s)/inventor(s), individual application identifying numbers, and the full, disaggregated IPC class associated with each patent.

Importantly, we restrict our data to application type "A," which in PATSTAT represents basic patents, and we do two separate pulls, one to get all persons who are inventors and another to get all persons/entities who are applicants. Moreover, rather than restricting the sample to the first patent in a family, we consider every patent from the same family. There is merit to analyzing only the first patent in the family, as one can get a better sense of breakthrough innovation, since all further patents in that family are a variation of that initial invention. However, our goal is to create a more broad dataset that captures all innovation flows across the world because we seek to understand why patents are filed where they are. To this end, where the last patent in a family is filed holds just as much importance to us as where the first patent was filed.[7]

We make several adjustments to the raw data, which we explain next. First, we aggregate the IPC classifications to the 4-digit level. Second, in many instances, one application may feature multiple applicants/inventors from different countries. Similarly, for a majority of applications, a single patent belongs to multiple IPC technology classifications. To avoid counting the same applications multiple times for different origins/classifications, we employ a fractional counting method for both technology class and origin country. For example, if an application has four inventors, one from the US and three from Canada, then this will be counted as 0.25 patents from the US and 0.75 from Canada, as opposed

[7]The EPO defines a patent family as "A patent family is a collection of patent applications covering the same or similar technical content."

to four different applications. To ensure consistency, we implement built-in checks and crosscheck with the OECD, which also relies on a fractional method.

We use the same idea to avoid counting one patent that falls into multiple IPC classification as multiple different applications. If, as in our example above, the IPC classifications of the patent are G06F 1/04, G06F 1/16, and G08B 1/02, then 0.67 of the application is assigned to G06F and 0.33 is assigned to G08B. This means that in the case of the four inventors described above, the Canadian inventors receive credit for 0.75 of the patent, and 0.67 of that is assigned to the G06F classification. This results in a total of 0.5 patents assigned to the Canadian G06F class.[8]

Third, in several cases, applications are filed to regional patent authorities covering two or more countries rather than a single country. This is a decision made at the individual level. In some cases, applicants may opt for the cheaper upfront cost of applying to just one or two European countries, and others may decide to go the more expensive route and apply to the European Patent Office (EPO) as a whole, which is cheaper than applying to many countries individually. As recognized by WIPO, the major regional authorities are African Regional Intellectual Property Organization (ARIPO), EPO, Eurasian Patent Organization (EAPO), Gulf Cooperation Council (GCC) Patent Office, and Organisation Africaine de la Propriété Intellectuelle (OAPI).[9] Under these jurisdictions, applicants can send one application to these authorities for a singular granting process and receive the possibility of protection in all fully ascended member states.

We attempt to take the regional patent authority application totals and disperse them in favor of individual member country applications. To this end, we make the reasonable assumption that not all member states of an authority are attracting patent applications equally. For example, it is likely that far more applications filed with the EPO are intended to be used to protect IP in a large, traditionally innovative country, such as Germany, than in a smaller member, such as Slovenia or Liechtenstein. Therefore,

[8]With the fractional method, the first step is divide into pieces based on the IPC codes. Then, the goal is to divide one more time and split the pieces evenly amongst applicant origins. The number of observations yielded through this method is 1,275,488. If we only split by IPC codes (ignoring applicant origins), the data would have 1,276,794 observations. On the other hand, if we only split by applicant origin (ignoring IPC codes), the data would contain 4,724,619 observations.

[9]`https://www.wipo.int/patents/en/topics/worksharing/regional-patentoffices.html`

when measuring the main destinations of cross-border patents, equating all patents to the EPO to count as one for each and every member state would paint a skewed image of technology transfer. This approach could make small countries that are part of a large regional authority seem like more of a technology destination than they are in reality.

To address this issue, we employ a weighted-dispersion method in which we allocate patent applications, from an origin to a regional authority, across the individual member states of that regional authority. We base this dispersion probability on the share of direct patent applications from each origin country to each individual member state in that same year and technology class. To visualize this point, imagine a hypothetical regional patent authority, UKESPDEU, which consists of only the United Kingdom, Spain, and Germany. Suppose that applicants from Australia filed 100 patents in the textiles industry with UKESPDEU in 2022. Suppose that, also in 2022, Australian applicants filed 25 textile patents directly in Germany, 10 textile patents directly in Spain, and 5 textile patents directly in the United Kingdom. Out of these 40 directly filed patents, Germany received 62.5%, Spain received 25%, and the United Kingdom received 12.5%. These shares serve as the probabilities of the intended final destination of patents filed to the regional authority. We use these probabilities as our weights to disperse out the patents filed to UKESPDEU. Following this method, dispersing the 100 Australian textile patents filed to UKESPDEU and adding them to the direct totals would result in 87.5 patents to Germany, 35 patents to Spain, and 17.5 patents to the United Kingdom.

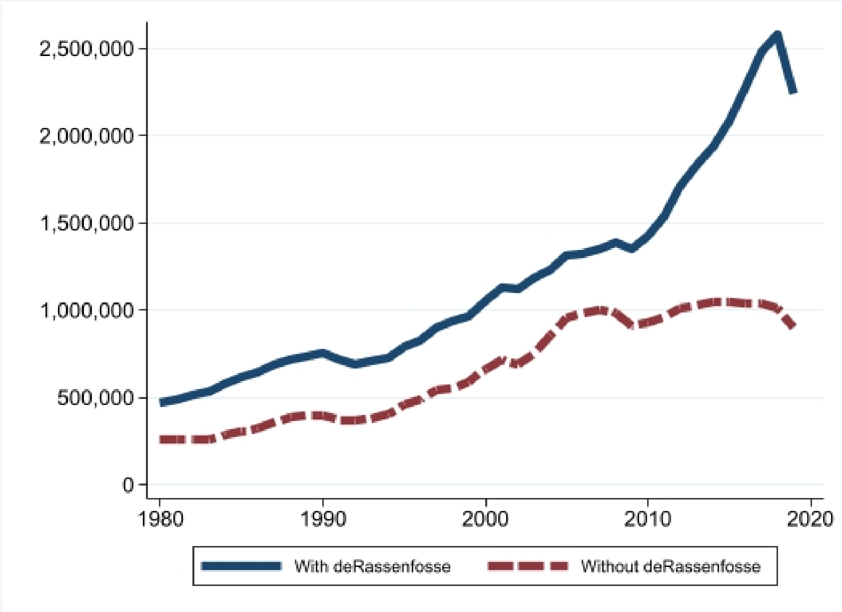
Fourth, we address a commonly discussed problem of PATSTAT database. Since PATSTAT is maintained by the EPO, they are unable to edit the data voluntarily provided to them by other authorities that are sometimes lacking in detail. This results in a prevalence of missing data in a number of categories, including in the country of the applications' applicant(s)/inventor(s), as documented by De Rassenfosse, Kozak, and Seliger (2021). We follow two steps for imputing blank origin countries. In the first step, we use the SQL code provided by De Rassenfosse, Kozak, and Seliger (2021) to impute missing values in the raw PATSTAT data. Before imputation, there are over 26 million applicants with a known origin from 1980-2019 and nearly 24 million inventors in our dataset; after applying their method, we have over 46 million applicants and 44 million inventors.

Figure 1 showcases the differences in known origins before and after imputation for each year in our sample. They use familial linkages between worldwide applications to impute the origin that is missing, based on data found in related patents filed elsewhere. Patents for the same technology are often filed in more than one jurisdiction (or even in the same jurisdiction for a slightly different but related technology). One authority may report incomplete information on the origin of a patent, but another authority may report more complete information for the same (or similar) technology in the same family. PATSTAT provides data that can be used to link priority filings with subsequent filings across the globe, making it possible to take information from related patent applications to impute the missing information, which is precisely what their provided code does. In brief, their method can be summarized as the following: If the information is not available on the patent application, search for the information from direct equivalent patents in the same family. If the information cannot be found on those direct equivalents, search for the information in subsequent filings in the same patent family. This continues on until all possibilities are exhausted.

The De Rassenfosse, Kozak, and Seliger (2021) method, although impressively comprehensive, is unable to account for all missing origins. In the second step, instead of simply dropping the remaining blank origin data, we use the aggregate bilateral data from WIPO to disperse the remaining "origin missing" applications. At the current stage, after applying all the edits stated above, our dataset contains authority, IPC 4-digit class, year, and origin. However, for every authority, in each year, in each ISIC industry there exists a blank country of origin with some patents attributed to it. Our goal is to assign all these remaining patents to origin countries rather than simply lose that data.

One possibility would be to follow a method similar to the one described above for the dispersion of regional authorities' applications. That is, dispersing applications based on shares of the applications, which are already assigned. However, this method might be a biased way of dispersing the "missing origin" applications. Some origins have more robust

Figure 1: Imputing missing values with De Rassenfosse, Kozak, and Seliger (2021)



Note: This figure shows, for each year, the difference between the applicants we have with known origins before and after using the De Rassenfosse, Kozak, and Seliger (2021) imputation method.

patent families for the De Rassenfosse, Kozak, and Seliger (2021) method to pull from. In addition, some authorities report better data than others, and these authorities receive applications from different origins at different rates. For example, Japan reports Japanese origins very well but is less reliable on reporting cross-border patents. Additionally, in recent years, China rarely reports origin countries at all to the EPO. As a result, using shares derived from our existing dataset would be reinforcing established biases in the data.

To account for this problem, we instead use the WIPO aggregate bilateral data as a proxy. We take the authorities from WIPO and compute the share of total patents for each authority that originate from each origin country for a given year. We then, as with the regional authorities described above, apply those probabilities to the "missing origin" data, and distribute them based on these WIPO weights. A key assumption with this approach is that the probabilities are assumed to be constant across all technology classes for each origin/authority/year relationship. Roughly 9% of our observations by applicant are dispersed with this method and 11% of our patents by inventor.

Finally, the 4-digit IPC technology classes are converted into ISIC rev. 3 2-digit industries using a crosswalk that can be found in Goldschlag, Lybbert, and Zolas (2016).[10] Our patent numbers for each technology class are multiplied by the probability weights provided and then summed by industries to give us a bilateral patenting dataset by country and industry rather than technology class.

2.2 Salient Features of Patent Flows

INPACT-S uncovers several interesting features of patent flows across countries, industries, and over time. Among other facts, we highlight the rise of Asia as both an origin and a destination of patent applications over the past decades. Asian countries increasingly becoming destinations for patent applications suggests a flow of technology from traditionally innovative countries. This exchange has the potential to drive development in Asia, as the countries gain access to advanced knowledge, methodologies, and technologies from developed countries. Indeed, we also observe that more Asian countries are becoming origins of patents, implying they are becoming more innovative themselves.

The origins of innovation. Figure 2 shows the worldwide distribution of patent applications per million people filed by (i) domestic inventors inside the country—domestic applications—in the upper panel, and (ii) domestic inventors to the world—cross-border applications—in the bottom panel, throughout the decade of the 2010s.[11]

The figure shows that innovation is concentrated in a few countries, mainly in Europe, the United States, and Eastern Asia. While Europe and North America have traditionally been innovation hubs, our data show the rise of Eastern Asian countries as new innovators. In terms of domestic patent applications, China stands out as the main innovator. Indeed, 37% of all patent applications being filed around the world in the 2010s can be attributed to Chinese domestic applications. The other leaders of total domestic applications were Japan, the United States, South Korea, and Germany in that order.

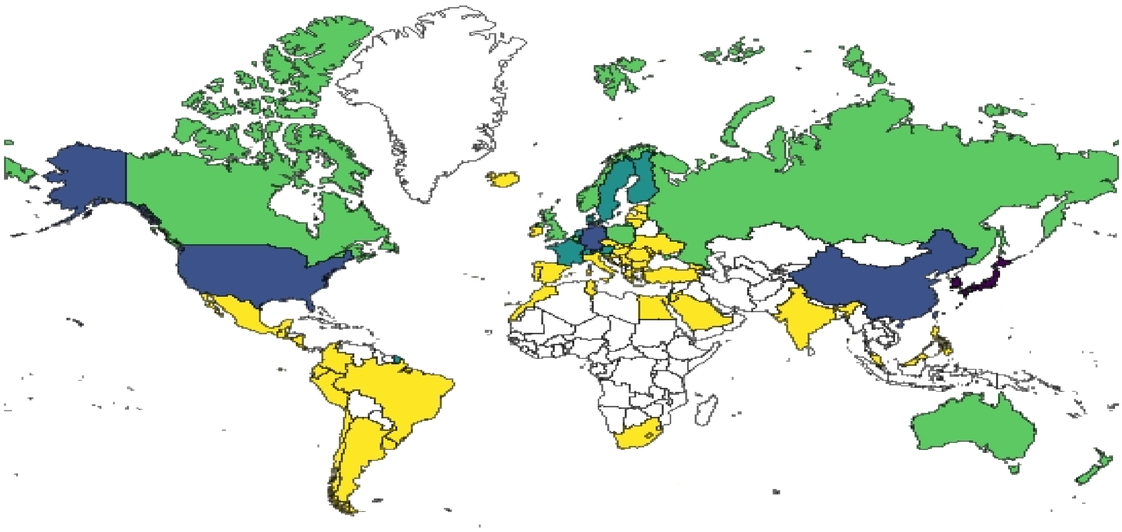
The rise of Eastern Asian countries on the world innovation stage centers around four

[10] `https://sites.google.com/site/nikolaszolas/PatentCrosswalk.`

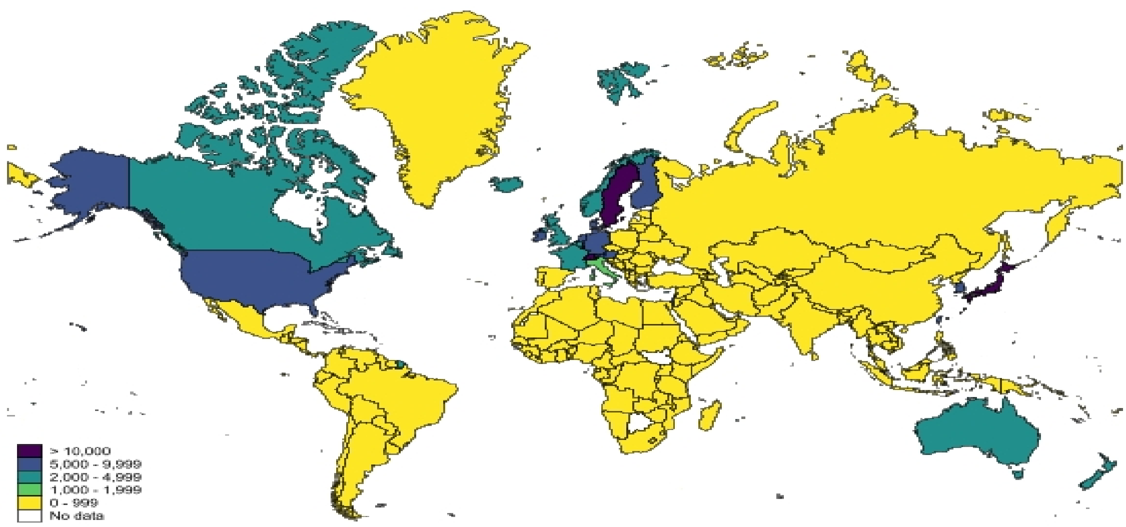
[11] Population is calculated by taking the average across the decade.

Figure 2: Origins of Innovation

(a) Domestic patent applications



(b) Cross-border patent applications



Note: The upper panel (a) shows the number of domestic patent applications per million people in the 2010s; the bottom panel (b) shows cross-border patent applications per million people in the 2010s. Blank countries do not have data available as authorities.

countries: Japan, South Korea, China, and Taiwan.[12] Japan and South Korea are more traditionally innovative countries, with their technology sectors dating back decades, while China and Taiwan have become new powerhouses of innovation, with a growth in the number of domestic patent applications between 1995 and 2018 by a factor of 13 in Taiwan and by a factor of 162 in China. China's explosion in terms of domestic patenting is unprecedented and even suspicious. In fact, there is reason to believe that this remarkable growth can be attributed to China's generous patent subsidy programs. However, on January 27, 2021, the China National Intellectual Property Administration (CNIPA) announced that these subsidies are to be phased out by 2025. It is yet to be seen the impact this will have on China's incredible patent growth.

Aside from the domestic market, these 4 countries have also become important sources of cross-border patent applications. Japan, likely due to the age of its technology sector, dominates the other eastern Asian countries in the number of patent applications filed abroad. However, South Korea has become more prominent in the international patent market beginning in the 90s, followed by Taiwan and China in the 2000s. Interestingly, China's unprecedented domestic patent growth has not been replicated on the international level in terms of the total number of cross-border patent applications filed, but the growth rate has. Though their cross-border patent applications are dwarfed by their domestic applications, China has still seen an increase by a factor of 230 in terms of cross-border applications filed from 1995 to 2018. These trends indicate an increase in these countries' presence in terms of innovative activity.

If we isolate the analysis to cross-border patent applications, applications filed by each origin country to the world excluding domestic applications, the picture looks slightly different. The main discrepancy lies in China, where Chinese innovators seek protection mainly domestically. Indeed, out of all patent applications to the world during the 2010s, only 1.5% are accounted for by Chinese cross-border patents. Again, this discrepancy is an indication of the market intervention introduced by the Chinese government in the

[12]Eastern Asia also includes Hong Kong and Mongolia but their overall values are small and inconsequential so we focus on the four mentioned.

late 2000s that sought to incentivize patent applications through a subsidy.[13] During the decade, the main innovators seeking protection abroad are Japan, the United States, Germany, and South Korea. From the 1990s to the 2010s, we find that Japan, South Korea, and Taiwan are the countries that have experienced the largest increase in the number of cross-border patent applications per million residents filed.[14]

To better illustrate the emergence of new regions as origins of patent applications, we group the countries in our dataset into 17 regions.[15] Figure 3 shows the evolution of patenting across the different regions. In the upper panel, we show the total number of cross-border patents filed by the countries in each world region, whereas in the bottom panel we show the same but for the countries that make up Eastern Asia.[16]

What stands out in the upper panel is Eastern Asia catching up to North America in terms of foreign applications filed in the mid-2000s and maintaining the lead since. Also notable is the speed at which Eastern Asia caught up to North America, closing the gap with the United States very quickly after the turn of the century. No other region comes close to matching Eastern Asia's growth over this time frame. From the bottom panel, we can see that this growth was largely due to Japan and, to a lesser extent, South Korea. However, while Japan, South Korea and Taiwan have plateaued in recent years, China has begun their own rapid growth as inventors of cross-border patents.

If we focus on patent applications to other countries within the region, we observe a strong bias of Eastern Asian countries to file cross-border patents that remain within the region (Figure 4). Beginning in 1990s, Eastern Asian countries began rapidly filing patents to other countries in the region, and after 2000, around the time of China's ascension to the WTO, this grew even faster while most other regions either stayed steady or increased just moderately over this span. This is consistent with an overall

[13] `https://www.stlouisfed.org/on-the-economy/2018/february/china-overtaken-us-terms-innovation`.

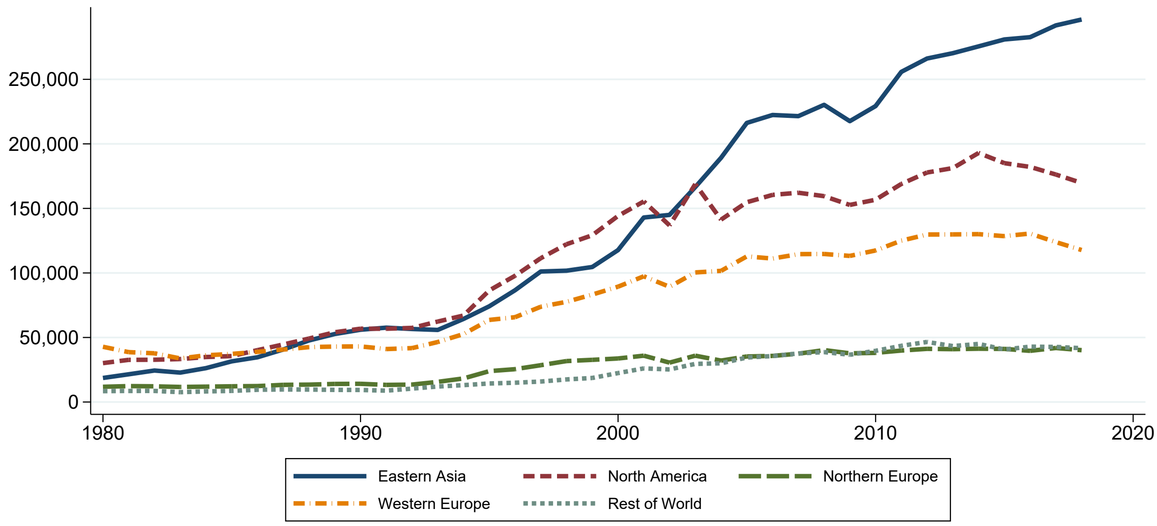
[14] This is excluding countries that are commonly labeled as "Tax Havens", which typically saw incredible patent growth over this period.

[15] We use regions as defined by the UN `https://unstats.un.org/unsd/methodology/m49/`: Australia and New Zealand, Central Asia, Eastern Asia, Eastern Europe, Latin America, Melanesia, Micronesia, Northern Africa, Northern America, Northern Europe, Polynesia, South-eastern Asia, Southern Asia, Southern Europe, Sub-Saharan Africa, Western Asia, Western Europe.

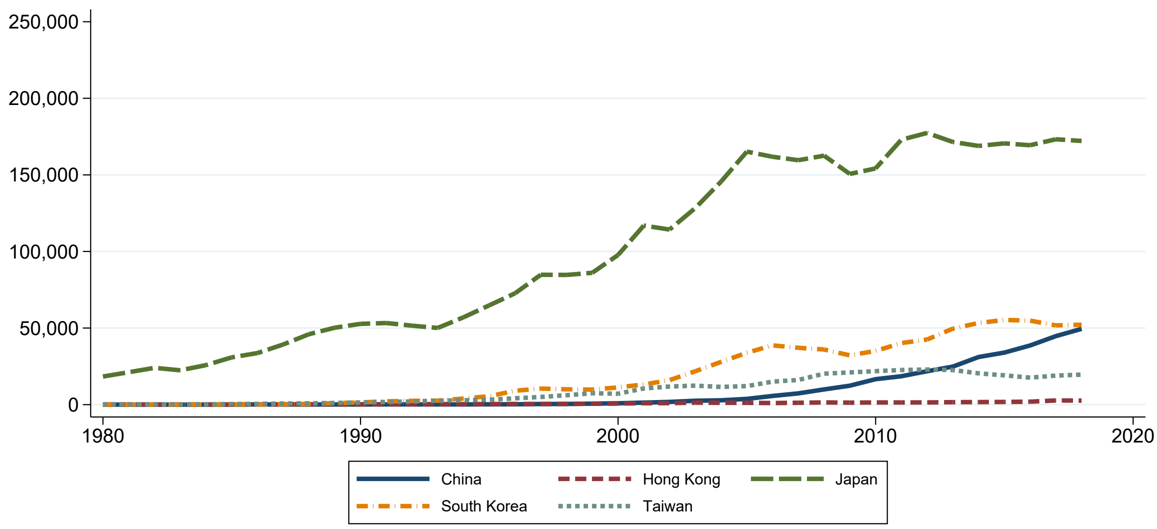
[16] Cross-border patents are determined at the jurisdiction level rather than the regional level. A patent from South Korea to Japan occurs in the same region but is a cross-border application.

Figure 3: Patent Evolution by Region of Origin

(a) World Regions



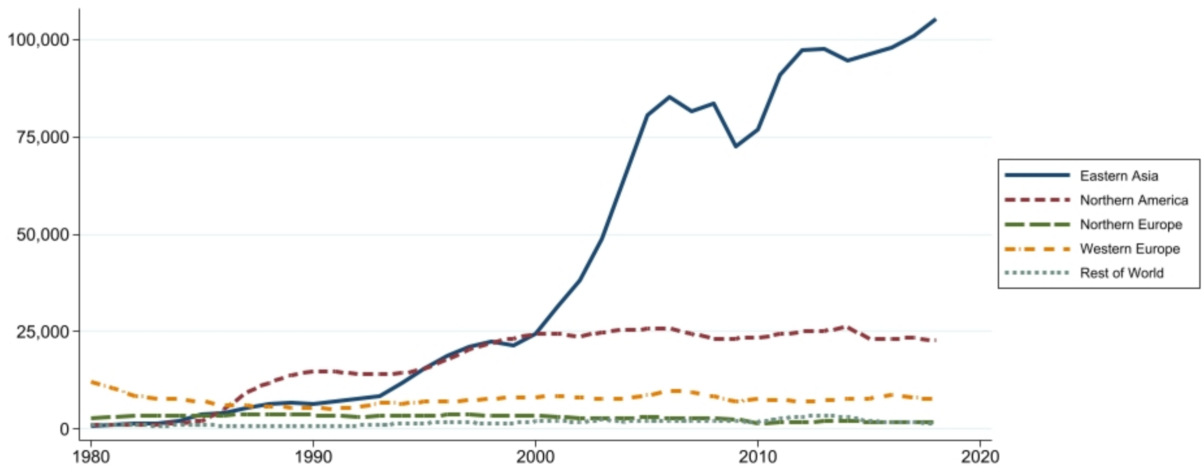
(b) Eastern Asia



Note: The upper panel (a) shows cross-border patent applications filed by the countries of different regions of the world; the bottom panel (b) plots patent applications filed by each country in eastern Asia.

trend towards eastern Asia in the global patent market. As seen in Figure 5, the world received more patents as a whole, Eastern Asian countries saw larger increases in the number of patents received compared with the rest of the world. This is explored further in the next section.

Figure 4: Regional Bias in Patenting Applications

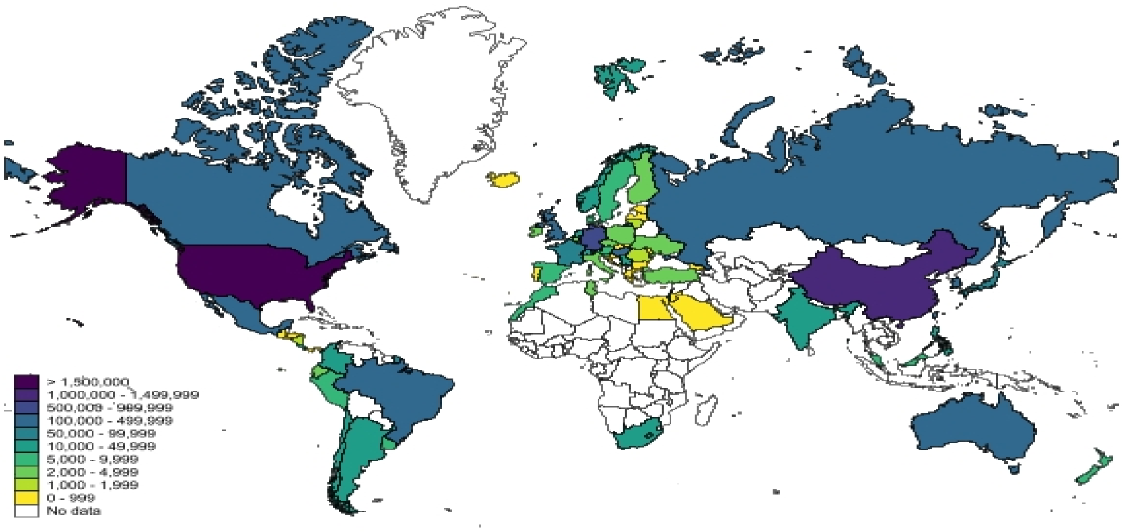


Note: This figure shows the number of cross-border patents filed by countries of a region which were filed elsewhere in that same region. The Rest of the World is an aggregate where we first compute this intra-region number for each region individually and aggregate them together.

The Destinations of Innovation. So far, we have documented the rise of Asia as an innovation hub. In this section, we investigate the following question: Where are innovators seeking protection for their ideas? Figure 5 shows that Eastern Asia has risen as a destination of cross-border patent applications in addition to being an innovator. This suggests that the region may be seen as a competitor destination to traditionally innovative countries, such as the US. By the 2010s, the USA and China have dominated as destinations where inventors seek protection of their IP, being the only countries to attract more than a million cross-border patent applications in the decade?

Additionally, South Korea attracted the 4th most cross-border patents, while Taiwan attracted the 7th most. There are two possible reasons for this: either these countries are becoming more innovative and competing with western innovation such that innovators want to ensure their technology is protected from imitation here, or typically innovative

Figure 5: Main Destinations of Cross-border patents in the 2010s



Note: This figure shows the number of cross-border patent applications received in the2010s. Blank countries do not have data available as authorities.

countries are doing more business in these countries, leading to an increased need for ensuring business assets are protected. Later, in the model section, we explore which of these effects dominates.

International Patenting Across Industries. Next, we leverage the industry dimension of the data and ask the following question: In what industries are innovators seeking international protection? Taking the United States as the world innovation leader, we find that patent applications from the United States to the world are concentrated in a few industries: Chemicals, Computers and electronics, and Medical and optical equipment. These are also R&D-intensive industries in that they account for most of the R&D spending and number of patents being created around the world. Second, we find that nine countries account for more than 80% of cross-border patent applications filed by United States applicants to the world: China, Canada, Great Britain, Australia, Germany, South Korea, Taiwan, Brazil, and Mexico. In Table 1, we report the share of patent applications from the US to each of these countries across five of the most R&D-intensive industries.

The table shows that about one-third of the patents filed by the US in Mexico, one-fourth of the patents filed in Canada and South Korea, and one-fifth of those filed in

Table 1: Industries of Patents filed by the US, Post 2000

Destination	Chemical Mfg	Computing	Machinery n.e.c.	Medical /Optical Equip	Radio/TV/ Comms Equip
Australia	27%	12%	5%	14%	3%
Brazil	30%	9%	7%	11%	3%
Canada	25%	10%	7%	13%	3%
China	16%	13%	9%	15%	10%
Germany	9%	13%	14%	18%	7%
UK	16%	21%	5%	18%	6%
Korea	23%	17%	4%	15%	11%
Mexico	32%	7%	6%	9%	3%
Taiwan	19%	17%	3%	15%	19%
ROW	17%	21%	6%	16%	7%

Notes: The table reports the share of patent applications from the US to each of the countries across five of the most R&D-intensive industries. R&D intensity is computed as the proportion of patents generated by each industry in relation to the overall number of patents across all industries.

China, UK, and Taiwan are in the chemical industry. However, in the case of Germany, US inventors seek protection mainly in the medical and optical equipment industry. Also notable is Taiwan, where 19% of US patents in Taiwan after the turn of the century were in the radio, television, and communication equipment industry, far higher than shares to other countries. This is notable because of Taiwan's importance in the semiconductors industry and the fact that this industry comprised just 7% of US patents to the rest of the world over this same period. Additionally, 14% of patents filed in Germany were in machinery, which is more than double its share of US patents filed in the rest of the world. Differences in patent applications across industries and countries could be explained by supply chain linkages requiring countries to seek protection in a particular industry, depending on the particular position in the supply chain.

2.3 Comparison with Alternative Datasets

Our novel INPACT-S dataset complements and improves on existing patent data publicly available from the United States Patent and Trademark Office (USPTO), the Organisation for Economic Co-operation and Development (OECD), and the World Intellectual

Property Organization (WIPO), along several dimensions.

While the USPTO only accounts for patents filed in the United States, the OECD database is slightly more comprehensive, including patents that have been filed in the United States (USPTO), in the European Union (EPO), and under the Patent Cooperation Treaty (PCT). In contrast, our dataset covers 91 patent offices around the world. This extension is important in capturing the innovation trends observed in the past four decades, in addition to the rise of new players in the knowledge sector.

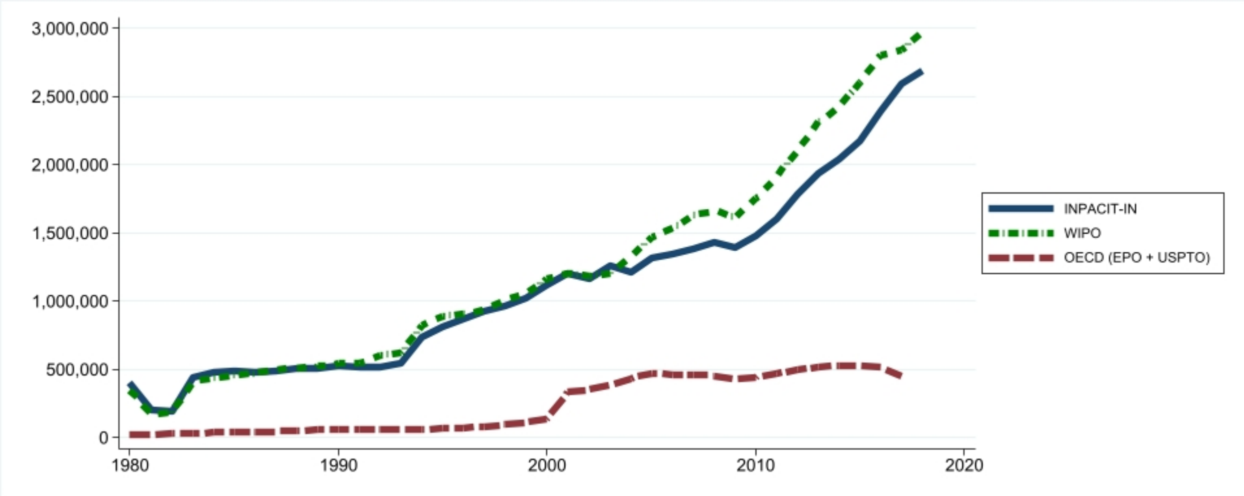
The WIPO dataset is closer to ours, as it includes patent applications filed in all patent offices for which data are available, but it does not report the data at the industry level and differs in the way it imputes some data points, as we elaborate on later. Hence, our dataset is more comprehensive than other existing publicly available datasets on international patenting flows. Beyond just these improvements, we provide data on citations across country-sector pairs, which allows us to compute a measure of quality-adjusted patent applications, as explained in the Appendix, used in the robustness tests.

Figure 6 shows the comparison of our dataset with the OECD and WIPO. For this comparison, we use the patent applications by applicant counts.[17] One important difference between the WIPO and our dataset is that the WIPO dataset does not have an industry dimension. Therefore, for comparison we must aggregate across our industries by bilateral relationships and year in our final dataset. We also aggregate OECD patents filed to USPTO and EPO.

Aside from a slight divergence in the early 2000s, our method matches the aggregate trends in the WIPO data extremely closely. However, there are a few important differences between the methods used to derive our data and the WIPO data. First, as described above, we find it unrealistic to assume each country in a regional authority is equally attracting patents to that authority. WIPO instead chooses to count the patents for these authorities by assigning one patent to each jurisdiction in the region. Furthermore, rather than fractionally dispersing out the patent equally amongst the applicant(s)/inventor(s), WIPO chooses to assign the patent to the country of the first

[17] By definition, fractional counting creates identical totals in aggregate whether summed by applicant or inventor despite the fact individual bilateral relationships may differ.

Figure 6: Comparisons with similar datasets



Note: The green line represents the aggregate WIPO world patent totals by year while the blue line represents INPACT-S totals after using the methods described above. The red line represents OECD, EPO, and USPTO patents. OECD has a group of applications filed under the PCT, but since they do not provide indicators as to where those applications are going we leave them out. Including those just increases their total slightly each year.

applicant, under-counting the number of patent applications originating in some jurisdictions. Since we are focused on understanding where innovators from different countries seek protection for their inventions, we see value in recording the origin of every applicant/inventor rather than just one.[18]

The discrepancies between our dataset and the OECD dataset reflect the additional patent offices we capture with ours, while the OECD restricts the sample to patent applications filed to just the EPO and USPTO patent offices.[19] Note that these differences are increasing over time, as new countries begin attracting more patent applications and are becoming new innovation powerhouses.

Having identified some of the main differences between our data and the OECD and WIPO datasets, we want to make sure that the choices and assumptions made in the construction of our dataset are reasonable and do not yield aggregate numbers that differ significantly from those reported by these more established datasets.

We begin by reporting in Table 2 the correlation between INPACT-S and publicly

[18] https://www.wipo.int/ipstats/en/help/

[19] The OECD also has PCT patents, but those provide no value in discerning bilateral patent trends.

available OECD and WIPO datasets. For comparison to the OECD, we take the raw PATSTAT data and employ our fractional method to patent applications filed to the EPO and the USPTO, the only jurisdictions available in the OECD dataset. Moreover, we do not impute any missing country codes. Here, we are only attempting to measure the accuracy of our fractional counting of the raw patent data.

Table 2: Correlations

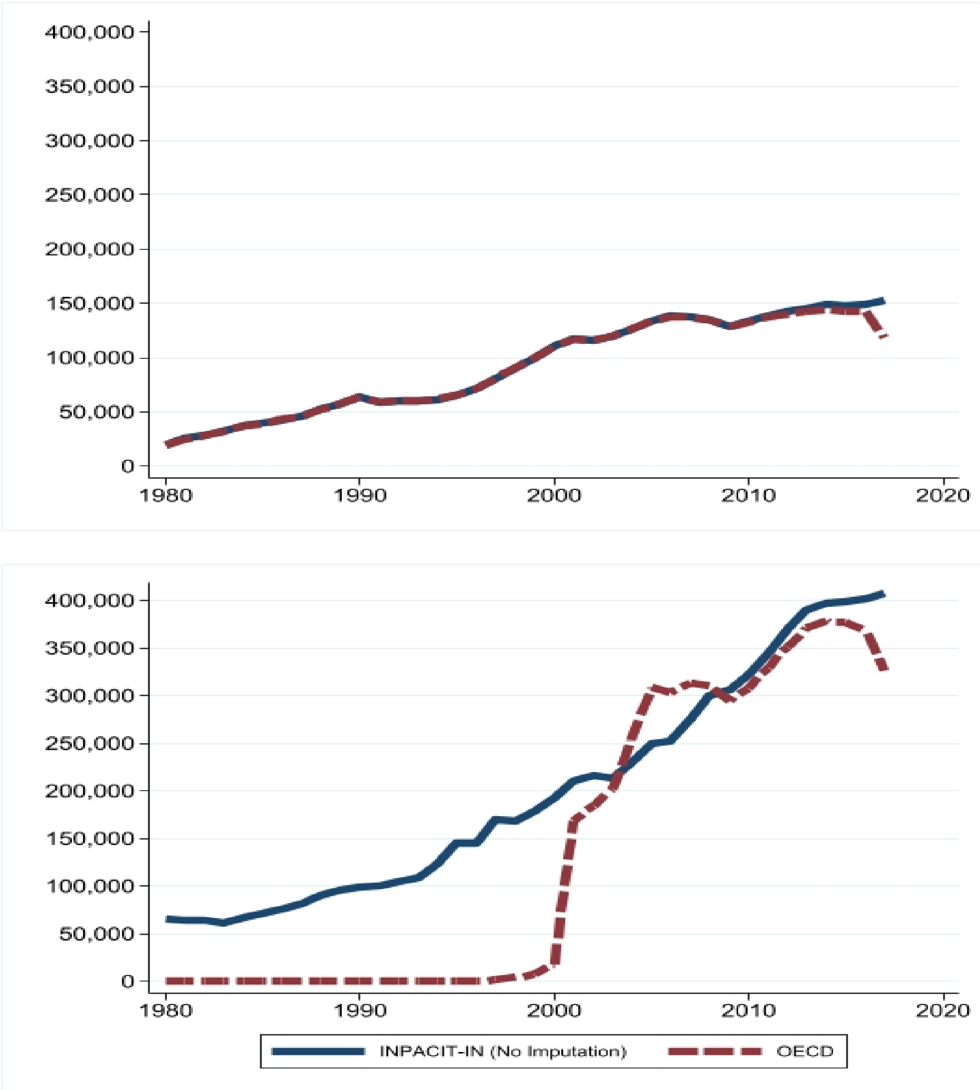
	OECD		WIPO	
	(1)	(2)	(1)	(2)
	Full Sample	2010-2018	Full Sample	2010-2018
Applicants	0.89	0.99	0.99	0.99
Inventors	0.91	0.99	0.99	0.99

Notes: The table reports the correlation between INPACT-S and publicly available OECD and WIPO datasets. To compare with the OECD, we take only patents that are filed with the USPTO and EPO and apply our fractional method. We do not impute any missing values.

For the full sample period, 1980-2019, our data are consistent with the OECD data, with a correlation of around 90% for international patenting by both applicants and inventors. When we restrict the sample to just 2010-2018 the correlation is nearly 100% for both. The reason is twofold: (i) Patent data have a lag in reporting and are only reliable after a few years, so dropping 2019 helps clean some of the noise, and (ii) the data provided by the OECD that cover patents filed to the USPTO are exceptionally poor prior to 2010.

Figure 7 shows the evolution of patent applications to the EPO and USPTO over the period of analysis using our data and those provided by the OECD dataset. In the upper panel, we observe that, overall, our dataset perfectly tracks world applications to the EPO from 1980 until very recently, when the OECD reports a sharp decline not found in our data. In contrast, our dataset captures a steady rise in patent applications to the EPO even in the most recent years. In the bottom panel, we restrict to patent applications to the USPTO as reported by the OECD dataset. The data provided by the OECD are poor prior to 2000 and follow an unrealistic growth trend afterward, while our dataset provides a more realistic growth pattern.

Figure 7: INPACT-S (without imputation) vs OECD



Note: The upper panel shows patents filed to the EPO from the world according to our data (blue line) and the OECD data (red line); the bottom panel plots patents filed to the USPTO filed by the world according to our data (blue line) and the OECD data (red line).

3 Cross-border Patents and Globalization

Motivated by the patterns that we have uncovered and described based on the INPACT-S database, in this section we ask the following questions: “What drives cross-border patenting?” and “What are the implications of the changing cross-border patenting patterns for welfare, development and inequality in the world?” To answer these questions, in Subsection 3.1, we develop a stylized model of imperfect IPR, cross-border patenting, globalization, and development, which guides our empirical analysis. An important byproduct of our theory is the derivation of a gravity model for cross-border patent flows, which enables us to obtain our own partial estimates of the effects of globalization and policy with the INPACT-S database. In Subsection 3.2, we capitalize on the latest developments in the trade, migration, and FDI gravity literature to translate our theoretical model into an estimating gravity equation for cross-border patents.

3.1 A Model of Cross-border Patenting and Globalization

Existing theories of innovation propose that countries engage in international patenting to protect their innovative ideas, especially in destinations where they plan to commercialize these innovations. The likelihood of patenting in a foreign country is influenced by the presence of mechanisms ensuring patent enforcement. The hypothesis is that globalization and improved IPR enforcement contribute to increased cross-border patenting. Our theoretical framework focuses on technology licensing as a motive for patenting. In this context, patent protection incentivizes innovators to disclose inventions, facilitating knowledge transfer across countries and promoting development, particularly in the South. While the emphasis is on technology licensing, the model's insights can extend to other motivations like trade and foreign direct investment (FDI). For example, an alternative perspective could consider patenting prior to exporting, as discussed in Brunel and Zylkin (2022).

We build a mode in which the world is composed of two countries–North (N) and South (S)–that exchange goods and ideas. Time is discrete and indexed by $t \in \{0, \infty\}$. The trade model consists of an Armington framework where each country produces differentiated intermediate goods that are traded internationally, subject to trade costs. The level of technology in each country is determined by the number of intermediate goods it produces. Technology evolves endogenously through two processes: (i) innovation in North and (ii) technology licensing from North to South.

North invests resources in R&D to generate new ideas, which are blueprints for producing differentiated goods. All ideas created in North are licensed to domestic intermediate producers, who pay royalty fees for their use. However, only a fraction of these technologies are licensed to producers in South. Before licensing the technology, North files a patent in South to mitigate the risk of imitation. The number of patented technologies corresponds to the number of technologies that will be licensed.

South imitates a portion of the newly patented technologies from North. In these cases, North does not receive royalty payments. If North did not patent the technology before licensing, it would be imitated with certainty. Patenting ensures that only a fraction of the licensed technologies will be imitated, and this fraction depends on the level of IP enforcement in South. Policies that enhance IPR in South, such as regional trade agreements with IP provisions, will increase cross-border patenting.

3.1.1 International Trade Model: Static Equilibrium

Given the level of technology and trade costs, an Armington trade model determines the static equilibrium.

Final Production. A final producer in each country $n \in N, S$ uses intermediate goods, both domestic and foreign, to produce a final good with a CES technology

$$Y_{nt} = \left(\int_{j=1}^{Z_{Nt}} X_{nN,t}^{\frac{\sigma-1}{\sigma}}(j)dj + \int_{j=1}^{Z_{St}} X_{nS,t}^{\frac{\sigma-1}{\sigma}}(j)dj \right)^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where Z_{it} is the number of products being produced by country $i \in \{N, S\}$, $X_{ni,t}(j)$ is the amount of good j from country i demanded by country n, and σ is the elasticity of substitution across varieties. The demand for intermediate goods from country i by final producers in country n is given by

$$X_{ni,t}(j) = \left(\frac{P_{ni,t}(j)}{P_{nt}} \right)^{-\sigma} Y_{nt}. \tag{2}$$

Intermediate Production. Each intermediate good j is produced by a monopolistic competitive firm subject to

$$y_{nt}(j) = l_{nt}(j). \tag{3}$$

The firms choose labor and prices taking as given the demand by final producers. Prices are set as a constant markup of the cost, which is given by wages. The mark up is $\bar{m} = \frac{\sigma}{\sigma-1}$ and prices are given by

$$P_{ni,t}(j) = \bar{m} W_{it} d_{ni}, \tag{4}$$

where W_{it} is the wage, and d_{ni} is an iceberg transport cost from selling goods from country i to country n. The resulting price of the final good producer is given by

$$P_{nt} = \left(Z_{Nt} p_{nN,t}^{1-\sigma} + Z_{St} p_{nS,t}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} . \tag{5}$$

Trade Shares. Given technology, Z_{nt}, and trade costs, d_{in}, the share of goods that are imported by country i from country n is given by

$$\pi_{in,t} = Z_{nt} \frac{(\bar{m} W_{nt} d_{in})^{1-\sigma}}{P_{it}^{1-\sigma}}. \tag{6}$$

The previous equation is a version of the standard gravity model of trade.

3.1.2 Innovation and Technology Licensing

Technology evolves endogenously through two processes: innovation in North and technology licensing of North's innovations to South. As technologies are non-rivalrous, both North and South can produce differentiated goods using the same technology. After the development of new technologies in North, a fraction of these technologies is licensed to firms in South, allowing them to produce new intermediate goods. Before the licensing process takes place, North files patents for these technologies in South, incurring a cost. South firms then pay royalties to North's innovators for utilizing the licensed technology. However, due to imperfect IPR enforcement in South, a portion of these licensed technologies is subject to imitation. If a technology is imitated, South can produce an intermediate good with that technology without paying any royalties to North. When deciding what fraction of their innovations to patent, innovators consider factors such as the probability of imitation, the cost of patenting, and the expected royalty payments.

In this model, patenting occurs prior to licensing.

Innovation and International Technology Licensing. Innovators in North create new technologies at the rate $\gamma_N(H^r_{N,t})^{\beta_r}$, with $H^r_{N,t}$ representing investment into R&D. Assuming full depreciation of new technologies, the number of newly created technologies every period is

$$Z_{Nt} = \gamma_N(H^r_{N,t})^{\beta_r}. \tag{7}$$

Technologies are blueprints that can be used in the production of a differentiated intermediate good. All ideas are used in the North, and innovators receive the profits from the domestic intermediate producers as royalty payments. The royalty payments received from North intermediate producers are

$$R_{NN,t} = \Pi_{Nt}, \tag{8}$$

where Π_{Nt} are the profits made by North intermediate producers. A fraction of these technologies is licensed to adopters in South, which then use them to produce an intermediate good in exchange for royalty payments. A fraction of licensed technologies is imitated, which implies that North innovators receive only a fraction $\phi_{SN,t}$ of the profits that South makes with North technology. Royalty payments from South to North are given by

$$R_{SN,t} = \phi_{SN,t}(\phi_{SN,t})\Pi_{St}, \tag{9}$$

where Π_{St} are the profits that South makes using licensed North technology and $(1 - \phi_{SN,t})$ is the probability of imitation, which depends on a policy parameter reflecting IPR enforcement, $\tau_{SN,t}$. In the model section, we maintain a general form for this function, which satisfies $\phi'(\tau) < 0$; in the quantitative analysis, we propose a specific functional form.

Cross-border Patenting. Before licensing the technology, North files for a patent in South. Some of these ideas are imitated by adopters in South, depending on the enforcement of IPR there. Innovators in North choose the fraction of technologies to patent in South, $\lambda_{SN,t}$, in order to maximize the value of an innovation,

$$\lambda_{SN,t}\phi_{SN,t}(\tau_{SN,t})V_{SNt} - C(\lambda_{SN,t}), \tag{10}$$

where V_{SNt} is the value of a patented technology that is not imitated, and $C(\lambda)$ is the cost of patenting. The idea behind the probability of imitation component is that if we patent, there is a probability $1 - \phi_{SN,t}$ of getting imitated whereas if we do not patent the probability is 1. If the probability of imitation is very high, the innovator may be indifferent between patenting or not. In that case, all technologies are imitated. Then, the FOC for the fraction of patented technologies is

$$C'(\lambda_{SN,t}) = \phi_{SN,t}(\tau_{SN,t})V_{SNt}, \tag{11}$$

where $V_{SN,t} = \Pi_{St}/Z_{St}$. Assuming the following functional form

$$C(\lambda) = \frac{1}{\xi}\bar{\lambda}(\lambda)^{\xi}, \quad \xi > 1, \tag{12}$$

enables us to express

$$\lambda_{SN,t} = \bar{\lambda}_{SN}^{-1/(\xi-1)} \left(\phi_{SN,t}(\tau_{SN,t})V_{SN,t})\right)^{1/(\xi-1)}. \tag{13}$$

Then, the number of patented technologies is

$$\text{Pat}_{SN,t} = \lambda_{SN,t}Z_{Nt}. \tag{14}$$

We assume that patented technologies are then licensed. Thus, the value of patenting a technology are the royalties received from adopters in South that are paying for the technology, $\Pi_{SN,t}$. The royalty payments are then

$$R_{SN,t} = \phi_{SN,t}\Pi_{St}. \tag{15}$$

Optimal R&D Investment. The first-order condition for R&D investment is derived from the following problem:

$$Z_{Nt}V_{Nt} - P_{Nt}H^r_{Nt}, \tag{16}$$

subject to the expression for $Z_{Nt} = \gamma_N(H^r_{N,t})^{\beta_r}$. The value of innovation can be expressed as

$$V_{Nt} = \frac{\Pi_{Nt}}{Z_{Nt}} + \phi_{SN,t}\lambda_{SN,t}\frac{\Pi_{St}}{Z_{St}}. \tag{17}$$

Finally, the profits are given by

$$\Pi_{nt} = (\bar{m} - 1)W_{nt}L_{nt}. \tag{18}$$

Welfare. Assuming linear utility in consumption and a discount factor β, welfare of country i is given by

$$W_{it} = \sum_{t=0}^{\infty} \beta^t C_{it}, \tag{19}$$

where C_{it} is consumption of country i in period t. Globalization forces and increases in IP protection have beneficial effects for both North and South. By decreasing the probability of imitation, North receives more royalty payments from South, which increases R&D investment and hence output in North. South benefits from more technology transfer from North. As a result, South produces more varieties, which increases output. There are also terms-of-trade effects that impact consumption. All these forces imply welfare effects that are heterogeneous across countries, as we will show in our quantitative analysis.

Market Clearing Conditions. To close the model we impose the following market clearing conditions:

(i) For final output in North: $Y_{Nt} = C_{Nt} + H_{Nt}$; and in South: $Y_{St} = C_{St}$;

(ii) For the labor market: $\bar{m}W_{nt}L_{nt} = \sum_{i \in N,S} \pi_{in,t}Y_{it}$;

(iii) For the total number of varieties in North: Z_{Nt}; and in South: $Z_{St} = \lambda_{SN,t}Z_{Nt}$, where all intermediate goods in South will be produced with foreign technology that has been patented.

A structural gravity equation for cross-border patents. The gravity equation for patenting of N in S is derived from combining

$$\lambda_{SN,t} = \left(\bar{\lambda}^{-1/(\xi-1)} \left(\phi_{SN,t}(\tau_{SN,t})\rho\Pi_{St}\right)^{1/(\xi-1)}\right) \tag{20}$$

and

$$\text{Pat}_{SN,t} = \lambda_{SN,T} Z_{Nt}, \tag{21}$$

which enables us to define the following proposition.

Proposition 1. ***(Gravity for Cross-border Patents.)*** *Cross-border patenting from country N to country S at time t is given by*

$$Pat_{SN,t} = \left(\frac{\bar{\lambda}}{\rho}\right)^{-1/(\xi-1)} \left(\phi_{SN,t}(\tau_{SN,t})\right)^{1/(\xi-1)} \left(\Pi_{St}\right)^{1/(\xi-1)} Z_{Nt}. \tag{22}$$

Unsurprisingly, like most other bilateral flows between countries (e.g., trade, migration, FDI, etc.), cross-border patent flows obey the law of gravity, i.e., the “closer” and the ‘’larger” two countries are, the more cross-border patents they would exchange.

More specifically, according to our theoretical gravity model, cross-border patenting depends on several determinants. First, it depends on the characteristics of the origin of patents, Z_{Nt}, which reflect the number of technologies being developed there. From equation 7, this term is determined by the efficiency of innovation and the investment in R&D by country N. Second, cross-border patenting depends on the characteristics of the destination where patents are filed, $(\Pi_{St})^{1/(\xi-1)}$, which determines the destination country’s profits.

Finally, patenting is impacted by country-pair specific characteristics that depend on policy and globalization factors, as we emphasize in our empirical analysis, $(\phi_{SN,t}(\tau_{SN,t}))^{1/(\xi-1)}$. Bilateral patenting frictions are composed of two terms: (i) globalization forces and (ii) policy parameters that impact the probability of imitation, as discussed in the next subsection. The underlying assumption is that globalization can introduce mechanisms for the enforcement of IPR as more countries improve their IPR regimes. Moreover, there has been a proliferation of international agreements and treaties that address IPR protection. Through such agreements, countries commit to harmonizing their IPR laws and implementing effective enforcement mechanisms. All these forces act to reduce the probability of imitation.[20]

3.2 Estimating Gravity for Cross-border Patent Flows

The objective of this section is to set an econometric model for cross-border patent flows. Guided by our theory (as summarized by equation (22)) and capitalizing on developments from the gravity literature on trade, migration, and FDI, we specify the following estimating equation:

$$\text{Pat}_{ni,t} = \exp[\chi_{i,t} + \pi_{n,t} + \overrightarrow{\mu}_{ni} + BRDR_{ni,t} \times \beta + POLICY_{ni,t} \times \alpha] \times \epsilon_{ni,t}, \quad \forall\, i, n. \tag{23}$$

The dependent variable in equation (23), $\text{Pat}_{ni,t}$, denotes the total number of patents from source i to destination n at time t.[21] To take full advantage of our dataset and to improve estimation efficiency, we allow for patent flows from any source i to any destination n; however, consistent with our theory and as will become clear shortly, in the empirical analysis we will be able to zoom in on the patent flows from North to South.

Since our dependent variable is based on count data, Poisson is the natural choice for our estimator. Moreover, owing to the seminal work of Santos Silva and Tenreyro (2006), the Poisson Pseudo Maximum Likelihood (PPML) has become the workhorse estimator for trade gravity models because of two properties that also apply to our analysis of patent flows. First, due to its multiplicative form, the PPML estimator would enable us

[20] An alternative modeling approach involves considering globalization forces as a factor that decreases the cost of patenting. Whether we incorporate these forces into the likelihood of imitation or factor them into the overall cost of patent acquisition, the fundamental structure of the gravity equation for cross-border patenting would remain consistent.

[21] In the robustness analysis, we also obtain estimates at the industry level.

to include and take advantage of the information contained in the zeros in our sample; i.e., when there is no patent flow from a given country to another. Second, and probably more important, Santos Silva and Tenreyro (2006) demonstrate that the PPML estimator successfully handles heteroskedasticity in trade flows data, which, due to Jensen's inequality, actually renders the corresponding OLS estimates inconsistent.[22]

In addition to cross-border ($i \neq n$) patent flows, our dependent variable also includes domestic ($i = n$) patents. This is important for our analysis for two related reasons. First, the use of domestic patents will enable us to estimate the impact of (de-)globalization, defined broadly as the effects of factors not explicitly captured in our model, on international relative to domestic patent flows.[23] To capture such effects, we introduce to our specification a series of time-varying border indicators, which are defined and discussed in more detail below. In addition, the use of domestic patents will enable us to obtain estimates of the globalization effects for different groups of countries (e.g., poor vs. rich, or South vs. North in our notation), country-specific globalization effects (e.g., for China vs. US), and directional estimates of the effects of globalization (e.g., for patents moving from North to South, from North to North, etc.). The latter is particularly important for our purposes because, consistent with our theory, we would be able to isolate the impact of globalization on cross-border patent flows from North to South.[24]

Turning to the variables on the right-hand side of our estimating equation, specification (23) includes three sets of fixed effects. $\chi_{i,t}$ denotes a full set of source-time fixed effects, which are motivated by and would absorb the theoretical term Z_{Nt}. In addition, these fixed effects will control for and absorb any other source-time-specific characteristics (e.g., institutional quality, national regulations, taxes, etc.) that may impact patent flows. Similarly, $\pi_{n,t}$ denotes a full set of destination-time fixed effects, which are mo-

[22]We refer the reader to Santos Silva and Tenreyro (2021) for a recent summary and discussion of the benefits of PPML for gravity regressions. We view PPML as the appropriate estimator for our purposes. Therefore, we employ it to obtain our main results. However, in the robustness analysis, we also replicate our main findings with the OLS estimator.

[23]Our theory can be generalized to include domestic patenting flows, as well as licensing from South to North. For instance, we could allow both North and South to do innovation with different intensity. Since we are not doing a full quantitative analysis of the model, and instead we are using it to guide our empirical analysis, we leave these extensions for future work.

[24]We refer the reader to Yotov (2022) for a summary of the benefits of using domestic flows in trade gravity regressions.

tivated by the theoretical term, $(\Pi_{St})^{1/(\xi-1)}$, and will control for and absorb any other destination-time-specific characteristics that may impact patent flows. In combination, $\chi_{i,t}$ and $\pi_{n,t}$ will comprehensively account for all possible country-time characteristics on the source and the destination side, thus enabling us to focus on the bilateral determinants of cross-border patents, which are of central interest to us.[25]

The third set of fixed effects that we employ includes country-pair fixed effects, $\overrightarrow{\mu}_{ni}$, which also vary depending on the direction of the patent flows. Motivated by Baier and Bergstrand (2007), and consistent with the average treatment effect methods of Wooldridge (2010), country-pair fixed effects are typically used in trade gravity models to mitigate potential endogeneity concerns with bilateral policies. The same logic should hold for bilateral policies that impact cross-border patents. On a related note, the country-pair fixed effects would absorb and comprehensively control for all time-invariant bilateral patent frictions that are part of the theoretical term $(\phi_{SN,t}(\tau_{SN,t}))^{1/(\xi-1)}$.[26] We also allow for the pair fixed effects in our model to vary depending on the direction of the patents. Baier, Yotov, and Zylkin (2019) demonstrate that this has significant implications for the estimates of free trade agreements, which can be very asymmetric and biased if the pair fixed effects are not allowed to vary depending on the direction of trade flows. Applied to our setting, the use of directional pair fixed effects could be crucial for proper identification of the impact of globalization and liberalization policies for the directional patent flows from North to South.

The next term in specification (23) is particularly important for our analysis. Specifically, $BRDR_{ni,t}$ denotes a vector of time-varying border indicators, which take a value of one for international patents and are equal to zero for domestic patents for each year in

[25]The use of fixed effects would, of course, also absorb the theoretical constant term $\left(\frac{\bar{\lambda}}{\rho}\right)^{-1/(\xi-1)}$.

[26]Egger and Nigai (2015) and Agnosteva, Anderson, and Yotov (2019) demonstrate that the 'standard' gravity variables (e.g., distance, contiguity, common official language, etc.) do well in predicting relative bilateral trade costs, however, they fail to capture the level of bilateral trade costs (e.g., they underpredict the bilateral trade costs for the poor countries and overpredict them for the more developed countries). In the robustness analysis, we obtain estimates of the effects of of the 'standard' gravity variables on the cross-border patent flows to find that while some of them are similar to the corresponding estimates from the trade gravity literature (e.g., the effects of distance and language), others are quite different (e.g., the effects of contiguity and common language). Given our focus on the time-varying bilateral determinants of patent flows, we rely on country-pair fixed effects to obtain our main results, and we delegate the estimates of the standard gravity variables, along with their discussion, to the Appendix.

our sample. The estimates on these dummy variables would capture the impact of globalization, trends that have affected the flow of cross-border patents relative to domestic patents. The flexible definition of the border dummies would enable us to identify the common (across countries) impact of globalization as well as the effects of globalization for specific groups of countries and depending on the direction of patent flows (e.g., from North to South). Bergstrand, Larch, and Yotov (2015) demonstrate that failure to control for such globalization effects may result in severely biased estimates of the effects of trade agreements in gravity regressions (e.g., because they may erroneously capture globalization trends). This may also be the case for the effects of policies that target cross-border patents. Moreover, it is possible that the evolution of cross-border patent flows may be driven by factors beyond the observable covariates in our model. As demonstrated in our empirical analysis, the flexible specification with borders would enable us to account for such effects.

Before we continue to describe the rest of the covariates in specification (23), we discuss two technical items in relation to the globalization dummies in our model. First, we cannot obtain estimates of the impact of globalization without the domestic patents in our sample. If we only had international patents, then the impact of globalization would be controlled for but buried in the country-time fixed effects. Second, due to perfect collinearity with the pair fixed effects in our preferred specification, we cannot estimate all border effects, so we need to drop one of them. Our choice will be the border dummy for the first year in our sample, 1985. Thus, all globalization effects that we will obtain would be relative to those in 1985.

In addition to accounting for globalization trends, we also include in our econometric model several policy variables that were designed to affect international patent flows. Specifically, the vector $POLICY_{ni,t}^{k}$ in equation (23) includes the following time-varying bilateral policy covariates. $RTA_{in,t}$ is an indicator variable that takes a value of one if countries i and n have a regional trade agreement (RTA) in force at time t.[27] In addition, we rely on Martínez-Zarzoso and Chelala (2021) to distinguish between RTAs with and

[27]Data on RTAs have been updated using the code provided by de Sousa (2012), who coded free-trade agreements using WTO data and complementary national sources.

without technology provisions ($RTA\_TECH_{in,t}$ vs. $RTA\_NO\_TECH_{in,t}$, respectively). $TRIPS_{in,t}$ is an indicator for the TRIPS agreement, which has been built using the information provided at the WTO website.[28] Since the generated TRIPS dummy variable is almost collinear with WTO membership, we include only the former in the empirical specification. Finally, $PCT_{in,t}$ is an indicator for membership in the Patent Cooperation Treaty (PCT).[29] Consistent with our theory and similar to our treatment of the effects of globalization, we would allow for heterogeneous effects of each of the policy variables in our model depending on the direction of patent flows (e.g., from North to South).

Finally, following the standard approach in the gravity literature, in our main specifications we cluster the standard errors by country pair, i.e., $Cov[\varepsilon_{int}, \varepsilon_{ind}] \neq 0$ for all t, d, and zero elsewhere. However, motivated by Egger and Tarlea (2015) and Pfaffermayr (2019), in the robustness analysis we also experiment with three-way clustering by source, destination, and year.

4 Estimation Results and Comparative Statics

This section presents the findings from our estimation analysis (in Subsection 4.1) and discusses comparative statics results (in Subsection 4.2).

4.1 Estimation Results

This subsection reports our main findings regarding the impact of globalization and various bilateral policies on the flow of patents across international borders. To highlight

[28]The agreement states that developing countries and those in the process of transformation from a centrally-planned into a market economy would have a five-year transition period, until 2020. Least-developed countries (LDC) were granted a longer transition period of a total of eleven years (until 1 January 2006), with the possibility of an extension. The transition period has been extended three times, and now runs until 1 July 2034, or until a member ceases to be an LDC, whichever comes first.

[29]The PTC is an international treaty concluded in 1970, which was amended in 1979 and modified twice (in 1984 and 2001) and with 157 members in 2022. The members can obtain patent protection simultaneously in all Contracting States by filling and international patent application in the country of which the applicant is a national or resident and it can also be filled in the International Bureau of WIPO. Additionally, it can also be filled in the European Patent Office (EPO), the African Regional Intellectual Property Organization (ARIPO), the African Intellectual Property Organization (OAPI) or the Eurasian Patent Office (EAPO) if countries are members of the agreements and conventions related to each of those patent offices. See https://www.wipo.int/treaties/en/registration/pct for further information.

several important aspects of our data and identification strategy, we develop the analysis in three nested specifications. We start with a simple specification, where we impose common (across countries) globalization effects and do not account for any policy effects. Then, consistent with our theory, we allow for the effects of globalization to vary based on development and depending on the direction of patents (e.g., from North to South). Next, in addition to the heterogeneous globalization effects, we introduce policy variables and allow for their effects to be heterogeneous across the same dimensions as the globalization effects. We conclude the estimation analysis with a series of sensitivity experiments, which are designed to test the robustness of our main findings to alternative estimators and specifications, to generate richer policy implications, and to highlight the sectoral dimension of our new database.

Our first estimates are obtained from the following simple/naïve version of specification (23), which does not include any policy variables and where we impose common globalization effects (β_t) across all countries for each year in our sample:

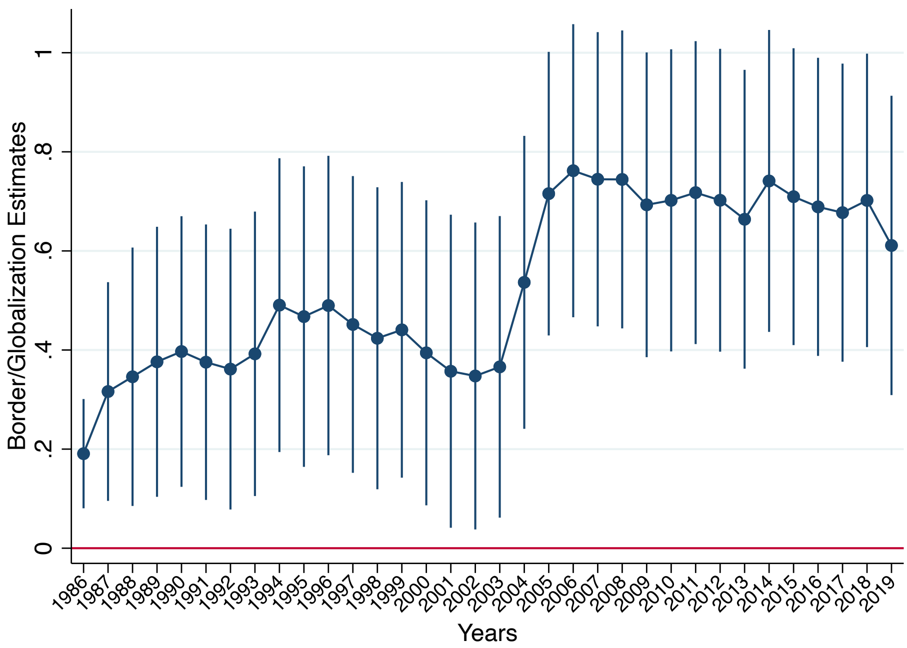
$$\text{Pat}_{ni,t} = \exp[\chi_{i,t} + \pi_{n,t} + \overrightarrow{\mu}_{ni} + \sum_{t=1986}^{2019} \beta_t \times BRDR_{ni,t}] \times \epsilon_{ni,t}, \quad \forall\, i, n. \qquad (24)$$

As noted earlier, due to perfect collinearity with the pair fixed effects ($\overrightarrow{\mu}_{ni}$) in our model, we cannot estimate all border/globalization effects and need to drop one of them. Our choice is to drop the border dummy for the first year in our sample, 1985. Thus, all globalization effects ($\hat{\beta}_t$) that we obtain would be relative to those in 1985.

For expositional purposes (e.g., due to the large number of border estimates that we obtain), instead of using a tabular format, we report our findings in Figure 8. Four findings stand out from the figure. First, all estimates are positive, implying that globalization forces have had significant positive impact on cross-border patent flows, relative to domestic flows. Second, using the border estimate for the last year in our sample ($\hat{\beta}_{2019} = 0.611$) implies that, during the sample period, globalization forces have led to an increase in the number of cross border-patents by 84.2% (std.err. 28.4).[30] Third, we

[30]Calculated as $[exp(0.611) - 1] * 100$, where the standard errors are obtained with the Delta method.

Figure 8: Globalization and Cross-border Patenting, 1985-2019



Note: This figure reports estimates of the impact of globalization on cross-border patent flows, which are obtained from specification (24). See text for further details.

notice two periods of faster growth – in the late 1980s to mid-90s and in the early 2000s – followed by slowdowns. It is tempting to speculate that the latter could be due to China and its accession to the WTO in 2001; however, this is not the case, and we will return to this result later on when we allow for heterogeneous globalization effects.

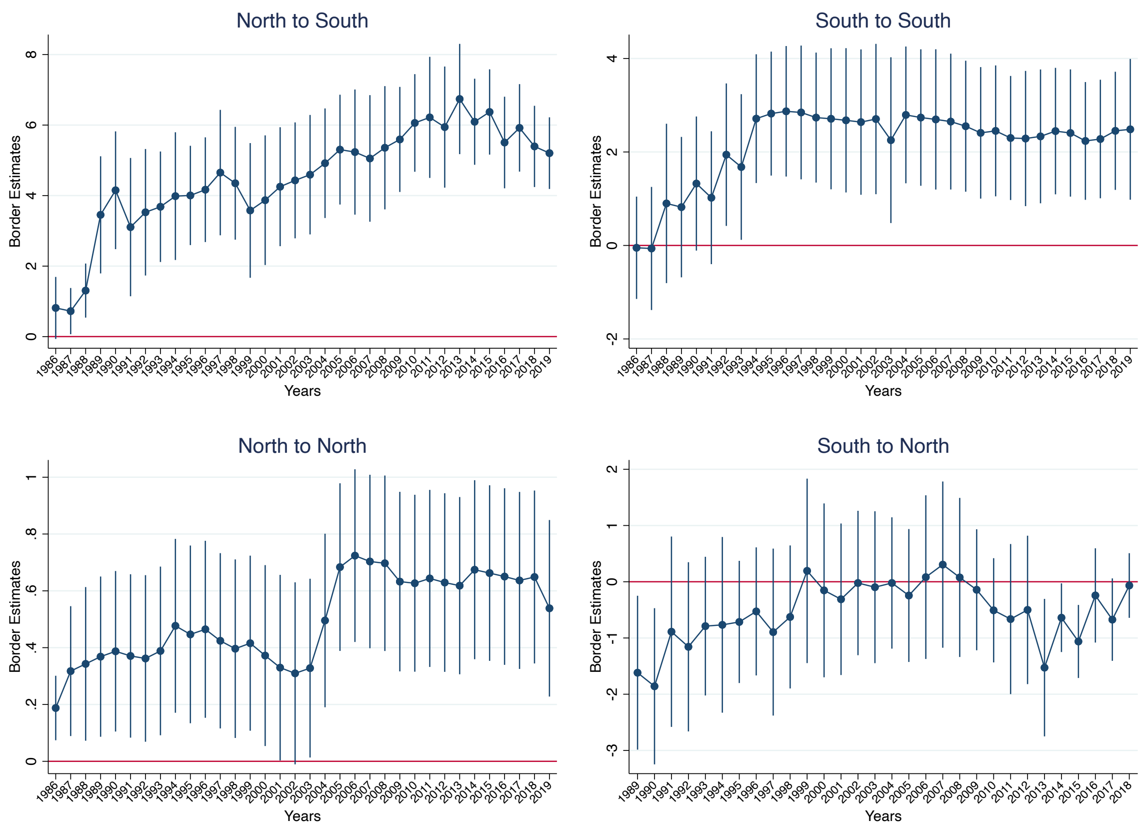
Motivated by our theory, in the next experiment, we refine specification (24) to allow for heterogeneous globalization effects depending on countries' development level and on the direction of patent flows. Specifically, we use the 2000 version of the income classification of the World Bank (WB) to categorize the countries in our sample in two groups – "North", which includes the "high-income" countries and the "upper-middle income" countries from the WB classification vs. "South", which includes the "lower-middle income" countries and the "low income" countries from the WB classification.[31]

[31] We chose the 2000 WB classification for two reasons (an alternative classification was built for 1990). First, because it is more complete. For instance, using data from 1990 fails to capture the emergence of post-Soviet countries like Russia, Ukraine, and the Baltic states. Second, because the year 2000 is closer to the middle of our sample. The classification can be downloaded at this link: https://datacatalogfiles.worldbank.org/ddh-published/0037712/DR0090754/OGHIST.xlsx. In the robustness analysis, we also experiment with two alternative classifications. First, we use all possible income groups categories. We prefer the two-group approach for expositional purposes and because it is consistent with our North-South theory. Second, we define "South" differently by only including the low

Then, based on the two country-specific income groups ("North" vs. "South") and the direction of patent flows, we construct four bilateral income groups of countries, which allow for heterogeneous globalization effects depending on whether the patent flows are from "North to North", "North to South", "South to South", and "South to North". Thus, in effect, we split the common globalization effects from specification (24) into four categories.[32]

Our findings appear in Figure 9. Perfect collinearity does not allow us to obtain

Figure 9: Globalization and North-South Cross-border Patenting



Note: This figure reports estimates of the impact of globalization on cross-border patent flows for four bilateral groups of countries, including "North to South" (top left panel), "South to South" (top right panel), "North to North" (bottom left panel), and "South to North" (bottom right panel). The country groups are based on the income classification of the World Bank, and all estimates are obtained from a single regression, which is based on specification (24) after allowing for heterogeneous effects for each of the four bilateral groups. See text for further details.

income countries in this category.

[32] In the robustness analysis, we also obtain group-specific estimates, i.e., for the impact of globalization on the countries in the North vs. South. The results are consistent with but less informative (since they do not distinguish between the direction of the patent flows) than our main estimates.

estimates of the impact of globalization on the patent flows from "South" to "North" (reported in the bottom right panel of the figure) for all years. However, the sets of estimates for the other three groups are complete. Several findings stand out from Figure 9. First, the new results reveal that the estimates in Figure 8 mask significant heterogeneity of the globalization effects. Second, the impact of globalization is the strongest (by far) for patent flows from North to South. This result is consistent with our theory and justifies our main focus on the cross-border patent flows from North to South. In terms of magnitude, our "North to South" results (in the top left panel of Figure 9) reveal that the estimate of the cumulative impact of globalization (i.e., the estimate for 2019) is seven times larger than the corresponding average index from Figure 8 but decreases in the latest 5 years of the sample.

Turning to the rest of the results in Figure 9, we observe very heterogeneous patterns. The second largest impact of globalization on cross-border patent flows is for the group 'South to South'. Similar to the 'North to South' group, the effects of globalization in this group are quite strong in the late 80s and until the mid-90s; however, they plateau and remain stable since then. As we show in the next experiment, these results are driven by various policies. The pattern that we see for the 'North to North' group, from the bottom left panel of the figure, is very similar to the overall pattern from Figure 8, suggesting that the weight on the 'North to North' group is significantly larger in our sample. Once again, we will see that the increase in the number of patents from 'North to North' is due to the impact of policies, which we introduce and discuss in our next experiment. Finally, based on the estimates for 'South to North' in the bottom right panel of Figure 9, we do not find any evidence for significant globalization effects for this group. This result is in sharp contrast with our findings from the top left panel of the figure, but it is consistent with and reinforces our theoretical assumptions.

The next specification delivers our main estimation results, which are obtained from specification (23), where, in addition to allowing the impact of globalization to be heterogeneous across the four bilateral income groups from our previous specification, we also introduce a series of policy variables, which we expect may affect cross-border patents.

Specifically, we estimate the effects of RTAs, which may or may not include technology provisions, the effects of the TRIPS agreement, and the effects of PCTs. Similar to the analysis of the effects of globalization, we also allow for heterogeneous effects of each of the policy variables across the four bilateral income groups in our sample. To detect possible correlations between the different policies and to decompose their effects, we introduce them sequentially in the four columns of Table 3. The estimates in each column are obtained with the full set of heterogeneous border variables and the full set of fixed effects from specification (23). The dependent variable is always the number of patents, the estimator is PPML, and the standard errors are clustered by country pair.

The estimates in column (1) of Table 3 reveal that, overall, RTAs have been effective in promoting cross-border patent flows. We also note that the RTA effects have been quite heterogeneous across the four bilateral income groups in our sample.[33] Specifically, the RTA effects are the largest for flows from 'North to North', while they are positive but smaller for flows from 'North to South' and from 'South to North', and they are insignificant for flows from 'South to South'. The results in column (2), where we distinguish between the effects of RTAs with and without technology provisions, reveal further heterogeneity. Specifically, we see that the RTAs with technology provisions have benefited all groups, except for "South to South", where we obtain a large, negative estimate.[34] Interestingly, according to our estimates, the patent flows from "South to South" have benefited tremendously from RTAs without technology provisions. The latter have also led to more patent flows between "North" countries.

In columns (3) and (4) of Table 3, we sequentially introduce the effects of the TRIPS agreement (in column (3)) and, in addition, the effects of the Patent Cooperation Treaty (PCT) (in column (4)). Since the introduction of the additional policy variables in each column does not significantly affect the estimates of the variables that were already

[33]The heterogeneous estimates that we obtain suggest that imposing common policy effects may lead to misleading policy implications. We offer such estimates in the robustness analysis in the Supplementary Appendix, where we also investigate the impact on the policy estimates from the use of domestic patent flows and from accounting for globalization forces.

[34]We find this result puzzling and intriguing from a policy perspective. A possible explanation is that the technology provisions from these RTAs may have been too strict to comply with for the poorer nations. Nevertheless, the magnitude of this effects decreases when adding other policy variables and it is only weakly significant.

Table 3: Preferential Agreements and Cross-border Patents

	(1) RTA	(2) TECH	(3) TRIPS	(4) PCT
RTA_S_N	0.150			
	(0.071)*			
RTA_S_S	-0.436			
	(0.303)			
RTA_N_N	0.280			
	(0.042)**			
RTA_N_S	0.126			
	(0.073)$^{+}$			
RTA_TECH_S_N		0.171	0.174	0.170
		(0.060)**	(0.059)**	(0.060)**
RTA_TECH_S_S		-0.838	-0.759	-0.587
		(0.339)*	(0.350)*	(0.234)*
RTA_TECH_N_N		0.263	0.255	0.217
		(0.042)**	(0.041)**	(0.040)**
RTA_TECH_N_S		0.132	0.130	0.115
		(0.072)$^{+}$	(0.071)$^{+}$	(0.067)$^{+}$
RTA_NO_TECH_S_N		-0.691	-0.684	-0.678
		(0.793)	(0.774)	(0.721)
RTA_NO_TECH_S_S		0.847	0.852	0.679
		(0.350)*	(0.348)*	(0.288)*
RTA_NO_TECH_N_N		0.645	0.637	0.604
		(0.093)**	(0.091)**	(0.094)**
RTA_NO_TECH_N_S		-0.210	-0.210	-0.202
		(0.158)	(0.156)	(0.154)
TRIPS_S_N			0.159	0.145
			(0.193)	(0.190)
TRIPS_S_S			0.413	0.421
			(0.218)$^{+}$	(0.194)*
TRIPS_N_N			0.190	0.208
			(0.119)	(0.123)$^{+}$
TRIPS_N_S			-0.012	-0.042
			(0.164)	(0.164)
PCT_S_N				1.134
				(0.344)**
PCT_S_S				1.607
				(0.252)**
PCT_N_N				0.627
				(0.087)**
PCT_N_S				0.150
				(0.194)
N	83228	83228	83228	83228

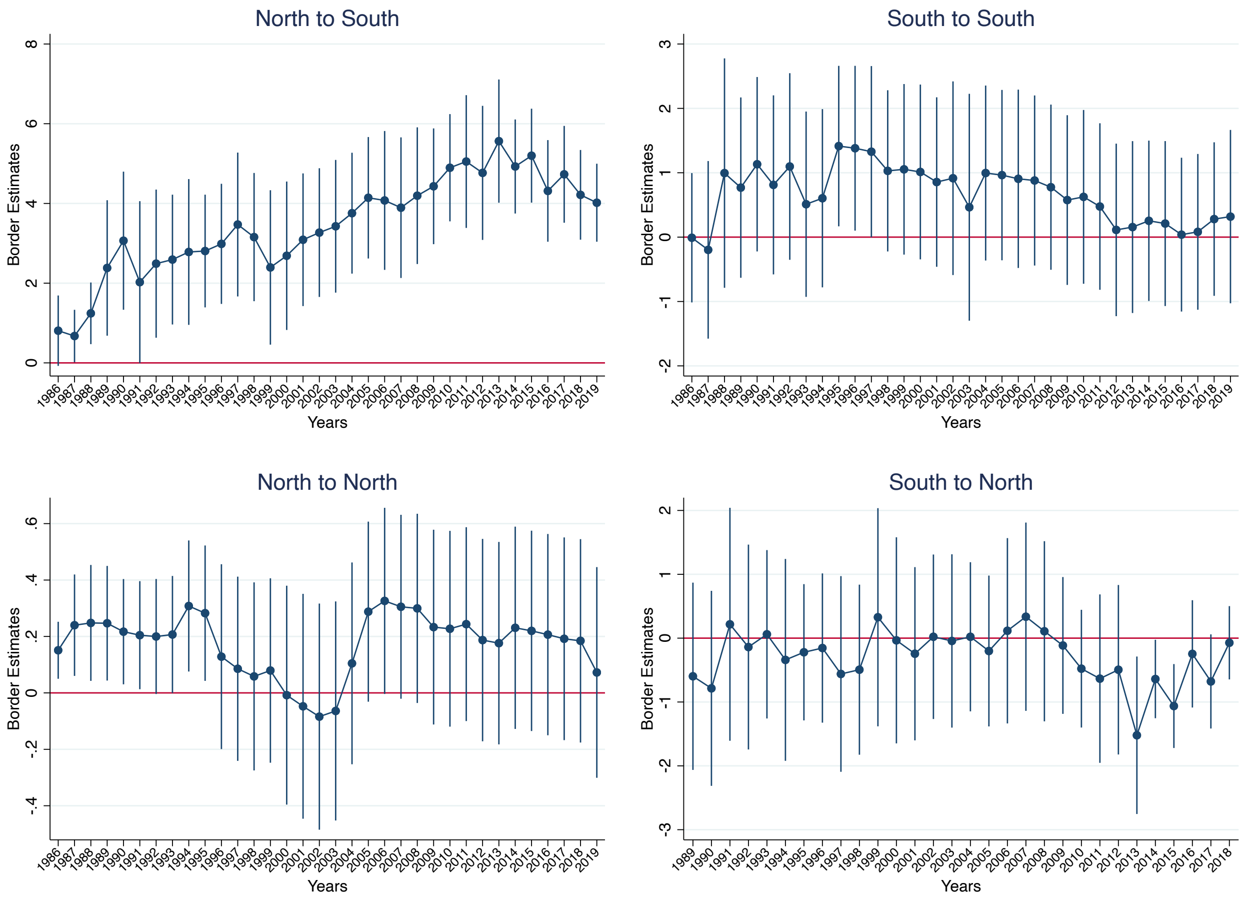
This table reports estimates of the effects of preferential agreements on cross-border patent flows. The estimates are obtained from specification (23), after allowing for the effects of globalization to vary across four bilateral groups (including "North to South", "South to South", "North to North", and "South to North", which are based on the income classification of the World Bank. In addition, each column of the table introduces a new policy variable, whose effects are also allowed to vary across the four bilateral income groups. Specifically, column (1) accounts for RTAs. Column (2) distinguishes between the effects of RTAs with and without technology provisions. In column (3) we add the TRIPS variables. Finally, in column (4), we also introduce the effects of the PCT. The dependent variable in each specification is the number of patents and the estimator is PPML. Standard errors in parentheses are clustered by country pair. $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

included in our specification, we focus our discussion on the results from column (4), which presents our main and most comprehensive findings. The estimates of the effects of RTAs with and without technology provisions are almost unchanged. According to our estimates, TRIPS has led to more patent flows from 'South to South' and from 'North to North', but the effects from the other two groups are not statistically significant. Finally, we find that the PCT has been very effective in promoting patent flows, especially from 'South to North', but also from 'South to South' and from 'North to North'. The only PCT estimate we obtain that is not statistically significant is for patent flows from "North to South."

Figure 10 reproduces the results from Figure 9 for the impact of globalization on cross-border patent flows for the four bilateral income groups of countries in our sample. However, the new globalization estimates are obtained from the econometric specification from column (4) of Table 3, which also includes the full set of our policy variables. We draw three conclusions based on the estimates from Figure 10. First, consistent with the significant policy estimates that we just obtained and discussed, we see that the policy variables in our model account for almost all of the globalization effects on cross-border patent flows from 'South to South' and from "North to North": almost all of the estimates for these two groups (see the top right and the bottom left panels of Figure 10) are no longer statistically significant. Second, an expected result is that the estimates of the effects of globalization remain non-significant for the 'South to North' group, as they were before including policy variables.

Finally, and most important for our purposes, the estimates of the impact of globalization on the patent flows from 'North to South' (in the top left panel of Figure 10) remain very strong. Comparison between the corresponding results for the "North to South" group between Figures 10 and 9 reveals that the preferential agreements we account for have contributed to some of the globalization effects from Figure 9. However, most of these globalization effects, as well as their evolution over time, remain strong in Figure 10 and, therefore, cannot be attributed to the policy variables in our model. The significant, and in fact exhaustive, policy effects that we obtain for the "South to South"

Figure 10: Globalization, Policy, and North-South Cross-border Patenting



Note: This figure reproduces the results from Figure 9 for the impact of globalization on cross-border patent flows for the four bilateral income groups of countries in our sample. However, the new globalization estimates are obtained from the econometric specification from column (4) of Table 3, which also includes the full set of our policy variables. See text for further details.

and “North to North” groups suggest that the large globalization effects we obtain for the patent flows from “North to South” may not be driven by any specific policies and reinforces our decision to capture such effects with the flexible border variable definitions. The latter would also enable us to perform some comparative statics in the next section.

We conclude this section with a brief discussion of the results from the robustness experiments that we perform to test the sensitivity of our main findings and to highlight some additional dimensions of our new database. The corresponding estimates, along with a more detailed discussion, appear in the Supplementary Appendix. We reproduce our results: (i) Applying the OLS estimator. (ii) Using three-way clustering. (iii) Not controlling for the impact of globalization explicitly. (iv) Excluding domestic patents.

Importantly, in this setting we cannot estimate the effects of globalization. (v) Leaving out China from our estimating sample. (vi) Using an alternative definition for "North" vs. "South". Specifically, we defined "South" as including only the "low" income countries and "North" for all other countries. (vii) Using patent citations to construct a new, quality adjusted measure of cross-border patents, which is used as our dependent variable. (viii) Imposing common effects across the different income groups for each of the policy variables in our model. (ix) Not using pair fixed effects in order to be able to identify the effects of "standard" gravity variables. (x) We also reproduced our main results at the industry level and at the sectoral level too. Overall, our main conclusions are reinforced by the additional experiments that we performed, but we also observed some intuitive heterogeneity. Our estimates, along with a corresponding discussion are included in the Appendix.

4.2 Cross-border Patenting, Development, and Inequality

Our empirical results indicate that globalization forces have been important drivers of cross-border patenting, especially from North to South. To the extent that cross-border patenting helps with technology transfer, these trends could be fostering growth in the South, leading to welfare improvements. In this section, we study, through the lens of our model, the influence of globalization on cross-border patenting, innovation, and welfare. Building upon the main estimates presented earlier, we conduct a numerical exercise to address the following question: What would have been the trajectory of cross-border patenting from North to South between 1995 and 2018 if globalization trends had remained at their 1995 levels?

Calibration. To answer this question, we employ our comprehensive dataset on patenting, geographical factors, and R&D intensity to calibrate our model. Several parameters are calibrated from previous studies or taken directly from the data. The parameter of the Armington elasticity takes a value of 5, which implies a trade elasticity of 4, which is standard in the trade literature. The parameter for the elasticity of innovation is set to 0.5, which is consistent with previous studies in the literature (see Cai, Li, and Santacreu, 2022). Population is taken from the CEPII datbase. The iceberg transport costs are calibrated using data on trade flows, geography measures, Gross Domestic Product (GDP) and population from CEPII, and deploying gravity methods using PPML with high dimensional fixed effects (see Zylkin, 2018). The elasticity of patent frictions η is set to 2.15, whereas the elasticity of the patenting cost ξ is set to 4.5. The value of these parameters implies an estimate of the border coefficient of $\frac{\eta}{\xi-1} = 0.61$, which is consistent with our empirical findings. Finally $\bar{p}hi_{SN}$ is set to 0.25, which implies a probability of imitation at the peak of 40%.

Table 4 reports the parameter values. The remaining parameters, namely, the innovation efficiency in the North, γ_N, the parameter in the cost of patenting function, $\bar{\lambda}$, and the patenting frictions, $\tau_{SN,t}$, are calibrated to match data on R&D intensity for the United States (serving as a proxy for R&D intensity in the North), denoted as H^r_{Nt}; the share of patents from the North filed for protection in the South, represented as $\lambda_{SN,t} = \frac{\text{Pat}SN,t}{ZNt}$; and the border effect obtained from our main specification in column (4) of Table 3 and Figure 10, corresponding in theory to $\phi_{SN,t}^{\frac{1}{\xi-1}}$.

Table 4: Parameter Values

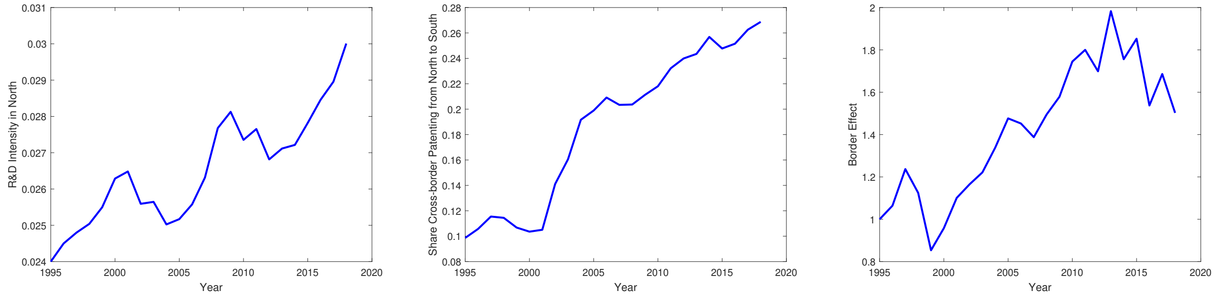
Parameter	Value	Description
σ	5	Armington elasticity
d_{NS}	3.5	Iceberg trade costs from S to N
d_{SN}	3.5	Iceberg trade costs from N to S
β_r	0.5	Elasticity of innovation
L_N	0.8	Population N
L_N	1	Population S
ξ	4.5	Elasticity in the cost of patenting
η	2.15	Elasticity of patenting friction
$\bar{\phi}_{SN}$	0.2	Prob imitation

The evolution of these variables is depicted in Figure 11. Over the analyzed period, the R&D intensity of the United States increased by a factor of 1.25, rising from 2.4 percent in 1995 to over 3% in 2018. Additionally, the share of patent applications from the North to the South tripled, surging from 8 percent in 1995 to 28 percent in 2018.

Furthermore, globalization trends proxied by the border effects witnessed an increase of approximately 1.5 times.

Figure 11: Observable Variables

(a) R&D intensity (b) Cross-border patent shares (c) Border effect



Notes: The figure plots the evolution of three observable variables we use to calibrate the parameters γ_{Nt}, $\tau_{SN,t}$, and $\bar{\lambda}_{SN,t}$ in the model. The left panel represents R&D intensity in the United States (in basis points), which we use as a proxy of R&D intensity in North; the middle panel represents data on the share of cross-border patenting from North to South, classified as in Section 2, and defined as the ratio of cross-border patents to the total number of patents being filed by North anywhere in the world; the right panel represents data on the border effect obtained from our estimation strategy.

We calibrate γ_{Nt}, $\tau_{SN,t}$, and $\bar{\lambda}_{SN,t}$ to match the trends depicted in Figure 11. To calibrate the patenting frictions, we specify the following functional form for the probability of imitation, $1 - \phi_{SN,t}$, as a function of patenting frictions:

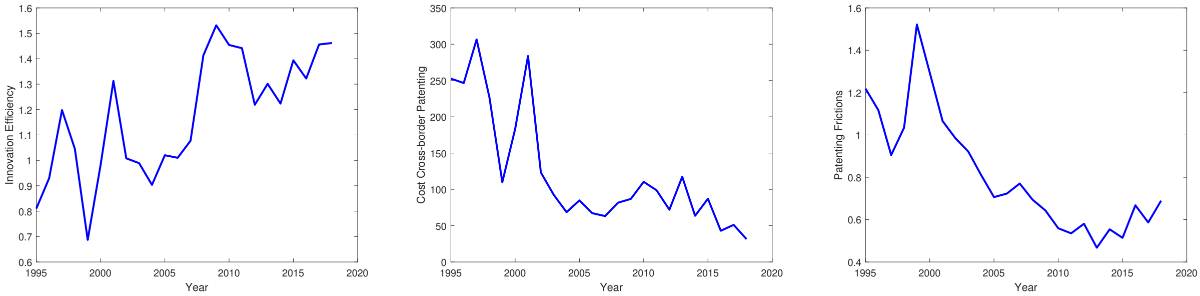
$$\phi_{SN,t} = \bar{\phi}_{SN} \tau_{SN,t}^{-\eta},$$

where $\tau_{SN,t}$ represents the policy variable that characterizes patenting frictions and can be inferred from the estimated border effects in our empirical analysis.

The calibrated parameters are reported in Figure 12. Throughout the analyzed period, there has been an increase in innovation efficiency in the North, corresponding with a rise in R&D investment. The decline in the cost of patenting aligns with the observed increase in the share of cross-border patenting from North to South. Lastly, the decrease in patenting frictions corresponds to the upward trend in globalization.

Figure 12: Calibrated Parameters

(a) Innovation efficiency, γ_{Nt} (b) Cost of patenting, $\bar{\lambda}_{SN,t}$ (c) Patent friction, $\tau_{SN,t}$



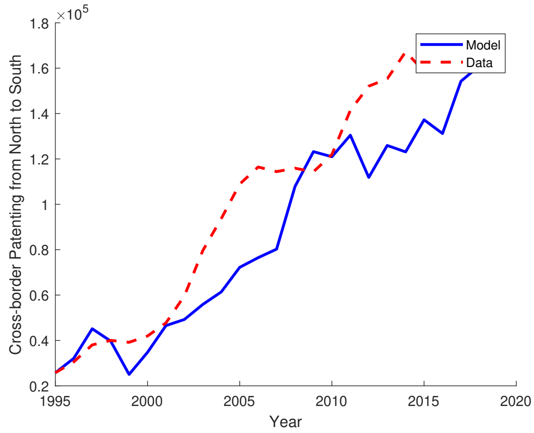
Notes: The figure plots the parameters γ_{Nt} (left panel), $\tau_{SN,t}$ (bottom panel), and $\bar{\lambda}_{SN,t}$ (right panel) which we calibrate to match the data from Figure 11.

Validation: Untargeted moments. Next, we provide validation of our model by assessing how well it fits two variables that were not explicitly targeted in the calibration process: the number of patents, Z_{Nt}, and the number of cross-border patent applications, $\text{Pat}_{SN,t}$. Although we used data on $\lambda_{SN,t}$ for parameter calibration, the model also accurately matches these two additional variables that determine this share. The correlation between patents created by North in both the data and the model is 0.80. Regarding cross-border patenting, the correlation between the model and the data is 0.95. The evolution of the variables is plotted in Figure 13.

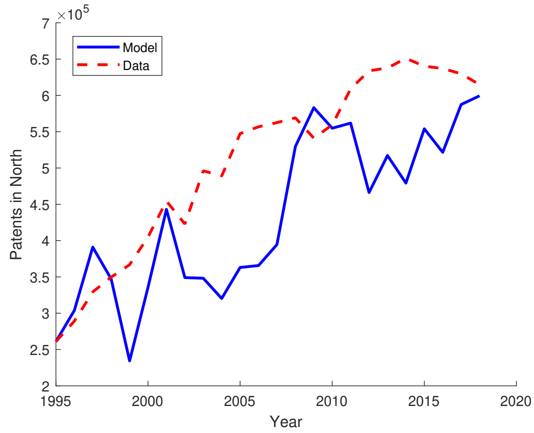
Counterfactual Analysis: Cross-border patenting and welfare. We then proceed with our main counterfactual analysis. Relying on the estimates of the globalization effects on patent flows from "North to South," we simulate a scenario in which globalization had absolutely no impact on these patent flows. To achieve this, we use the vector of patent frictions from 1995 to 2018 as our baseline and set all border estimates to their corresponding values for 1995. Figure 14 plots the evolution of cross-border patenting from North to the South in both the data and the counterfactual scenario where patenting frictions (i.e., border effects) are set to their 1995 levels, effectively removing the influence of globalization. Our findings indicate that, in the absence of globalization trends, cross-border patenting would have been significantly lower. Specifically, in 2018, cross-border patenting would have been 43 percent lower absent globalization trends observed in the

Figure 13: Validation: Untargeted moments

(a) Cross-border patenting: ‘North to South’



(b) Number of innovations in North



Notes: The figure shows the evolution of two variables in the data (solid line) and in the model (dashed line). The left panel shows the evolution of cross-border patenting from North to South; the right panel shows the evolution of the number of patents being filed by North to the world.

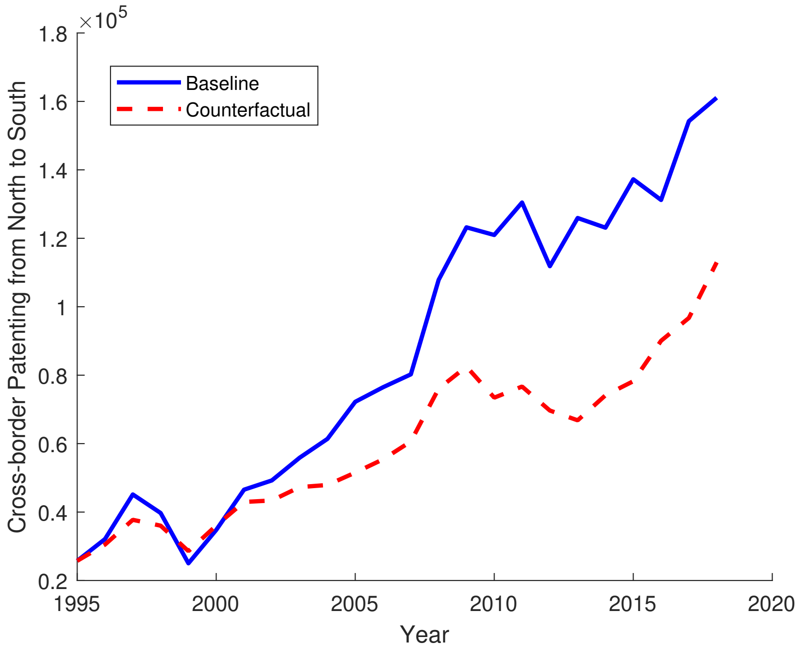
data. These finding reveal that globalization trends help knowledge transfer through cross-border patenting, fostering development in the South.

We complete the analysis by evaluating welfare changes between our baseline and our counterfactual where patent frictions stay at the levels of 1995. We assume a discount factor of $\beta = 0.98$ and compute welfare using equation (19).

Figure 15 plots the evolution of real consumption in North and South both in the baseline scenario and in the counterfactual. Absent globalization effects, consumption would have been lower in both countries. The effects are bigger after 2001, which is consistent with the empirical results displayed in Figure 10.

Table 5 reports welfare changes for North and South between the baseline and the counterfactual, following equation (19). Our results indicate that welfare in both countries would have been lower if the world had not experienced an increase in globalization and the welfare impacts are heterogeneous across countries. These results reveal two intuitive but, nevertheless, important results. First, we see that both the rich and the poor countries in the world have gained from the transfer of patents across international borders. Second, we note that the gains for “South” are actually larger if we focus on the period after the 2000s. The implications are that poor countries have become richer

Figure 14: Counterfactual: The Role of Globalization

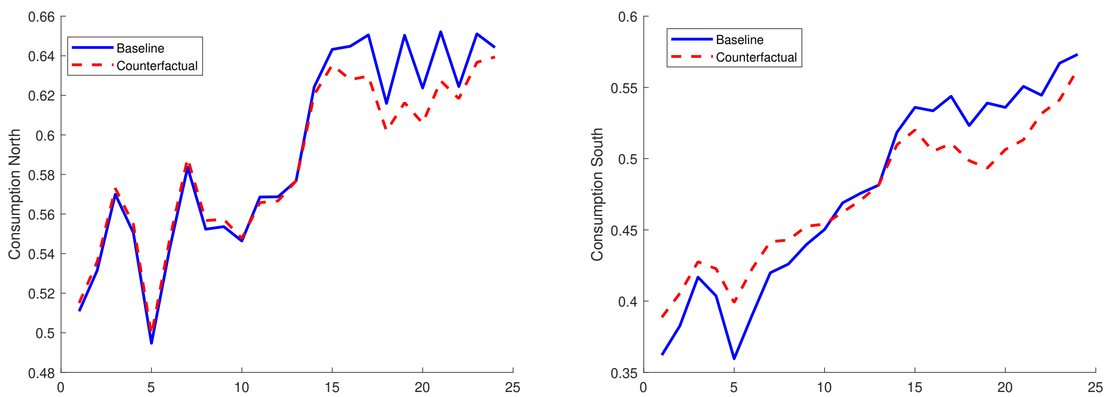


Notes: The figure shows the evolution of cross-border patenting from North to South in the data (solid line) and in a counterfactual in which we set patent frictions fro the period to their levels in 1995 (dashed line).

Figure 15: Welfare Analysis

(a) Consumption North

(b) Consumption South



Notes: The figure shows the evolution of consumption between 1995-2019 in North (left panel) and South (right panel), both in the baseline scenario (solid line) and in the counterfactual in which border effects had stayed at their levels of 1995 (dashed line).

and, probably more important for our current purposes, that cross-border patenting has led to a decrease in the real income gap between poor and rich countries. The natural policy implication is to look for further opportunities to stimulate such activities.

Table 5: Welfare gains (percent)

	Welfare gains (all period)	Welfare gains (2001-2019)
North	-0.72	-2.30
South	-0.12	-4.47

Notes: The table reports welfare gains between the baseline and a counterfactual in which border effects stay at their levels in 1995 for North and South, during the period 1995-2019 (first column), and during the period 2001-2019 (second column).

5 Final Remarks

This paper empirically explores the drivers behind firms seeking international patent protection and links cross-border patenting to development and inequality. To this end, we compiled a new dataset on cross-border patenting across industries, enabling us to gain insights into global patenting behavior. To guide our analysis, we have developed a stylized model that links globalization trends, trade policies, and cross-border patenting, emphasizing the importance of patent transfers from developed to developing countries. The model yields a structural gravity model, which we have estimated using the latest techniques from empirical trade literature, allowing us to account for the impacts of globalization and various policy factors.

Our analysis yields estimates of globalization effects that vary between North and South regions. Notably, globalization-driven patent flows from North to South have had a more favorable impact on the South after the 2000s, reducing global income inequality. Regarding policy variables, we find that RTAs play a role, especially those with IPR provisions that increase patent flows from North to South. Furthermore, both TRIPs and the PCT promote cross-patenting, though the impact varies depending on the groups of countries involved in the origin and destination of patent flows.

A counterfactual exercise shows that absent globalization forces, welfare would have been lower in both North and South, which larger welfare losses in South after 2000.

While our analysis focuses on the connection between globalization, trade policy and cross-border patenting, there may be other channels that influence firms' decisions to seek international patent protection. These could include FDI and cross-border licensing. It would be interesting to explore the extent to which these channels influence firms' behavior and how they interact with regional trade agreements and IPR regimes. We leave these questions for future research.

References

Agnosteva, Delina E., James E. Anderson, and Yoto V. Yotov. 2019. "Intra-national Trade Costs: Assaying Regional Frictions." *European Economic Review* 112 (C):32–50.

Arregui, Daniela and Inmaculada Martínez-Zarzoso. 2022. "Do Trade Agreements Contribute to Technology Internationalization?" *INFER WP Series* (3).

Baier, Scott L. and Jeffrey H. Bergstrand. 2007. "Do Free Trade Agreements Actually Increase Members' International Trade?" *Journal of International Economics* 71 (1):72–95.

Baier, Scott L., Yoto V. Yotov, and Thomas Zylkin. 2019. "On the Widely Differing Effects of Free Trade Agreements: Lessons from Twenty Years of Trade Integration." *Journal of International Economics* 116:206–226.

Bergstrand, Jeffrey H., Mario Larch, and Yoto V. Yotov. 2015. "Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation." *European Economic Review* 78:307–327.

Bond, Eric W and Kamal Saggi. 2020. "Patent protection in developing countries and global welfare: WTO obligations versus flexibilities." *Journal of International Economics* 122:103281.

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

Branstetter, Lee, Raymond Fisman, C Fritz Foley, and Kamal Saggi. 2007. "Intellectual property rights, imitation, and foreign direct investment: Theory and evidence." Tech. rep., National Bureau of Economic Research.

Brunel, Claire and Thomas Zylkin. 2022. "Do cross-border patents promote trade?" *Canadian Journal of Economics/Revue canadienne d'économique* 55 (1):379–418.

Cai, Jie, Nan Li, and Ana Maria Santacreu. 2022. "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors." *AEJ: Macroeconomics* 14 (1):104–145.

Coelli, Federica, Andreas Moxnes, and Karen Helene Ulltveit-Moe. 2022. "Better, faster, stronger: Global innovation and trade liberalization." *The Review of Economics and Statistics* 104 (2):205–216.

Coleman, Elijah. 2022. "The Effects of Multilateral IP Treaties on Patenting Behavior." .

De Rassenfosse, Gaétan, Marco Grazzi, Daniele Moschella, and Gabriele Pellegrino. 2022. "International patent protection and trade: Transaction-level evidence." *European Economic Review* 147:104160.

De Rassenfosse, Gaetan, Jan Kozak, and Florian Seliger. 2021. "Geocoding of Worldwide Patent Data." *Available at SSRN 3425764* .

de Sousa, José. 2012. “The currency union effect on trade is decreasing over time.” *Economics Letters* 117 (3):917–920.

Diwan, Ishac and Dani Rodrik. 1991. “Patents, appropriate technology, and North-South trade.” *Journal of International Economics* 30 (1-2):27–47.

Egger, Peter H. and Sergey Nigai. 2015. “Structural gravity with dummies only: Constrained ANOVA-type estimation of gravity models.” *Journal of International Economics* 97 (1):86–99.

Egger, Peter H. and Filip Tarlea. 2015. “Multi-way clustering estimation of standard errors in gravity models.” *Economics Letters* 134 (C):144–147.

Glass, Amy Jocelyn and Kamal Saggi. 1998. “International technology transfer and the technology gap.” *Journal of development economics* 55 (2):369–398.

Goldschlag, Nathan, Travis J Lybbert, and Nikolas Jason Zolas. 2016. “An ‘algorithmic links with probabilities’ crosswalk for uspc and cpc patent classifications with an application towards industrial technology composition.” *US Census Bureau Center for Economic Studies paper no. CES-WP-16-15* .

Grossman, Gene M and Edwin L-C Lai. 2004. “International protection of intellectual property.” *The American Economic Review* 94 (5):1635–1653.

Helpman, Elhanan. 1993. “Innovation, Imitation, and Intellectual Property Rights.” *Econometrica* 61:1247–80.

Hémous, David, Simon Lepot, Thomas Sampson, and Julian Schärer. 2023. “Trade, innovation and optimal patent protection.” .

Hoekman, Bernard and Kamal Saggi. 2007. “Intellectual property provisions in North-South trade agreements.” .

Howard, Jacob, Keith E Maskus, and William C Ridley. 2023. “The Impacts of Intellectual Property–Related Preferential Trade Agreements on Bilateral Patent Applications.” .

Keller, Wolfgang. 2004. “International technology diffusion.” *Journal of economic literature* 42 (3):752–782.

Kwan, Yum K and Edwin L-C Lai. 2003. “Intellectual property rights protection and endogenous economic growth.” *Journal of Economic Dynamics and Control* 27 (5):853–873.

Lai, Edwin L-C. 1998. “International intellectual property rights protection and the rate of product innovation.” *Journal of Development Economics* 55 (1):133–153.

Lai, Edwin L-C and Larry D Qiu. 2003. “The North’s intellectual property rights standard for the South?” *Journal of International Economics* 59 (1):183–209.

Lybbert, Travis J. and Nikolas J. Zolas. 2022. “Getting Patents & Economic Data to Speak to Each Other: An ‘Algorithmic Links with Probabilities’ Approach for Joint Analyses of Patenting & Economic Activity.” *WIPO Economic Research Working Papers 05* .

Martínez-Zarzoso, Inmaculada and Santiago Chelala. 2021. "Trade agreements and international technology transfer." *Review of World Economics* 157 (3):631–665.

Maskus, Keith Eugene. 2000. *Intellectual property rights in the global economy*. Peterson Institute.

Pfaffermayr, Michael. 2019. "Gravity models, PPML estimation and the bias of the robust standard errors." *Applied Economics Letters* 26 (18):1467–1471.

Santacreu, Ana Maria. 2022. "Dynamic Gains from Trade Agreements with Intellectual Property Provisions." .

Santos Silva, J.M.C. and Silvana Tenreyro. 2006. "The Log of Gravity." *Review of Economics and Statistics* 88 (4):641–658.

Santos Silva, J.M.C. and Silvana Tenreyro. 2021. "The Log of Gravity At 15." School of Economics Discussion Papers 0121, School of Economics, University of Surrey. URL `https://ideas.repec.org/p/sur/surrec/0121.html`.

Tanaka, Hitoshi and Tatsuro Iwaisako. 2014. "Intellectual property rights and foreign direct investment: A welfare analysis." *European Economic Review* 67:107–124.

The World Bank. 2015. "Foreign direct investment and development: Insights from literature and ideas for research." *http://blogs.worldbank.org/psd/foreign-direct-investment-and-development-insights-literature-and-ideas-research, submitted by C. Qiang. Co-authors: R. Echandi and J. Krajcovicova* .

UNCTAD, Commodities. 2023. "International supply networks: A portrait of global trade patterns in four sectors." .

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press, 2nd ed.

Yang, Guifang and Keith E Maskus. 2001. "Intellectual property rights, licensing, and innovation in an endogenous product-cycle model." *Journal of International Economics* 53 (1):169–187.

Yotov, Yoto V. 2022. "On the role of domestic trade flows for estimating the gravity model of trade." *Contemporary Economic Policy* URL `https://doi.org/10.1111/coep.12567`.

Zylkin, Thomas. 2018. "PPML_PANEL_SG: Stata module to estimate structural gravity models via Poisson PML." .

Supplementary Appendix

This appendix reports and discusses the results from a series of experiments we performed to test the robustness of our main findings and to highlight the dimensions of our new data. The experiments appear in the order in which they were mentioned in the main text. The first two experiments we perform are related to the heterogeneous estimates of the effects of globalization.

- In our first experiment we obtain directional globalization estimates based on all four income categories of the World Bank, including, "high income", "upper-middle income", "lower-middle income", and "low income". Figure A1 reports our estimates, which are intuitive and as expected. Most importantly, the strongest effects of globalization are from rich to poor countries. The relationship between these estimates and our main results is that the latter are essentially weighted averages of the results in Figure 9.

- In Figure A2 we reproduce our main results from Figure 9, but without allowing for directional effects, i.e., just for 'North' vs. 'South'. The results are expected and are consistent with our main findings.

Table A1 shows the results for a number of robustness checks, obtained by estimating variations of the main empirical model form column (4) of Table 3. For clarity and expositional simplicity, we report only the globalization effects for "North" to "South" and the estimates from all specifications are included in a single Figure A3.

- The first variation consists of estimating a linear model using the natural log of patent counts as a dependent variable. As can be seen in column (1), the estimated coefficients for RTA present higher standard errors than for the PPML model with the dependent variable in levels. Nevertheless, the few coefficients that are accurately estimated present the same sign and similar magnitudes.

- The second column of Table A1 presents the results when standard errors are clustered by three dimensions: origin, destination, and time (instead of by pair). The statistical significance of the coefficients decreases slightly, but most interpretations remain valid.

- Column (3) shows the results when the model is estimated excluding the terms that proxy for globalization. In this case, the coefficients slightly decrease in magnitude compared with column (2) but remain within similar confidence bands.

- Column (4) excludes domestic patents, which results in comparable coefficients for the RTA and TRIPS variables, but the significance for the PTC vanishes. Importantly, this specification does not allow us to obtain estimates of any of the globalization effects that have been of central interest to us.

- Finally, column (5) shows the results excluding China, showing that this affects the significance for the NS and SN policy variables, whereas coefficients for the North to North, and South to South pairs remain similar.

In Table A2 five additional robustness checks are presented.

- In column (1) a different classification of North and South is used, placing in South exclusively low-income countries, whereas North contains the other WB categories: upper- and lower-middle- and high-income. A noticeable change in the RTA coefficients is that agreements with technology provisions between North and South have a higher coefficient than with the baseline classification, when the RTA contains technology clauses, whereas when those RTAs are between low-income countries the effect on cross-patenting is negative and now significant at the 5 percent level (10 percent in the main results). This latter outcome is probably driven by less than a handful of agreements.

- A common concern when analyzing patents trend is controlling for the quality of patents. There are some cases when patents might be filed en masse for reasons other than increasing innovation. For example, if a country improves its IPR quickly and significantly, it may cause a surge in patenting, as innovators rush to take advantage of this new IP protection. This could lead to a situation where the patenting activity of a country increases by much more than their innovation level would indicate. Another notable example would be the case of government subsidies for filed patents. This would incentivize the filing of many patents regardless of if they are actually of any merit. These factors make it important to consider the notion of "quality patents"; in other words, patents that are actually the result of innovation and result in a new useful knowledge base being created. This motivates our next experiment, in which we construct a quality-adjusted patent count.

There are many ways to adjust for quality, but we follow one of the methods developed by Coelli, Moxnes, and Ulltveit-Moe (2022) in which they look at the number of citations created as a share of the number of patents filed. We calculate this relative quality of patent flows as follows: Let c_p denote the number of citations that occur within the first three years after patent p was filed. Let μ_f be the average number of citations within the first three years after filing across that DOCDB family such that:

$$\mu_f = \frac{\sum_{p \in \Xi_{pf}} c_p}{\sum_{p \in \Xi_{pf}} p},$$

where Ξ_{pf} is the number of patents, p, in that DOCDB family, f. The sum of citations for each origin is then:

$$Q_{it} = \sum_{p \in \Xi_{pf}} \mu_f,$$

where Ξ_{ft} is the set of country i's families filed in year t. The full average quality for origin i in year t is given by:

$$\hat{Q_{it}} = \frac{Q_{it}}{P_{it}},$$

where P_it is the total patents filed by country i in year t.

Our new estimates with the quality-adjusted dependent variable appear in column

(2) of Table A2, and they show a decrease in the statistical significance of the RTA coefficients for agreements between NS and SN. The corresponding globalization effects for the group North to South are shown in the bottom-left panel in Figure A3, showing positive and significant effects for 2002 onward.

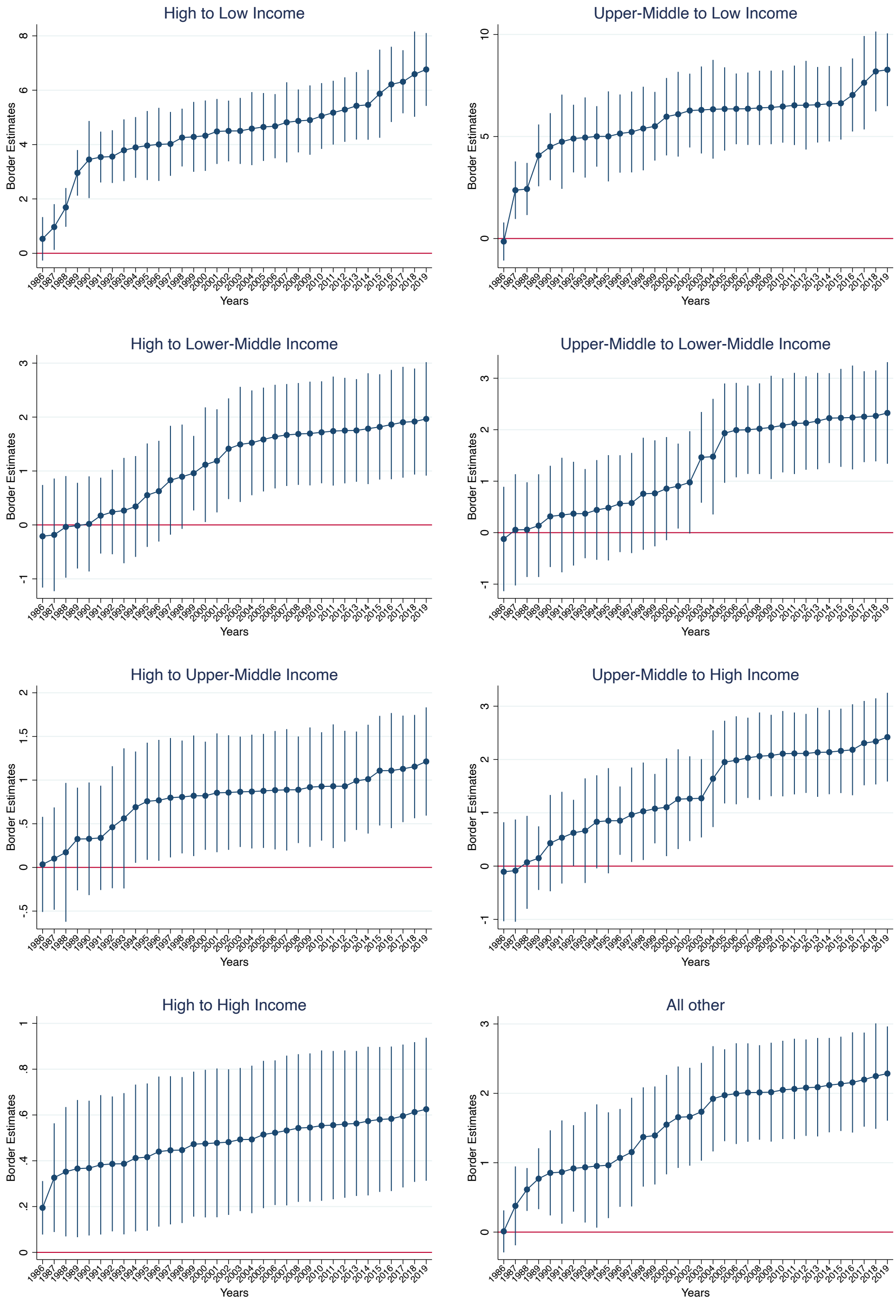
- Next, in column (3) the averaged result for each policy variable are shown, without considering the level of development of the pair or countries. The results indicate that, on average, the effect of having an RTA with or without technology provisions is positive, similar in magnitude and more accurately estimated for those with provisions. However, an important message from these results, in combination with our main findings, is that the average agreement estimates may be masking significant heterogeneity in the impact of policy on cross-border patent flows.

- Finally, in column (4) we replace the pair fixed effects with a set of "standard" gravity variables, including the logged distance between countries (weighted by population), which is allowed to have heterogeneous effects for domestic vs. international patents, and dummy variables for sharing a common border, having the same official language and a past or present colonial relationship. Consistent with the trade gravity literature, our estimates for cross-border patents reveal that cross-border patenting decreases with distance and increases when countries share an official language. In fact, the estimate on common official language is significantly larger in magnitude than the standard estimate from the trade literature. We find this result intuitive, as language is potentially a more important factor for patent sharing.

  We also obtain some results that are different from those for trade. For example, we obtain negative and significant estimates for the effects of common borders and colonial ties, while the corresponding estimates from the trade literature are mostly positive and statistically significant. In addition, unlike the trade literature, we obtain a larger negative impact of domestic distance as compared to international distance. We are not aware of existing estimates for cross-border patents against which we can benchmark our findings. Moreover, these results are not important for our main purposes. Nevertheless, we find them interesting and possibly worth further investigation.

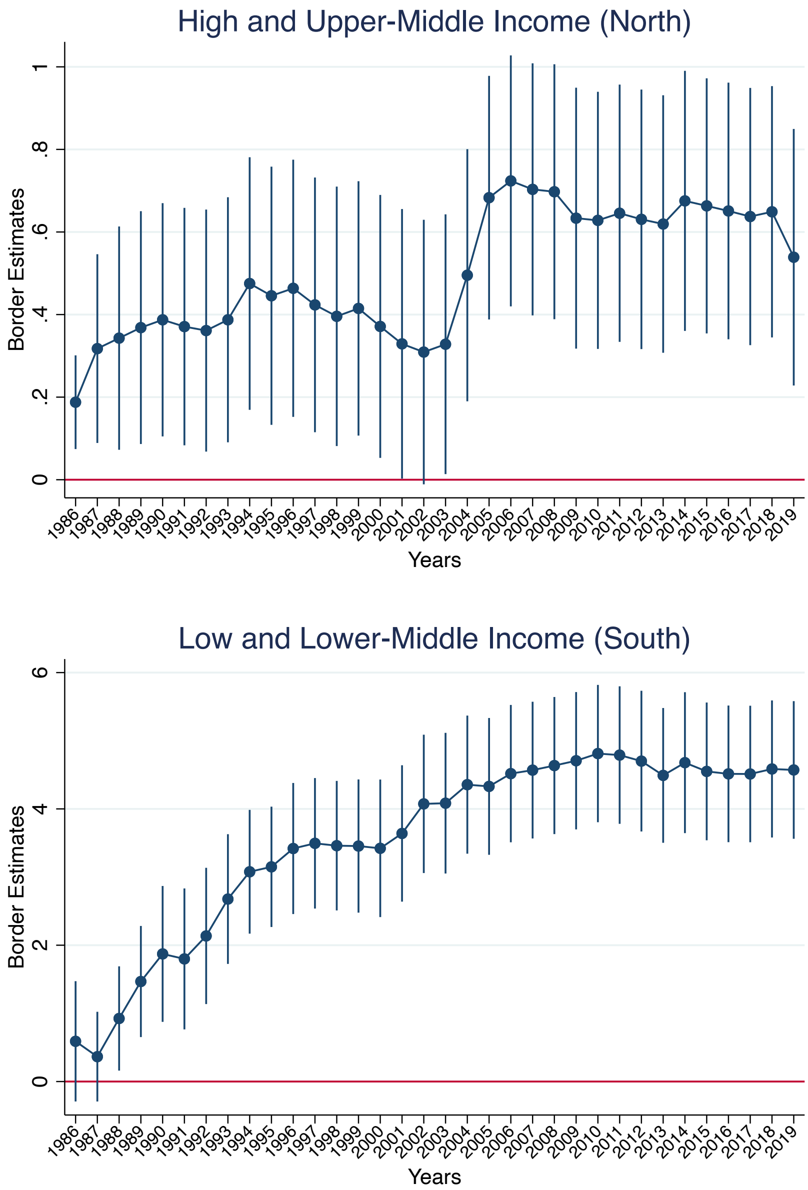
Finally, to highlight the sectoral dimension of our new dataset, we also obtain disaggregated estimates. Table A3 reports estimates at the sectoral level. Estimations are presented for all sectors in column (1) and for specific groups of manufacturing sectors, according to the Standard Industrial Classification Revision 3, in columns (2)-(5). The results shown in column (1) permit us to discard the existence of aggregation bias in our main results. In addition, results for specific groups of sectors are presented according to their level of sophistication. S1, S2 and S3 denote respectively sectors 15-19, 20-29 and 30-37, respectively. S1 includes food and beverages, tobacco, textile and apparel, leather and footwear; S2 includes paper and printing, chemicals and metals, among others; and S3 are office and computing machinery, communication equipment, vehicles and medical, precision and optical instruments, for example. The effect of RTA with technology provisions has a significantly higher magnitude for flows of patents going from North to North or to South in S3, whereas those without technology provision show a negative and significant effect only for innovators in South patenting in North offices and for S3. The effect of other policy variables, TRIPS and PCT, do not vary much across sectors. Two main policy implications stand out from this analysis. First, the heterogeneity that we document across the broad sectors implies that serious policy analysis should be performed at the disaggregated level, potentially even more disaggregated than presented here (for which our data allow). Second, for RTA to facilitate/promote cross-patenting between rich and poor countries, the agreements must contain specific chapters on IPR and innovation.

Figure A1: Globalization and Cross-border Patenting, Directional



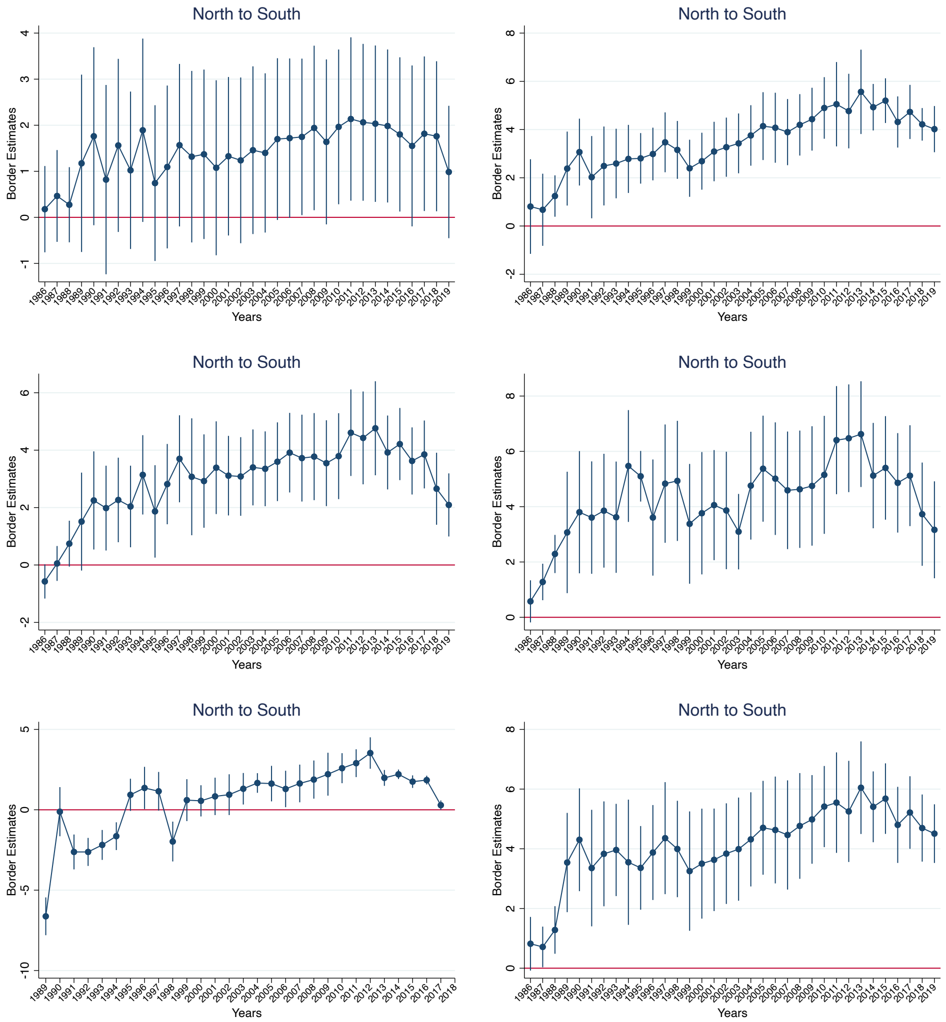
Note: This figure reproduces our main directional globalization estimates but based on the four income groups from the 2000 classification of the World Bank, including 'high income', 'upper-middle income', 'lower-middle income', and 'low income'. All estimates are obtained from a single regression, which is based on specification (24) after allowing for heterogeneous effects for each of the four bilateral groups. See text for further details.

Figure A2: Globalization and Cross-border Patenting, North vs South



Note: This figure reproduces our main globalization estimates for the 'North' vs. 'South' group of countries but without allowing for directional effects. All estimates are obtained from a single regression, which is based on specification (24) after allowing for heterogeneous effects for each of the four bilateral groups. See text for further details.

Figure A3: Globalization and Cross-border Patenting, Robustness



Note: This figure reproduces our main directional globalization estimates for the group 'North to South' but based on the robustness checks presented in tables (A1) and (A2), including in this order from top to bottom and left to right: 'OLS', 'CLUST', 'NOCHN', 'POOR', 'QALTY' and 'CMMN'. See the notes below tables (A1) and (A2) for further details. The graphs are obtained from from specification (23), after allowing for the effects of globalization to vary across four bilateral groups ('North to South', 'South to South', 'North to North', and 'South to North'. See text for further details. Notice that only results for the first group are presented, since there were the only group showing significant effects, after including in the empirical model the policy variables.

Table A1: Preferential Agreements and Cross-border Patents: Robustness I

	(1) OLS	(2) CLUST	(3) NOGLOB	(4) NODOM	(5) NOCHN
RTA_TECH_S_N	-0.030	0.170	0.141	0.194	-0.159
	(0.087)	(0.079)*	(0.074)$^{+}$	(0.046)**	(0.161)
RTA_TECH_S_S	-0.088	-0.587	-0.667	-0.629	-0.630
	(0.114)	(0.257)*	(0.166)**	(0.212)**	(0.280)*
RTA_TECH_N_N	0.009	0.217	0.233	0.059	0.235
	(0.041)	(0.066)**	(0.037)**	(0.035)$^{+}$	(0.039)**
RTA_TECH_N_S	-0.059	0.115	0.125	0.019	0.142
	(0.064)	(0.081)	(0.073)$^{+}$	(0.048)	(0.169)
RTA_NO_TECH_S_N	-0.031	-0.678	-0.519	-0.747	0.233
	(0.121)	(0.795)	(0.476)	(0.690)	(0.193)
RTA_NO_TECH_S_S	0.134	0.679	0.548	0.338	0.602
	(0.188)	(0.377)$^{+}$	(0.209)**	(0.254)	(0.331)$^{+}$
RTA_NO_TECH_N_N	0.531	0.604	0.625	0.330	0.626
	(0.057)**	(0.168)**	(0.097)**	(0.081)**	(0.094)**
RTA_NO_TECH_N_S	0.097	-0.202	-0.198	-0.263	0.056
	(0.063)	(0.308)	(0.147)	(0.144)$^{+}$	(0.126)
TRIPS_S_N	0.348	0.145	1.049	1.399	0.258
	(0.105)**	(0.124)	(0.246)**	(0.216)**	(0.156)$^{+}$
TRIPS_S_S	0.312	0.421	-0.114	0.813	0.008
	(0.111)**	(0.202)*	(0.190)	(0.227)**	(0.140)
TRIPS_N_N	0.007	0.208	0.173	0.280	0.202
	(0.057)	(0.211)	(0.055)**	(0.110)*	(0.121)$^{+}$
TRIPS_N_S	0.165	-0.042	0.081	0.815	-0.229
	(0.086)$^{+}$	(0.170)	(0.183)	(0.218)**	(0.125)$^{+}$
PCT_S_N	0.248	1.134	1.940	-0.151	1.391
	(0.104)*	(0.518)*	(0.340)**	(0.394)	(0.237)**
PCT_S_S	0.026	1.607	1.524	-0.179	1.954
	(0.135)	(0.476)**	(0.195)**	(0.306)	(0.313)**
PCT_N_N	0.305	0.627	0.636	-0.052	0.612
	(0.072)**	(0.159)**	(0.079)**	(0.099)	(0.088)**
PCT_N_S	0.137	0.150	0.402	-0.468	0.288
	(0.079)$^{+}$	(0.282)	(0.222)$^{+}$	(0.207)*	(0.163)$^{+}$
N	71465	83228	83228	80932	79825
r2	0.920				

This table reports a number of robustness checks. The estimates are obtained from specification (23), after allowing for the effects of globalization to vary across four bilateral groups ('North to South', 'South to South', 'North to North', and 'South to North', which are based on the income classification of the World Bank. Each column of the table introduces a variation of the main model. Specifically, in column (1) uses a linear specification that is estimated by OLS with the dependent variable in natural logs. Column (2) clusters standard errors differently (multi-clustering). In column (3) the globalization effects are excluded from the specification. In column (4), estimates the model without domestic patents. Finally, column (5) excludes China from the sample. The dependent variable in each specification is the number of patents and the estimator is PPML in all but column (1). Standard errors in parentheses are clustered by country pair in all columns but . $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table A2: Preferential Agreements and Cross-border Patents: Robustness II

	(1) POOR	(2) QALTY	(3) COMMN	(4) GRAV
RTA_TECH_S_N	0.185	0.208		-0.410
	(0.065)**	(0.077)**		(0.326)
RTA_TECH_S_S	-0.806	-0.540		1.559
	(0.200)**	(0.190)**		(0.585)**
RTA_TECH_N_N	0.215	0.115		-0.392
	(0.039)**	(0.036)**		(0.186)*
RTA_TECH_N_S	0.131	0.010		-0.567
	(0.069)+	(0.046)		(0.244)*
RTA_NO_TECH_S_N	-0.525	0.272		-0.163
	(0.554)	(0.255)		(0.330)
RTA_NO_TECH_S_S	0.381	0.288		2.066
	(0.257)	(0.420)		(0.470)**
RTA_NO_TECH_N_N	0.605	0.965		0.245
	(0.094)**	(0.196)**		(0.166)
RTA_NO_TECH_N_S	-0.195	-0.354		0.121
	(0.150)	(0.171)*		(0.356)
TRIPS_S_N	0.957	-0.205		1.676
	(0.251)**	(0.255)		(0.404)**
TRIPS_S_S	-0.225	0.770		-0.263
	(0.229)	(0.248)**		(0.320)
TRIPS_N_N	0.169	0.116		0.309
	(0.112)	(0.151)		(0.240)
TRIPS_N_S	0.036	0.411		0.534
	(0.184)	(0.254)		(0.413)
PCT_S_N	1.533	0.439		0.686
	(0.301)**	(0.165)**		(0.521)
PCT_S_S	1.363	1.188		1.945
	(0.233)**	(0.338)**		(0.562)**
PCT_N_N	0.624	0.407		-0.948
	(0.086)**	(0.107)**		(0.355)**
PCT_N_S	0.308	-0.316		1.531
	(0.211)	(0.169)+		(0.487)**
RTA_TECH			0.153	
			(0.040)**	
RTA_NO_TECH			0.179	
			(0.090)*	
TRIPS			0.310	
			(0.180)+	
PCT			1.074	
			(0.146)**	
LN_DIST				-0.469
				(0.080)**
LN_DIST_DOM				-1.041
				(0.131)**
CONTIG				-0.434
				(0.182)*
COMLANG				1.023
				(0.151)**
COMCOL				-0.713
				(0.308)*
N	83224	73366	83228	83814

This table reports a number of robustness checks. The estimates are obtained from specification (23), allowing for the effects of globalization to vary across four bilateral groups ('North to South', 'South to South', 'North to North', and 'South to North', which are based on the income classification of the World Bank, in columns (1), (2) and (5). Each column of the table introduces a variation of the main model. Specifically, column (1) uses a different classification of North and South countries, with South including only low income countries. Column (2) uses as dependent variable the number of patents weighted by the number of citations. In column (3) we present average common effects of the policy variables. In column (4), we introduce "gravity" variables: the natural log of distance weighted by population (distinguishing between international and domestic distance), common border, common language and past or present colonial link; instead of pair FE. Finally, column (5) introduces the same "gravity" variables in the model that allows for heterogeneous effects. The dependent variable in all specifications but (2) is the number of patents and the estimator is PPML in all columns. Standard errors in parentheses are clustered by country pair in all columns. $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Table A3: Preferential Agreements and Cross-border Patents: Sectors

	(1) ALL	(2) S1	(3) S2	(4) S3
RTA_TECH_S_N	0.189	0.280	0.161	0.210
	(0.050)**	(0.068)**	(0.069)*	(0.074)**
RTA_TECH_S_S	-0.443	-0.367	-0.546	-0.260
	(0.192)*	(0.242)	(0.267)*	(0.291)
RTA_TECH_N_N	0.222	0.153	0.183	0.278
	(0.029)**	(0.043)**	(0.040)**	(0.047)**
RTA_TECH_N_S	0.118	0.011	0.097	0.156
	(0.048)*	(0.125)	(0.061)	(0.084)$^{+}$
RTA_NO_TECH_S_N	-0.758	-0.322	-0.494	-1.342
	(0.453)$^{+}$	(0.462)	(0.643)	(0.540)*
RTA_NO_TECH_S_S	0.800	1.285	0.667	0.908
	(0.223)**	(0.533)*	(0.281)*	(0.382)*
RTA_NO_TECH_N_N	0.596	0.576	0.562	0.628
	(0.068)**	(0.106)**	(0.096)**	(0.112)**
RTA_NO_TECH_N_S	-0.205	-0.124	-0.189	-0.358
	(0.106)$^{+}$	(0.152)	(0.123)	(0.263)
TRIPS_S_N	0.113	0.273	0.153	0.040
	(0.143)	(0.165)$^{+}$	(0.170)	(0.279)
TRIPS_S_S	0.413	0.626	0.405	0.440
	(0.161)*	(0.317)*	(0.220)$^{+}$	(0.246)$^{+}$
TRIPS_N_N	0.197	0.101	0.252	0.148
	(0.086)*	(0.124)	(0.117)*	(0.130)
TRIPS_N_S	-0.085	0.149	-0.111	-0.046
	(0.131)	(0.132)	(0.146)	(0.269)
PCT_S_N	1.027	0.610	1.271	0.681
	(0.233)**	(0.326)$^{+}$	(0.293)**	(0.541)
PCT_S_S	1.624	1.912	1.621	1.482
	(0.202)**	(0.267)**	(0.262)**	(0.380)**
PCT_N_N	0.638	0.727	0.685	0.548
	(0.060)**	(0.086)**	(0.082)**	(0.107)**
PCT_N_S	0.248	0.465	0.223	0.295
	(0.137)$^{+}$	(0.220)*	(0.167)	(0.320)
N	205585	64463	79791	61331

This table reports results for disaggregated data. The estimates are obtained from specification (23), after allowing for the effects of globalization to vary across four bilateral groups ('North to South', 'South to South', 'North to North', and 'South to North', which are based on the income classification of the World Bank. In column (1) results are presented for all manufacturing sectors at 2 digit-level of the International Standard Industrial Classification (ISIC). Column (2) presents the result for sectors 15-19. In column (3) for sectors 20 to 29. In column (4), sectors 30-37 are grouped. The dependent variable in each specification is the number of patents and the estimator is PPML. Standard errors in parentheses are clustered by country pair in all columns but . $^{+}$ $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.