Patents' time to grant - A study of China and India

Satyaki Chakravarty*[1], Shubhashis Gangopadhyay†[2,3,4,5], and Taylan Mavruk‡[3]

[1]University of North Carolina Greensboro, USA

[2]India Development Foundation, India

[3]Gothenburg University, Sweden

[4]University of Groningen, Netherlands

[5]Indian School of Public Policy, India

This is an older draft. Please click here for the latest version.

Abstract

China and India reported highest growth in high-tech exports for the past two decades (World Bank) and hence are increasingly becoming *the* countries to file for protection of intellectual property. This paper examines whether there is a delay in time to grant of a patent when applied by domestic versus foreign applicants in China and India. A delay in time to grant on the basis of nationality violates the TRIPs agreement. Further, delay in grant of a patent generates dead-weight loss for the society, hence effectively deterring future innovation. The results show that China delays foreign patents, however India does exactly the opposite, indicating that there is a significant variation in patent examination times across applicant characteristics, technology fields and countries. Thus invisible trade barriers do exist in these markets. Applicants can use the details in this paper and plan their

*s_chakr2@uncg.edu

†sg@idfresearch.org

‡taylan.mavruk@handels.gu.se

filing strategies, e.g. taking into account average grant lag.

JEL codes: F53, F68, K39, L52, L63, O25, O34

Keywords Patent grants, statistical discrimination, China, India, duration analysis

1 Introduction

This paper examines and compares the delay in grant of patents applied for by nationals and foreign applicants in India and in China. Delay in grant of a patent would generate dead-weight loss for the society because of its unrealized spillover effects. Delay of grant would also negatively affect comparative advantage for firms, which is related to how quickly the patent becomes impactful and generates citations. China signed the TRIPs (Trade Related Intellectual Property Rights) agreement at the end of 2001, while India signed it at the start of 1995. While for China, TRIPS came into effect immediately after signing, India, being a developing country, was allowed a 10 year period to become TRIPS ready. India, therefore, became TRIPS compliant in 2005. Thanks to Article 3 (National Treatment) and Article 4 (Most Favoured Nation) of the agreement, both countries agreed upon equal treatment to foreign and domestic applicants and any favour granted to any member of the treaty is immediately applicable to all other members. Hence, unlike other commodities, a nation cannot put barriers to entry of patents not produced by nationals.

The time taken to grant a patent by any country, once an application has been submitted in that country, will depend on a number of factors covering (a) the number and quality of personnel working in the national patent office(s) and (b) the volume and quality of the applications (see Schuett 2013). Thus, on average, patent applications may take different amounts of time before they are granted in different jurisdictions. Indeed, even a particular patent may take less time in one jurisdiction compared to another, simply because the patent office in one country is more adept at the technologies covered by a patent, compared to another country. However, the statistical discrimination against foreign patent applications have been well documented by Rassenfosse et al. (2019) and Webster, Jensen, and Palangkaraya (2014) for various patent offices accross the world.

Given Articles 3 and 4 of the TRIPS agreement, the time taken to grant patents by a country should not, in principle, be different for foreign and domestic applicants. If, at all, at least some foreign patent applications should be granted faster than similar applications from domestic entities. Why? Most countries require that national entities first apply in their own jurisdictions before applying to other jurisdictions. This is because as part of the national strategic requirements, countries require entities within their jurisdictions to apply first in these countries. This is because countries may want to hide certain breakthroughs in ideas which have strategic value to the country.[1] If such ideas are allowed to be patented, the idea will become common knowledge since patent applications, under uniform patent laws under TRIPS, must be published within 18 months (see Deepak and Luo 2018). Consider countries A and B both of which have signed onto TRIPS. Suppose a patent is applied for in country B by an entity from A. Then the patent office in B knows that A is already considering whether this patent can be granted. Since A and B are both TRIPS signatories and have uniform patent regimes, B need not duplicate the efforts being put in by A's patent office. Instead, if B's patent office resources are already stretched, it may not want to go through the same levels of investigation since A is already putting in that effort. This may quicken the time for granting a patent in B that has already been granted in A.

Alternatively, a patent office may deliberately delay a patent applied for by foreign

1. The decision to file for patents in a foreign country could also be influenced by the prior applications of actors in other 'reference' third-countries, such as countries located in the similar geographic macro-region and countries with a similar export product structure. Thus there could be spatial dependence in the geography of crossborder patenting, which may also lead to a faster grant for some foreign patents (see Perkins and Neumayer 2011). The discussion may extend to the citations. Localization of citations may not be driven by localized knowledge transmission, inventors may build on remote inventions (see Arora, Belenzon, and Lee 2018).

applicant to suppress competition for a short time for the domestic patent companies to work upon and create a substitute of that suppressed patent. In this context, strategic delay in granting of a patent becomes important. Strategic delay here refers to hindrance created for foreign applicants to help domestic innovation prosper (Maskus 2000). Delay in granting due to few patent examiners itself makes way for opportunities to infringe and/or innovate upon a yet to be granted patent and effectively render the patent useless even before it is granted. A strategic delay is useless only if the following happen:

1. no substitutes are generated in the process of delay;
2. the patent is granted exactly as it was applied for, that is to say each and every claim in the patent application is granted.[2]

Suppose a patent application is infringed i.e. the claims of an applied patent documents are infringed. If the set of claims in the applied version of the patent document is different from the set of claims in the granted version of the patent document, the infringement claim by the inventor may become weaker.

Though not many empirical studies exist which deals specifically with time to grant and national treatment of intellectual property by countries, there are a few papers worth mentioning, more so, since this paper borrows and extends ideas discussed in those papers; specifically from Rassenfosse et al. (2019), Rassenfosse and Raiteri (2016), Rassenfosse, Raiteri, and Bekkers (2017), and Harhoff and Wagner (2009). The first empirical paper pointing out an association between nationality and grant of patent was by Yang (2008), which finds that US does not prefer domestic applicants to foreign applicants; in other words no significant difference in probability of grant if the applicant is domestic versus

2. Infringement is considered upon claims of a patent and not the patent as a whole. After a patent is applied for protection, it is possible that only some of the claims are granted protection and some are not. Alleged infringement by patent holder on the claims which were not granted protection cannot be called infringement. This has been confirmed through patent lawyers in India, though similar logic can be found here: http://euro.ecom.cmu.edu/program/law/08-732/Patents/PatentLitigation.pdf

foreign. On the other hand, China provides preferential treatment to domestic applicants in terms of higher probability or certainty of patents being granted as compared to foreign applicants, as shown by data from 1985 to 2002.

Webster, Jensen, and Palangkaraya (2014) focus on European Patent Office (EPO) and Japan Patent Office (JPO) and show greater probability of grant for domestic applicants in both countries for the years 1990 to 1995. Further, they show that if Europe and Japan have an edge in certain technology domain than other countries, then the domestic inventor have stronger probability of grant in that technology domain. Rassenfosse et al. (2019) acts as an extension to Webster, Jensen, and Palangkaraya (2014), incorporating patent applications from three additional offices, viz. China National Intellectual Property Administration (CNIPA), Korean Intellectual Property Office (KIPO) and United States Patent and Trademark Office (USPTO) for the years 2000-2006. The paper concludes that there exists a statistical bias against foreign applicants in each of the offices studied. The bias is reduced if the foreign applicant follows the route of Patent Cooperation Treaty (PCT) to file their application.

Rassenfosse and Raiteri (2016), Rassenfosse, Raiteri, and Bekkers (2017), and Thoma (2013) discuss rejection and delay in granting of patents exclusively for CNIPA. Specifically, Rassenfosse and Raiteri (2016) examine technology domains which are relatively important to China in terms of their expertise and show that probability of grant of patents related to those technology domains decrease when applied for at CNIPA. Rassenfosse, Raiteri, and Bekkers (2017) show that probability of grant of a patent decreases and a granting of patent is delayed when applied for at CNIPA if declared as a Standard Essential Patent (SEP) before entering the examination phase at SIPO. Thoma (2013) finds that the probability of grant for foreign multinational firms is less than domestic firms although foreign patents are stronger regarding the prior art. The study on delay in grant of patents by Liegsalz and Wagner (2013) uses a shorter period of time, i.e. 1990 to

2002 and follows a different method, a method which this paper follows as well, modelling time to grant as survival function to show that there indeed exists a gap in the time to grant of patents when it comes to domestic versus foreign applicants.

Some empirical papers deal with value of patent and its time to grant in various countries. Harhoff and Wagner (2009) use survival analysis to show that valuable patents are likely to be granted faster than the non-valuable ones. The problem of determining value of patent was first addressed by Harhoff, Scherer, and Vopel (2003), attacking the problem in the most direct way possible; finding out the actual value of a patent in terms of its price. Impact of citations on stock market valuations computed by Hall, Jaffe, and Trajtenberg (2005) show that increase in one citation boosts market value by 3 percent. One possible explanation for this result is that a larger number of citations, used as a proxy for quality and innovation efficiency, provides communication, allowing investors to look beyond the short-term profits of firms to their intangible assets and strategy. In other words, it allows monitoring the firms to understand the long-term firm value. Gambardella, Harhoff, and Verspagen (2008) try to explain value using citations, references, claims and countries in which patents were applied for. Régibeau and Rockett (2007) deals with a similar question on value of a patent theoretically and corroborate their findings through a smaller dataset on US genetically modified crop patents. Their main finding is that when uncertainty in future value of a patent reduces, the time taken to grant a patent also reduces. Time taken to grant is also reduced when its importance increases. The quality of a patent, or its importance and anticipated value, is defined according to citations, size of patent family, number of claims, the number of years for which the patent was renewed along with other variables used in the literature. In our paper, we also use these variables but interpret them as indicators of complexity, which, in turn, could be related to delay in time taken to grant. For example, when a patent has a large number of novel claims, a patent office needs to evaluate each claim made in the application. For a resource poor patent office, this may take more time than one with a lesser number of claims. Alternately, a higher 'quality' patent may have more claims than one of a lesser quality, because it is usable in a number of different value generating products. Lemley and Sampat (2008) find that over time patent offices become overworked and increasingly grant "bad patents". Firms are likely to learn some of the characteristics of a "good patent" and mimic those without having quality. Thus we refer to the control variables used in this paper as indicators of complexity.

Our paper differs from previous studies in three important ways. First, we compare the working of the patent offices in India and China. To ensure comparability, we study both the countries during the time that they have been TRIPS compliant. Second, and related to the first, we have a larger dataset as we cover a longer time period, ending with a more recent end-date, than previous studies. Third, we compare differences in the time taken to grant between domestic and foreign applicants for the two countries given the patent has spent t days in an office.

Formally, the paper tests the following null-hypothesis:

H_0: Probability of grant of a patent on any day, given it has spent t days in office, is the same for domestic and foreign applicants.

We test this hypothesis with data collected from Derwent Innovation for patents applied and granted in China and India from 2000 to 2016 and from 2005 to 2016, respectively. These periods correspond to the time when each of the countries has been TRIPS compliant. We test the hypotheses using survival analysis and find that the hazard rate of grant for domestic applications in India is less than foreign applications in India, suggesting that foreign applications get granted faster. Unlike the results for India, the hazard rate for domestic applications in China is higher than foreign applications in China, indicating that in China there exists statistical discrimination against foreign patent applications. Further, we consistently find that the delay in granting foreign patents is especially prominent if a patent belongs to a technology domain in which Chinese patents have relative expertise compared to all other patent applicants in China. For India, relative expertise in certain technology domains does not seem to differentially affect domestic and foreign patents.

The results suggest that these two countries differ in their treatment of domestic and foreign patent applications in terms of the time taken to grant patents.

The paper is organised as follows: Section 2 first provides information on overall patent activity, second it describes the data used in this study, and finally, it specifies the econometric model. While Section 3 discusses the results from granted patents only, Section 4 discusses the results from granted and pending patents. Finally, Section 5 concludes the paper by discussing the implications.

2 Data and method

2.1 World patenting activity

Patenting activity all over the world has increased dramatically. Some studies argue that this growth is related to the different tax policy or other patent regimes at the country level, which influenced the extent and location of innovation and patent ownership (Bradley, Dauchy, and Robinson 2015; Sanyal 2015). Table 1 presents some statistics about this growth. In 2016, more than 3.1 million applications were filed around the globe (WIPO 2017). Between 2006 and 2016, Asia grew the most among the top 20 patent offices ranked by the number of applications filed. During this period, Asia grew by 8.50 per cent per year, followed by North America, Latin America, Africa, Oceania and Europe. However, if China and Japan are taken out of Asia, the remaining jurisdictions of Hong Kong, India, South Korea, Russia and Singapore, together grew by 2.51 per cent per year. This is less than the Latin American growth rate (World Intellectual Property

Organisation Statistics Database). The numbers are not surprising, given that the growth of high technology exports from these countries surpassed others during this period, confirming the results from Hall, Jaffe, and Trajtenberg (2005), Arora, Gambardella, and Fosfuri (2008), Lybbert and Zolas (2014), Kanwar and Hall (2017), and Deepak and Luo (2018) who suggest that patents facilitate the development of markets for technology, infuses more information into the market, augments trading opportunities, and enhances Tobin's q. High technology products are those developed using high Research and Development intensity such as aerospace, computers, pharmaceuticals, scientific instruments and electrical machinery. Table 2 shows growth in high technology exports.

[Insert Table 1 here]

[Insert Table 2 here]

China has been a major driver of growth in Asia. In terms of high technology exports, between 1996 and 2006, China grew by 32.96 per cent annually, while between 2006 and 2016 its growth was 6.15 per cent. In the same two periods, India grew by 11.36 per cent and 10.58 per cent, respectively. India grew the most in high technology exports between 2006 and 2016. However, India has much smaller absolute numbers compared to China. We also observe that though growth rate of Iran seems to be the highest, in absolute terms the country cannot be put on the same page as other countries. The difference between average high-tech exports of all countries in Table 2 barring Iran and the high-tech export of Iran is about 81 billion current US dollars for the year 2011. The difference if calculated for 1997 comes to be about 35 billion. The differences are by far the largest compared to any other country.

2.2 Data for this study

Our primary data source is Derwent Innovation. We cover all patents applied for in China from 11/12/2001 (the date China became TRIPS compliant) to 31/12/2016, the last date of the data with us. For India, we use all patent applications from 31/12/2004 (the date when India became TRIPS compliant) till 31/12/2016. The total number of patents for both the countries is 7,210,968. Some patents are published more than once in the application phase (or before being granted) because of modifications in the patent document after its first publication. We use only the latest version of each patent publication and remove the references to all earlier versions. This led to the removal of many repeated observations. We also drop data with missing or incomplete or entries with application date coming after the publication date which leaves us with 6,855,770 unique patent applications.

Both India and China require domestic entities to apply to their respective domestic patent offices before applying for patent protection in other jurisdictions. Indeed, almost all countries require their domestic entities to seek permission from their domestic patent offices to allow patent applications to be published. This is because every country, in its national strategic interest, may want certain ideas to be kept secret. This allows us to classify Indian (Chinese) domestic patents to be those which have been first applied for in India (China) and foreign patent applications to be those which were first applied for outside India (China).[3]

Family size considers all the countries where the patent has been applied for plus all patentable improvements (of the original patent also called "parent" and "children" of

3. For China, see (Article 9) in http://english.cnipa.gov.cn/art/2014/3/31/art_1349_81671.html (accessed on 12/9/2020); for India, see (Chapter VII Secrecy Directions; point number 71 of page 32) in http://www.ipindia.nic.in/writereaddata/Portal/IPORule/1_70_1_The-Patents-Rules-2003-Updated-till-23-June-2017.pdf (accessed on 12/9/2020); for the USA, see (U.S.C. 184 Filing of application in foreign country) in https://mpep.uspto.gov/RDMS/MPEP/e8r9#/e8r9/d0e304599.html (accessed on 12/9/2020).

the focal patent) applied for in all countries.

We define the variable *Grant lag* as the number of days the application has been in the patent office starting from the application date till being granted.

Finally, we generate three technological domain groups — pharmaceuticals which covers 5.44% of the sample observations; technology which covers 18.08% of the sample data; and, the rest covering 76.47% of the data.[4] Patents at times are assigned in multiple technology domains. We choose the technology domain which has the highest frequency for a patent application. If for a patent application technology domains' frequencies tie for two or more technology domains, we choose one at random. These variables are further defined in Table A1.

2.3 Descriptive statistics

Table 3A gives the number of patent applications that are pending (yet to be granted) and those that have already been granted, in India and China. For each of the countries, the period covered is from the time they became TRIPS compliant till December 31, 2016. The first thing that jumps out is the difference in patent activity between China and India (as measured by the number of applications since becoming TRIPS compliant). First, China has significantly more patent activity compared to India; each year on an average 679,710 patents are filed in China compared to 34,990 for India. Second, in China, domestic invention is more intense compared to patent applications from outside

4. WIPO (World Intellectual Property Organization) defines 35 technology domains on the basis of IPC (International Patent Classification). We aggregate the 35 domains into three groups in the following way:
Technology: audio visual technology, basic communication processes, computer technology, digital communication, IT methods for Management and telecommunications
Others: analysis of biological materials, basic materials chemistry, biotechnology, chemical engineering, civil engineering, control, electrical machinery, energy, engines pumps and turbines, environmental technology, food chemistry, furniture games, handling, machine tools, macromolecular chemistry and polymer, measurement, mechanical elements, medical technology, metallurgy, microstructure and nanotechnology, optics, organic fine chemistry, other consumer goods, other special machines, semiconductors, surface technology, textile paper machines and thermal processes apparatus

the country; in India, on the other hand, foreign patents are significantly larger in number compared to domestic applications. Moreover, in China the number of domestic granted patents per year is approximately twice as the foreign grants whereas in India the number of foreign patents granted per year is 13 times larger than the domestic patents.

[Insert Table 3A here]

Table 3B gives us the descriptive statistics for granted patents in China and India by the nationality of applicants. T-tests with unequal variance are conducted to test for unconditional difference in means across domestic and foreign applicants in India and China. It is clear that there are significant differences in the average values of the variables for each group in each country. The most relevant difference for this study is that the time taken to grant a foreign patent in China is 692 days larger than that for a domestic patent application (as measured by Grant lag); in India, it is the opposite (-173 days)! The other difference we observe is that foreign patents tend to be qualitatively different from domestic patents in both China and India. Higher mean values for these variables could signal greater complexity or, superior quality, or value. Greater complexity would require patent officers to spend more time in unravelling the applications.

[Insert Table 3B here]

For completeness, we present the means of the control variables for all applications — those already granted and the ones that are yet to be granted. This is presented in Table 3C. Variable "Grant lag" and grant lags by technology domains are not in Table 3C, but all the other variables are the same as in Table 3B. While most of the variables have similar signs in case of difference between foreign and domestic patents for the two tables, the magnitudes are different. Similar to Table 3B, we observe foreign patents to be qualitatively different from domestic ones. In case of patents applied for in India, family size for foreign patents is higher compared to domestic patents, whereas for the granted patents, family size for foreign patents is lower compared to domestic patents. On an average, we observe higher variation in case of China in the two types of patent documents; granted patents and all applications.

[Insert Table 3C here]

The last three rows of Table 3B shows the average time taken to grant patents in pharmaceuticals and technology, compared to all other domains. While grant lag is larger for foreign patents in all domains in China, in India, the grant lag is larger for domestic patents in all domains. We observe that in case of pharmaceuticals, both India and China grant domestic patents quicker than domestic average (about 26 and 99 days quicker respectively) and for technology patents, both the countries take longer to grant domestic patents than the domestic average (about 240 and 392 days longer respectively). In case of foreign pharmaceuticals, China takes longer by about 398 days while India grants them quicker by 137 days compared to the foreign average. For foreign technology, both the countries take longer time to grant the patents compared to the average. Both pharmaceuticals and technology are important for China and India and so we consider them separately in our analytical exercises.

2.4 Method

Our starting point is that if a country is TRIPS compliant, then the time taken to grant a patent is determined by the volume of activity (number of patents applied for in any given period), the resources (including personnel) in the country's patent office(s) and the complexity of the patent applications. In particular, the time to grant is not determined by whether the application is by a foreign entity or a domestic entity. If there is any difference in the time taken to grant a domestic patent vis-a-vis a foreign patent, one reason could be that domestic patents are systematically different in their complexity compared to foreign patents and, hence, take different amounts of time for the granting process to be completed. If this was the case, then correcting for a patent's complexity will explain away the time difference in granting patents. If differences persist after correcting for complexity, it could be due to a country's differential treatment of domestic applications vis-a-vis foreign applications.

If Y_{ij} is a measure of the time taken to grant a patentable application i in country j, we postulate that it depends on two sets of variables: (a) the identity of the entity applying for the patent (domestic or foreign) and (b) the complexity of the patent. In particular, our presumption is that while an application may take more time to be granted in country i compared to country j due to country specific factors, the time taken to grant a foreign application is no different from that of a similar domestic applicant within the same country. The 'similarity' of two patent applications is measured by their degree of complexity. Various previous authors have used different sets of variables to identify the complexity, or quality, of a patent. For us they constitute the variables in set (b) above and these are our control variables. Our main focus is the sensitivity of Y_{ij} to the set of factors (a). The list and definitions of our control variables, or the set of factors in (b), is presented in Table A1.

Our exercise can be broadly divided into two parts depending on the sample we use. For both parts, our data cover the period when the two countries — China and India — became TRIPS compliant till the end of 2016. One part of the data exercise involves applications on which patents have already been granted by December 31, 2016. The other part consists of granted patents as well as those that were yet to be granted as on December 31, 2016.

For the sample of granted patents, we conduct three exercises — a non-parametric estimate of the probability of an application being granted on any given date given that the application is pending, a la Kaplan-Meier; a model where we regress Y_{ij} against factors (a) and (b) above; and, finally, a hazard rate model where the hazard 'event' occurs when a patent is granted (with the same controls as in the regression model). In the second part of the exercise we use the full sample (granted and yet to be granted). The main difference between the two parts is that in the first part, which has granted patents only, we can define the time taken to grant, or *Grant Lag*. This variable cannot be defined for pending applications (included in the second part) since there is no grant date for them. Hence, the non-parametric and regression models both of which use grant lag could not be carried out in the second part of the exercise. The hazard rate model, however, is implemented using both granted and yet to be granted observations.

3 Results with granted patents

3.1 Results from Non-parametric analysis

Figure 1 illustrates the probability of a patent being granted at time t given that it is yet to be granted. Each curve in Figure 1 represents an entity type. Generally, the probability of a patent application being granted increases with time. However, given the length of time that an application is pending, the probability of grant is highest for a foreign application made in India (IF), followed by a domestic patent in India (ID), followed by a domestic application in China (CD) with a foreign application in China having the lowest probability (CF). This is the Kaplan-Meier estimator of the hazard function where the 'event' happens when the patent is granted. We observe in the figure that hazard rate is increasing in duration. In summary, the Kaplan-Meier estimates imply that foreign applications in India are granted faster than domestic applications in India but the reverse is true for foreign and domestic applications in China.

[Insert Figure 1 here]

3.2 Results from OLS

Figure 1 assumes that all granted patents are similar, differing only in the nationality of the entity applying for the patent. In reality, foreign applications in China (India) may be systematically different from domestic applications in China (India). For example, foreign applications in China may be more complex than domestic applications while domestic applications in India may be more complex than foreign applications. Complex patents take more time to process and this systematically differing complexity of each group could be the reason for the Kaplan-Meier diagrams appearing as they do. Unravelling this constitutes our main exercise.

To figure out how complexity of the applications affects the results, we investigate how sensitive to application characteristics is the time taken to grant a patent. In other words, does a more complex application need longer processing time? If foreign patents are systematically more complex than domestic patents in China then this could be the reason why foreign patent applications in China take more time to be granted. Our dependent variable in this exercise, is 'Grant lag' or the number of days between the application date and the grant date. We test the following model:

$$
\begin{aligned}
[Grant\ lag]_i = \alpha &+ \beta_1 ID_i + \beta_2 CD_i + \beta_3 CF_i \\
&+ \beta_4 [Nb\ \ indep\ \ claims]_i + \beta_5 [Tot\ \ wordcount\ \ ic]_i + \beta_6 [Inv\ \ count]_i \\
&+ \beta_7 [Wordcount\ \ per\ \ c]_i + \beta_8 [Dep\ \ per\ \ ind\ \ claim]_i \\
&+ \beta_9 [Family\ \ size]_i + \beta_{10} [group]_i + \beta_{11} [RTA]_i + \varepsilon_i,
\end{aligned}
\tag{1}
$$

The control variables 'Nb independent claims', 'Tot wordcount ic', 'Inv count', 'Wordcount per c', Dep per ind claim' and 'Family size' have selectively been used by various authors as measures of complexity.[5] Independent claims are standalone statements describing the function of an invention. Higher the number of independent claims, higher is the scope of a patent. Number of words per independent claim and wordcount per claim also indicate scope. Dependent claims list limitations as well as elaborating on independent claims. Several dependent claims per independent claim indicates elaborate explanation of a patent. Family size of a patent indicates the number of jurisdictions the patent has been applied for, as well as the number of connected patents in terms of improvements. Higher the size of family, higher is the potential commercial value of the patent, signalled through the cost that applicant incurred in applying at multiple jurisdictions. RTA indicates relative expertise of domestic inventors relative to all inventors applying for in a country. Value of RTA greater than 1 indicates a higher proportion of patent filing by domestic inventors of country j in technology domain i for year t. Several papers, including Harhoff and Wagner (2009) and Liegsalz and Wagner (2013) have used similar control variables.

The variable 'group' denotes the technology domain of the patent. In section 2.2 we have described the classification of patents into those pertaining to pharmaceuticals,

5. The definition of these variables are given in section 2.2.

technology and to the remaining technology domains (or, 'others').

Patents may enter the study at different times but we observe some clustering around certain months and thus we cluster standard errors on grant months rather than treating each grant date independently.[6] Clustering on grant weeks yield similar results. We also include fixed effects for publication year. All the continuous variables are winsorised at 1 percent and 99 percent levels to rule out outliers. Table 4 presents the results.

[Insert Table 4 here]

In column 1 we report estimates of equation 1 without the variable 'group' while column 2 reports all the estimates of equation 1. Column 3 and 4 report results using RTA and RTA interacted with ID, CF and CD. Test of differences in the coefficients are presented at the bottom of the table. For column 1, our reference category is IF and for column 2 the reference technology is pharmaceuticals (while reference category is still IF). The conclusions drawn from the Kaplan-Meier diagram continue to hold making our results comparable to the results documented by Liegsalz and Wagner (2013). In column 2, we observe that grant lag is about 243 days longer for domestic patents in India. While grant lag is less for both foreign and domestic patents in China, it takes a lot less time (293 days) for domestic patents than for foreign patents in China compared with foreign patents in India (the reference group). Column 3 shows an increase in grant time if the domestic patents have higher relative expertise compared to other patents in a country in a particular technology domain on an average. However, we do not know what happens to RTA by countries and hence we interact RTA with our groups of ID, CF and CD in column 4. Column 4 shows if domestic inventors of a country have relative expertise in

6. Firms strategy on filing patents cluster around certain months which are independent to firms. However, examiners release examined patents in batches. We observe the clustering by computing variation in publication lag (defined as publication date of patent application prior to grant minus application date) by months which motivate our model to be clustered on publication month.

a technology domain, they are granted faster. This is true for India as well as China. For India, domestic patents are still delayed on an average by 144 days. But among the domestic patents, if the patents belong to an advantageous technology domain, they are granted quicker. For China, after controlling for RTA, the difference between domestic and foreign patents reduce and come down to 38 days. However, grant lag increases if a foreign patent is applied for in a technology domain in which domestic patents in China have relative expertise. Domestic patents are also granted quicker by about 80 days if they belong to those technology domains.

3.3 Results from hazard rate of grant

Here we assume that the probability of a patent being granted after spending time t in the patent office without being granted, is given by

$$\begin{aligned}\lambda(t,X)_i = \lambda_0(t) exp(&\beta_1 ID_i + \beta_2 CD_i + \beta_3 CF_i \\ &+ \beta_4[Nb\ indep\ claims]_i + \beta_5[Tot\ wordcount\ ic]_i + \beta_6[Inv\ count]_i \\ &+ \beta_7[Wordcount\ per\ c]_i + \beta_8[Dep\ per\ ind\ claim]_i \\ &+ \beta_9[Familysize]_i + \beta_{10}[group]_i + \beta_{11}[RTA]_i)\end{aligned} \tag{2}$$

where $\lambda(t,X)$ is the hazard rate of grant and X is the set of control variables. $\lambda_0(t)$ is the base line hazard rate and this is assumed to be increasing in t. We assume that the baseline hazard follows a Weibull distribution ($\lambda(t) = \lambda\gamma t^{\gamma-1}$), which allows both, flexibility of the model and different shapes of the hazard function. In the case of $\gamma = 1$, the exponential distribution can be obtained, but in our data we observe that hazard rate monotonously increases from zero at time zero to ∞ at time ∞ for $\gamma > 1$. [7] Also, we treat failure time as if it is continuous, not as being divided into discrete chunks or units, because our unit of analysis is patent id and day and we do not observe multiple events of the same patent id within each day. See for instance Azoulay, Ding, and Stuart (2007), who examine the individual, contextual, and institutional determinants of academic patenting, for a careful implementation of discrete time hazard rate models.

We use Akaike Information Criterion ($AIC := -2log(L) + 2(p)$) as a goodness-of-fit measure for choosing among the parametric models, although in Panel A of Table 5 both AIC and Bayesian information criterion (BIC) give similar results. Since we have many covariates in the models, AIC might better identify the best model among the candidates because it imposes a greater penalty for the number of parameters.[8] In AIC p is the number of model parameters and L is the model likelihood function. Our main focus is on parametric models among which, we examine and compare AIC for Weibull and Exponential distributions. Cox proportional hazard model runs the risk of absorbing any specification error which may be present in our model. Therefore, we only report those results as robustness check.[9] We do not consider lognormal (or anti-log of normal) distribution because it is evident from Figure 1 that the general form of hazard does not seem to increase to a maximum value and then begin decreasing toward 0.

The smaller is the Akaike information criterion (AIC), the more efficacious will the model be in identifying how the hazard rate differs among our groups. Although in general AIC values are high, Table 5 shows that Weibull models have the smallest AIC value (8,298,153) among parametric distributions which motivates the choice of these two models in our main analyses.

7. The hazard parameters in Table 7 shows that hazard is increasing in duration (For column 4, Lnp=0.49 and p= 1.6) and this is statistically significant (z-stat: 25.40)
8. See Burnham and Anderson 2004 for a comparison between AIC and BIC.
9. See section 4.4

[Insert Table 5 here]

The hypothesis is tested based on the estimation of β_1-β_3 of the key variables defined above, e.g. β_1 would indicate the increase or decrease in the instantaneous probability of grant of domestic application relative to foreign application in India, given survival to date. This estimates the impact on the hazard rate, rather than on the average duration time. A decrease in hazard means increase in the duration time. Put differently, decrease in hazard means more patent applications without decision or delay in grant.

As in the preceding exercise, we report clustered (around publication month) standard errors and Winsorize all continuous variables at 1 per cent and 99 per cent levels to rule out outliers.

4 Results with granted and pending patents

So far, we have used the information on granted patents only, which ignores the denied applications and may introduce bias in the analysis (see Balsmeier et al. 2016 for a detailed discussion). Our analysis does not use information on patents that are yet to be granted. To incorporate these, we take all applications till December 31, 2016 regardless of whether they have been granted or are still pending.[10]

4.1 Resuts from logistic regression

We define a variable *Granted* that takes the value 1 if the patent has been granted and 0 if it is pending. We run the following logistic regression:

10. Ideally, we should be able to distinguish among applications that have been rejected and those that are still being considered. Unfortunately, few countries report 'rejected' applications but report all granted patents. Those that are yet to be granted could, therefore be rejected or still being processed. We assume that since there is a cost of filing an application, no one will file an application knowing that it will be rejected! So we take all patents that are yet to be granted as potentially 'grantable' but the process of granting them is not yet complete.

$$
\begin{aligned}
Granted_i = \alpha &+ \beta_1 ID_i + \beta_2 CD_i + \beta_3 CF_i \\
&+ \beta_4[Nb\ indep\ claims]_i + \beta_5[Tot\ wordcount\ ic]_i + \beta_6[Inv\ count]_i \\
&+ \beta_7[Wordcount\ per\ c]_i + \beta_8[Dep\ per\ ind\ claim]_i \\
&+ \beta_9[Familysize]_i + \beta_{10}[group]_i + \beta_{11}[RTA]_i + \varepsilon_i
\end{aligned} \tag{3}
$$

[Insert Table 6 here]

Table 6 reports marginal effects from these regressions. There is about 6% less probability of grant for domestic patents compared with foreign patents in India and the probability of grant for foreign patents in China is about 1 percentage point higher than domestic patents in China. These results confirm our findings in the earlier tables. Including the domain dummy in column 2 yields consistent results. In column 3 we include RTA and in column 4 we interact RTA with our groups. The results show that after using both granted and pending patents, estimates for India differ but for China they still remain qualitatively the same. One reason might be due to the assumption that the patents which are not yet granted and the patents which are rejected are treated same in this analysis, while they might have a qualitative difference which the estimates do not capture. Hazard rate on the other hand mitigates part of the problem by taking time into consideration. Here, controlling for RTA shows on an average domestic patents in India do not have a lower probability of grant. Contrary to the previous results, domestic patents with relative expertise in certain technology domains do not have a positive probability of grant. For China, foreign patents have a greater probability of grant than the domestic patents on an average. If the patent belongs to technology domains where Chinese inventors have relative advantage, probability of grant of foreign patent in those domains reduces while for domestic patents, it increases. This result is in line with previous results.

4.2 Results from hazard rate of grant using all applications

As is to be expected, some applications are yet to be granted on the last date of our data set (i.e., as on December 31, 2016). So the time to event (grant date) is incomplete for these applications, leading to right censoring in the analyses. This is the most plausible cause of censoring as it is less likely that applications have not been granted because applicants have not responded to patent office queries and dropped out. It is more likely that our study ends at a point when, for some applications, the grant date is not observed.

This is a classic administrative censoring issue. We deal with this problem by setting the censored observations to the last observed date in the data, allowing us to assume that decision to grant time and (administrative) censoring time are independent. This further allows us to include censored observations in the likelihood function and assume that censoring distribution contains information about the parameters β_1-β_3 contained in equation 2. We thus repeat the estimation of 2 but now we use the full sample of all applications.

Column 1 in Table 7 shows only the results from a model where the explanatory variables are ID, CF and CD without any controls. Column 2 includes the control variables in the model. In column 3 we include RTA and, as before, in the last column, we interact RTA with our groups. Standard errors are clustered on grant months. A statistically significant coefficient, e.g. on the ID dummy, would indicate a difference in the instantaneous probability of grant of domestic application relative to foreign application in India, given survival to date. The table also reports the test of pairwise coefficient differences between ID and CF, ID and CD and, between CD and CF.

[Insert Table 7 here]

In column 2 we observe that the hazard rate for domestic applications in India is about 17% less than foreign applications in India, suggesting that foreign applications get granted faster in India. The hazard rate is 152% higher for foreign applications in China and 182% higher for domestic applications in China relative to foreign applications in India, showing that patents are granted faster in China than in India (see also the significant *chi2* values of the difference tests in the table). However, unlike the results for India, China seems to differentiate against foreign patent applications by taking longer to grant them compared to their domestic applications. This is exactly the opposite to what we observe in India. In column 4, the hazard rate of domestic applications in India become insignificant, hinting on the workload of the office. In the same column, we observe that for China there is no difference between domestic and foreign patents' hazard rate. But, if a foreign patent is from a technology domain in which Chinese domestic patents have relative expertise, the hazard rate falls by about 40 percent.

The results from China are consistent with those in Yang (2008), Liegsalz and Wagner (2013), Thoma (2013), and Rassenfosse, Raiteri, and Bekkers (2017) showing that China appears to give preferential treatment to domestic applications. China behaves similarly to Japan as demonstrated by Kotabe (1992). Yang (2008) also shows that the US is equal in granting patents to domestic and foreign applications. Though, Rassenfosse et al. (2019) show that USPTO also favors domestic over foreign patents. The results may be linked to the studies that explore the relation between inventor mobility and the technological spillovers (see Agarwal, Ganco, and Ham Ziedonis 2007).

One potential explanation for our result from India could be that the uncertainty for foreign applications is lower relative to a comparable domestic application. This is due to foreign applications having been submitted to their own countries first.[11] It could also

11. See footnote 4 and the discussion leading up to it.

be that patent offices simply favour foreign applications because of increasing demand for foreign products and services in India. A higher RTA indicates higher proportion of applications relative to all applications for a particular technology domain. If a patent office is already strained on the number of examiners for a specific expertise, it is plausible that patents might get delayed for that domain. One explanation for the results on China could be that foreign applications in China have to be filed through a government designated agency and this may cause delay in the process. Yang (2008) compares foreign applications filed through local authority in China with domestic applications and still the results from China does not change.

Thus our current analyses cannot rule out the differences in the administrative requirements or in the drafting of patent applications. However, for the validity of our results we rely on the work by Liegsalz and Wagner (2013) and particularly Webster, Jensen, and Palangkaraya (2014) who control for inventor experience and the number ex ante claims at each patent office and arrive at similar conclusions as we draw regarding our results of the patent applications in China. Their results suggest that domestic patent applications are on average 10-16 percent more likely to be granted than foreign applications and these results seem to be robust for alternative measures of domestic patent applications. In particular, the address of domestic applications tend to be a more influential factor than using surnames of the applicants when determining a domestic applicant.

Our control variables capture the complexity of the patent applications. Barring variable "Number of independent claims", all other control variables significantly increase the hazard rate suggesting that these reduce the time to grant. The results hardly change if RTA is considered, indicating robust estimates for these control variables. The consistency in the estimates also rule out possible omitted variables which capture complexity of a patent in our analyses. "Number of independent claims" decreases the hazard rate by about 6%, which is consistent with findings in Marco, Sarnoff, and deGrazia (2016) who show that narrower independent claims at publication are related to a higher probability of grant and a shorter examination process than broader claims. "Total wordcount ic" increases the hazard rate by 0.05%, which may not be that small an effect as it measures the marginal effect for one additional word. Addition of an inventor ("Inv count") increases the hazard rate by 9%. A unit increase in the ratio of dependent claims over independent claims ("Dep per indep claim") leads to a 3% increase in the hazard rate. Finally, one more word increase in the average number of "Words per c" increases the hazard rate by 0.2%.[12]

4.3 Results from technological domain groups

Our next goal is to examine the effects of Pharmaceuticals and Technology domains on hazard rate separately. Column All in Table 8 excludes the Pharmaceuticals to be used as a reference group and the other columns examine each domain group separately.

[Insert Table 8 here]

When we include the domain groups in column All, we observe that the difference in the hazard rate of domestic applications in India and foreign applications in India disappears. The hazard rates for domestic applications and foreign applications in China remain qualitatively the same as in Table 8. Interestingly, the hazard rate for Technology is about 28% higher than Pharmaceuticals and it is about 65% higher for Other domains than Pharmaceuticals. Also, the hazard rate seems to be higher for Other domains than Technology. These results imply that patents in Pharmaceuticals and Technology are a lot more complicated and hence it takes longer time for them to be granted. This finding is consistent with results in Popp, Juhl, and Johnson (2003) who suggest that technological

12. A part of the results from the complexity measures could be driven by the fact that some patents are a type of voluntary standard setting organizations (SSOs) patents, which could receive many more citations than an average patent (see Rysman and Simcoe 2005).

types and complexities are significant factors to prolong the granting process in the US. It seems to take longer time to examine patent applications from e.g. biotechnology and computer industries, and electrical, mechanical relative to miscellaneous technologies because the former mentioned applications require a broader search of prior art due to their more complex nature. In China, the domain of patents does not seem to matter for the hazard rate of domestic and foreign applications.

When we examine the technological domains separately, our results remain the same as in Table 8 for Pharmaceuticals and Technology. Domestic patent applications in India takes longer time to be granted than foreign applications in India while the reverse is true for China. However, the hazard rate for patent applications in other domains does not differ between domestic and foreign applications in India but, again, domestic applications in China are granted about 31 percentage points faster than foreign applications. This result also confirms the ones obtained from column All that domains other than Pharmaceuticals and Technology lead to equal hazard rate of domestic and foreign applications in India.

Another observation is that statistical discrimination against foreign patent applications in China is about 146 percentage points higher in Pharmaceuticals (a simple difference of 1.64-0.18 in column Pharmaceuticals) but this value is about 22 percentage points higher in Technology, indicating that statistical discrimination against foreign patent applications is about 124 percentage points higher in Pharmaceuticals than in Technology in China. This is also true against other domains in China supporting the findings in Liegsalz and Wagner (2013) who show that Chinese applicants receive disproportionally faster patent grants in areas of high technological relevance for China, indicating favourable treatment for domestic applicants. The results are again reversed for India, foreign patent applications in Technology are granted at a rate of 28 percentage points (difference in absolute values: 0.6866-0.4106) faster than foreign patent applications in

Pharmaceuticals.

4.4 Robustness

4.4.1 Results from a frailty model

We consider parametric heterogeneity in the hazard model (a frailty model) and assume that the heterogeneity is independent of the covariates and enters the hazard function multiplicatively. For our Weibull hazard model this yields ($\lambda(t) = v_i \lambda \gamma t^{\gamma-1}$) and $E(v_i) = 1$, the standard normalization. We assume gamma distribution for the heterogeneity but test also the sensitivity of the results by assuming inverse gaussian distribution. However, the data do not show convergence using a frailty model with inverse gaussian distribution. Moreover, we do not have a reason to assume a shared frailty model as patent applications are independent from each other so frailties are not common.

The main model in Table 7 assumes that the hazard function is fully determined by the covariate vector. However, there may be some unobserved variables that violate this assumption. Omitting these variables generate unexplained heterogeneity which can be captured via a frailty model and thus we can examine why patent applications with lower hazard rates more "frail" than those with higher hazard rates.

Since we observe each patent application once in the data we have no reason to believe that multiple observations of the same application always has the same value of the unobserved heterogeneity. Therefore, we assume that the frailty is not shared among a group of patent applications. Table 9 revises the main analyses and the analyses with technological domains in a frailty model which assumes gamma distributed frailties.

[Insert Table 9 here]

The observed significant theta (presented as in Ln theta) in the table confirms the presence of heterogeneity. This is also confirmed in our Likelihood ratio test of theta against zero (untabulated). The results obtained from column 1 is similar to the ones obtained in Table 7 but the magnitude of the hazard rates become higher with a frailty model. Similar conclusions are drawn from column 2 in which we control for domains. Also, controlling for publication year fixed effects does not qualitatively alter the results. Taken together, controlling for unobserved heterogeneity via a frailty model our conclusions are the same, suggesting that there is a statistical discrimination against foreign patent applications in China whereas India seems to grant foreign patent applications faster than domestic patent applications.

4.4.2 Results from Cox proportional hazard model

Next, we use a semiparametric model (the Cox proportional hazard model), which allows us to make no assumption about the shape of the baseline hazard function. The model also requires less assumptions on the distributions of covariates than above used parametric models. Since our dependent variables do not vary over time we can simply assume that the hazard ratio that compares patent grants is constant over time and hence use the Cox proportional hazard model. The results are presented in 10.

[Insert Table 10 here]

The results from the Cox proportional hazard model appear to tell the same story. The hazard rate for domestic applications in India is about 16% less than foreign applications in India so, again, foreign applications get granted faster. The opposite is true for domestic applications in China, which get granted about 31 percentage points faster than foreign applications in China. When we control for RTA, similar to results from Table

7, we find no difference in hazard rate for Indian domestic patents. It seems domestic patents in India are applied for in technology domains in which India has relative expertise and due to possible backlog, we observe a delay, given by a reduction in hazard rate by about 13 percent. Foreign patents in China do seem to have a higher instantaneous probability of grant. However, it reduces by about 40 percent if it is from a technology domain in which Chinese patents have relative expertise. Additionally in those domains, domestic patents' hazard rate improves by 16 percent.

4.4.3 Results from Harhoff and Wagner (2009) model

In this section, we revisit two sets of analyses that are shown by Harhoff and Wagner (2009) and apply them to our data set. First, we collect a set of covariates that closely mimic the control variables used in Harhoff and Wagner (2009). One reason for not using these control variables in the main analyses is that they reduce the number of observations drastically. A short definition of these control variables extracted for each patent are given in Table A1.

We note that control variables such as share of X, Y and D citations used in Harhoff and Wagner (2009) are only applicable to European Patent Office and hence are not used here. Also, measures of generality and originality as developed by Hall, Jaffe, and Trajtenberg (2001) are not used here.

[Insert Table 11 here]

In Table 11, we observe that in each column ID is negatively and both CF and CD are positively related to the hazard rate and the effect of CD is significantly greater than CF on the hazard rate. The sign and the significance of the coefficients found in each column are consistent with what we observe in Table 7, confirming our earlier results. A majority of the results from the control variables is also consistent with findings in Harhoff and Wagner (2009).

The second analysis replicates the two stage results shown in Harhoff and Wagner (2009), which would also, to a certain extent, take into account a potential type of endogeneity problem in patent applications caused by complexity of the applications. In the first stage, in a probit model, we estimate whether a patent was maintained (from application date) for more than or equal to 10 years using a set of controls as identifying variables explained in column 5 of Table 11. This variable would indicate the estimated value or quality of a patent. In the second stage, we regress the estimated probabilities obtained from stage 1 on the control variables in a hazard model. Since data on maintaining patent for 10 years is available only for China, we cannot replicate the results for India. The results are shown in Table 12.

[Insert Table 12 here]

Column 1 shows that a patent has a higher hazard rate of grant if it is maintained for more than or equal to 10 years and thus the quality of a patent seems to matter for a quicker grant, which confirms findings in Harhoff and Wagner (2009) suggesting that the grant of valuable patents is accelerated. Consistent with the earlier results, patents in more complicated domains such as Pharmaceutical, Technology, and Biotechnology, compared with other domains, take a longer time to get granted. Column 2 confirms, again, our earlier findings that the hazard rate for a Chinese domestic patent is 131 per cent higher compared to a foreign patent applied in China.

5 Conclusion

Our main results show that there is a significant variation in patent examination times across applicant characteristics, technology fields and countries. The foreign patent applications in India are granted faster than domestic patent applications in India. The results are reversed for patent applications in China, indicating that in China there exists a statistical discrimination against foreign patent applications and that the national treatment principle is not being enforced in China.

Although we control for applicant characteristics and technology fields, the findings might be due to a greater familiarity of domestic applicants with the Chinese patent system. Regardless of the reason, applications from foreign and domestic applicants do not seem to receive equal treatment. Though the results are all indicative, invisible trade barriers do exist in these markets. Delay of patents not only just violates the national treatment principle, it causes serious ramifications for the near and 'invisible' future of innovation. On the other hand, a quicker grant of a patent would provide with comparative advantage not only to firms but also to investors, triggering them to become long-term-oriented shareholders. This is because grant of a patent would allow investors to understand long-term factors such as intangible assets and strategy that may lead to effective monitoring. This investor capital would in turn positively affect the firm value because through better monitoring, long-term oriented investors may prevent managers from aiming at short-term earnings targets and hence earnings manipulation. Our findings should be relevant for applicants, who can use these details and plan their filing strategies, taking into account average grant lag. Further research could examine familiarity of domestic applicants with the patent systems.

References

Agarwal, Rajshree, Martin Ganco, and Rosemarie Ham Ziedonis. 2007. "Reputations for Toughness in Patent Enforcement: Implications for Knowledge Spillovers via Inventor Mobility." *Strategic Management Journal* 30 (13): 1349–1374.

Arora, Ashish, Sharon Belenzon, and Honggi Lee. 2018. "Reverse Citations and Localized Knowledge Spillovers." *Journal of Economic Geography* 18 (3): 495–521.

Arora, Ashish, Alfonso Gambardella, and Andrea Fosfuri. 2008. "Patents and the market for technology." Chap. 4 in *Intellectual Property Rights, Growth and Trade,* edited by Keith Maskus, 123–156. Amsterdam and Oxford: Elsevier Press.

Azoulay, Pierre, Waverly Ding, and Toby Stuart. 2007. "The determinants of faculty patenting behavior: Demographics or opportunities?" *Journal of Economic Behavior & Organization* 63 (4): 599–623.

Balsmeier, Benjamin, Guan-Cheng Li, Tyler Chesebro, Guangzheng Zang, Gabe Fierro, Kevin Johnson, Aditya Kaulagi, et al. 2016. "Machine learning and natural language processing on the patent corpus: data, tools, and new measures." Working Paper, UC Berkeley.

Bradley, Sebastien, Estelle P Dauchy, and Leslie A Robinson. 2015. "Cross-Country Evidence on the Preliminary Effects of Patent Box Regimes on Patent Activity and Ownership." Tuck School of Business Working Paper No. 2681433.

Burnham, Kenneth P, and David R Anderson. 2004. "Multimodel Inference: Understanding AIC and BIC in Model Selection." *Sociological Methods & Research* 33 (2): 261–304.

Deepak, Hegde, and Hong Luo. 2018. "Patent Publication and the Market for Ideas." *Management Science* 64 (2): 652–672.

Gambardella, Alfonso, Dietmar Harhoff, and Bart Verspagen. 2008. "The value of European patents." *European Management Review* 5 (2): 69–84.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg. 2005. "Market value and patent citations." *RAND Journal of economics,* 16–38.

Hall, Bronwyn H, Adam B Jaffe, and Manuel Trajtenberg. 2001. *The NBER patent citation data file: Lessons, insights and methodological tools.* Technical report. National Bureau of Economic Research.

Harhoff, D, FM Scherer, and K Vopel. 2003. "Exploring the Tail of the Patent Value Distribution." Chap. 12 in *Economics, Law and Intellectual Property,* edited by O. Granstrand, 279–309. Boston, MA: Kluwer Academic Publishers.

Harhoff, Dietmar, and Stefan Wagner. 2009. "The duration of patent examination at the European Patent Office." *Management Science* 55 (12): 1969–1984.

Kanwar, Sunil, and Bronwyn H Hall. 2017. "The Market Value of R&D in Emerging Economies: Evidence from India." *The BE Journal of Economic Analysis & Policy* 17 (1).

Kotabe, Masaaki. 1992. "A comparative study of US and Japanese patent systems." *Journal of International Business Studies* 23 (1): 147–168.

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

Liegsalz, Johannes, and Stefan Wagner. 2013. "Patent examination at the state intellectual property office in China." *Research Policy* 42 (2): 552–563.

Lybbert, Travis J, and Nikolas J Zolas. 2014. "Getting patents and economic data to speak to each other: An 'Algorithmic Links with Probabilities' approach for joint analyses of patenting and economic activity." *Research Policy* 43 (3): 530–542.

Marco, Alan C, Joshua D Sarnoff, and Charles deGrazia. 2016. "Patent Claims and Patent Scope." USPTO Economic Working Paper 2016-04.

Maskus, Keith E. 2000. "Lessons from studying the international economics of intellectual property rights." *Vand. L. Rev.* 53:2219.

Perkins, Richard, and Eric Neumayer. 2011. "Transnational spatial dependencies in the geography of non-resident patent filings." *Journal of Economic Geography* 1:37–60.

Popp, David, Ted Juhl, and Daniel KN Johnson. 2003. "Time in purgatory: Determinants of the grant lag for US patent applications." NBER Working paper w9518.

Rassenfosse, Gaétan de, Paul H Jensen, T'Mir Julius, Alfons Palangkaraya, and Elizabeth Webster. 2019. "Are Foreigners Treated Equally under the Trade-Related Aspects of Intellectual Property Rights Agreement?" *The Journal of Law and Economics* 62 (4): 663–685.

Rassenfosse, Gaétan de, E Raiteri, and RNA Bekkers. 2017. "Discrimination against foreigners in the patent system: evidence from standard-essential patents." Working Paper, Ecole Polytechnique Fédérale de Lausanne, Switzerland.

Rassenfosse, Gaétan de, and Emilio Raiteri. 2016. "Technology protectionism and the patent system: Strategic technologies in China." Working Paper, Ecole Polytechnique Fédérale de Lausanne, Switzerland.

Régibeau, Pierre, and Katharine Rockett. 2007. "Are more important patents approved more slowly and should they be?" *CEPR Discussion Paper No. 6178.*

Rysman, Marc, and Timothy Simcoe. 2005. "Patents and the Performance of Voluntary Standard Setting Organizations." NET Institute Working Paper No. 05-22.

Sanyal, Paroma. 2015. "Quantifying Changes in Innovation: Patenting Activity and Ipr Regimes." Working Paper, Brandeis University.

Schuett, Florian. 2013. "Inventors and impostors: an analysis of patent examination with self-selection of firms into R&D." *Journal of Industrial Economics* 61 (3): 660–699.

Thoma, Grid. 2013. "Quality and Value of Chinese Patenting: An International Perspective." *Seoul Journal of Economics* 26 (1): 33–72.

Webster, Elizabeth, Paul H Jensen, and Alfons Palangkaraya. 2014. "Patent examination outcomes and the national treatment principle." *The RAND Journal of Economics* 45 (2): 449–469.

WIPO. 2017. *World Intellectual Property Indicators 2017.* WIPO publication. World Intellectual Property Organization. ISBN: 978-92-805-2903-6.

Yang, Deli. 2008. "Pendency and grant ratios of invention patents: A comparative study of the US and China." *Research Policy* 37 (6-7): 1035–1046.

Table 1: Compounded annual growth rate of patents applied for in countries between 2006 and 2016 (Figures in percentages)

Region	Country	Country-wise growth	Region-wise growth	Asian growth without China and Japan
Asia	China	20.32	8.50	2.51
	Hong Kong	0.22		
	India	4.53		
	Iran	9.13		
	Japan	-2.47		
	South Korea	2.31		
	Singapore	1.83		
	Russia	0.99		
Oceania	Australia	0.88	1.89	
	Indonesia	6.35		
North America	Canada	-1.89	3.18	
	United States of America	3.58		
Latin American Countries	Brazil	3.51	2.54	
	Mexico	1.17		
Europe	France	-0.61	0.98	
	Germany	1.15		
	Italy	-1.04		
	United Kingdom	-1.53		
	European Patent Office	1.66		
Africa	South Africa	2.47	2.47	

Source: WIPO Statistics Database and Authors' calculation

Table 2: Compounded annual decadal growth rate in high technology exports (Figures in percentages)

Country/Year	1996-2006	2006-2016
China	32.96	6.15
Hong Kong	-9.70	-13.46
India	11.36	10.58
Iran	76.14	5.68
Japan	2.41	-3.25
South Korea	12.93	2.40
Singapore	7.96	0.13
Russia	5.66	5.56
Australia	5.81	3.09
Indonesia	10.17	-4.16
Canada	3.06	-1.08
United States of America	4.72	-3.51
Brazil	18.25	1.51
Mexico	11.76	2.69
France	6.70	2.45
Germany	10.35	1.52
Italy	3.87	0.75
United Kingdom	7.96	-5.43
South Africa	9.45	0.63

OECD classifies exports into four categories, high, medium-high, medium-low and low technology based on expenditure on R & D relative to gross output and value added. Please note that for Iran, data on high-tech exports were only available for the years 1997 to 2006, which is reported in the column 1996-2006 and 2010 to 2011, which is reported in the column 2006-16.
Source: World Bank Databank and Authors' calculation.

Table 3A: Patents applied and granted

		Domestic	Foreign
China	Applied	3,707,860	798,954
	Granted	1,314,473	633,396
India	Applied	72,198	287,053
	Granted	4,981	36,855

Table 3B: Descriptive statistics (granted patents)

		Domestic	Foreign	Diff	t-stat
Grant lag	China	960.12	1,652.00	691.88	926.24***
	India	2,231.22	2,057.78	-173.44	-14.63***
	Diff	1,271.10	405.78		
	t-stat	112.91***	107.16***		
Nb indep claims	China	1.40	2.80	1.41	490.19***
	India	1.93	2.04	0.11	5.05***
	Diff	0.53	-0.76		
	t-stat	26.79***	-74.18***		
Tot wordcount ic	China	228.18	336.90	108.72	278.04***
	India	203.84	312.79	108.94	38.47***
	Diff	-24.33	-24.11		
	t-stat	-9.98***	-16.17***		
Inv count	China	3.57	2.83	-0.73	-230.00***
	India	2.78	3.11	0.33	11.81***
	Diff	-0.78	0.28		
	t-stat	-30.36***	25.36***		
Dep per ind claim	China	4.83	6.29	1.46	219.57***
	India	5.91	10.18	4.27	55.16***
	Diff	1.09	3.89		
	t-stat	15.63***	111.60***		
Wordcount per c	China	102.95	63.93	-39.02	-440.00***
	India	53.73	57.39	3.66	6.25***
	Diff	-49.22	-6.55		
	t-stat	-91.87***	-25.99***		
Family size	China	2.39	12.89	10.50	971.39***
	India	1.08	1.03	-0.05	-2.20**
	Diff	-1.30	-11.86		
	t-stat	-58.22***	-1000.00***		
RTA	China	1.03	0.92	-0.11	-360.00***
	India	1.10	0.96	-0.14	-24.59***
	Diff	0.07	0.04		
	t-stat	13.15***	17.42***		
Pharmaceuticals	China	933.74	2,048.52	1,114.77	229.58***
	India	2,131.86	1,921.22	-210.64	-5.83***
	Diff	1,198.12	-127.29		
	t-stat	35.87***	-8.75***		
Technology	China	1,199.60	1,704.31	504.70	288.25***
	India	2,623.13	2,089.83	-533.30	-15.11***
	Diff	1,423.52	385.52		
	t-stat	41.63***	43.28***		
Others	China	913.93	1,617.61	703.68	866.30***
	India	2,191.91	2,060.99	-130.92	-9.89***
	Diff	1,277.98	443.38		
	t-stat	112.91***	107.16***		
Nb granted per year	China	101.14	44.08	-57.06	-1200.00***
	India	0.21	2.71	2.50	320.71***
	Diff	-100.93	-41.37		
	t-stat	-2500***	-1600***		

Column labelled difference reports difference in means between foreign and domestic applicants for each variable for only granted patents. Welch's t-statistic assuming unequal variance has been calculated and reported under the rows and column t-stat and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 3C: Descriptive statistics (all applications)

		Domestic	Foreign	Diff	t-stat
Nb indep claims	China	1.34	2.85	1.51	797.17***
	India	2.29	2.51	0.22	29.02***
	Diff	0.95	-0.34		
	t-stat	144.55***	-83.73***		
Tot wordcount ic	China	199.79	331.31	131.51	540.48***
	India	226.90	331.88	104.98	115.74***
	Diff	27.11	0.57		
	t-stat	35.54***	1.04		
Inv count	China	3.01	2.87	-0.14	-70.30***
	India	2.73	3.04	0.31	40.84***
	Diff	-0.27	0.17		
	t-stat	-40.83***	43.05***		
Dep per ind claim	China	4.28	6.31	2.03	479.25***
	India	5.36	8.69	3.33	172.69***
	Diff	1.08	2.39		
	t-stat	66.69***	212.47***		
Wordcount per c	China	94.85	61.89	-32.96	-710.00***
	India	54.12	59.34	5.22	34.36***
	Diff	-40.73	-2.55		
	t-stat	-300.00***	-30.44***		
Family size	China	1.58	11.05	9.46	1300.00***
	India	1.08	1.88	0.80	99.82***
	Diff	-0.50	-9.16		
	t-stat	-140.00***	-910.00***		
RTA	China	1.02	0.94	-0.08	-500.00***
	India	1.12	0.97	-0.15	-100.00***
	Diff	0.10	0.03		
	t-stat	77.25***	55.43***		
Nb applied per year	China	100.64	44.18	-56.46	-1300.00***
	India	0.25	2.41	2.16	353.97***
		Domestic+Foreign			
Nb applied per year	China	679.71			
	India	34.99			
	Diff	-644.72			
	t-stat	-5100.00***			

Column labelled difference reports difference in means between foreign and domestic applicants for each variable for all patents in our data set. Welch's t-statistic assuming unequal variance has been calculated and reported under the rows and column t-stat and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 4: Linear model using grant lag as dependent variable

Variable	(1)	(2)	(3)	(4)
ID	236.9093	242.7697	222.3132	402.0070
	(5.45)***	(5.68)***	(5.24)***	(5.70)***
CF	-633.2652	-651.5889	-648.2561	-826.2145
	(-13.62)***	(-13.96)***	(-14.03)***	(-17.36)***
CD	-932.0725	-944.0615	-953.2871	-864.6751
	(-19.32)***	(-19.40)***	(-19.79)***	(-16.72)***
Nb indep claims	57.6660	53.2102	52.5957	51.3542
	(51.28)***	(49.25)***	(49.21)***	(48.70)***
Tot wordcount ic	-0.1110	-0.1333	-0.1265	-0.1250
	(-17.73)***	(-23.90)***	(-25.25)***	(-25.43)***
Inv count	-8.7157	-7.2089	-7.6366	-7.4085
	(-19.82)***	(-18.86)***	(-20.06)***	(-19.75)***
Dep per ind claim	12.4596	11.7602	11.9176	11.4133
	(47.74)***	(43.57)***	(42.92)***	(42.19)***
Wordcount per c	-0.1262	-0.1570	-0.1667	-0.1708
	(-6.97)***	(-8.84)***	(-9.35)***	(-9.63)***
Family size	21.9535	22.6327	22.3859	21.6113
	(45.44)***	(46.87)***	(45.86)***	(46.42)***
Technology		163.4743	190.7347	175.5265
		(11.60)***	(18.15)***	(16.20)***
Others		26.5660	50.1034	38.1377
		(2.77)***	(7.61)***	(5.67)***
RTA			85.8109	47.7932
			(7.04)***	(2.59)**
RTA*ID				-144.2607
				(-3.46)***
RTA*CF				204.3465
				(10.28)***
RTA*CD				-80.2237
				(-3.52)***
Publication year FE	Yes	Yes	Yes	Yes
Cons	2,116.5988	2,096.4707	1,995.2739	2,056.3523
	(43.60)***	(41.39)***	(40.11)***	(39.87)***
R2/Pseudo R2	0.504	0.512	0.513	0.516
F-value/Wald-stat	1180	1127	1163	1034
N	1,979,798	1,979,798	1,979,326	1,979,326
Chi2 ID vs CF	491.51***	554.30***	542.87***	433.50***
Chi2 ID vs CD	774.39***	826.06***	824.63***	335.72***
Chi2 CD vs CF	773.43***	740.28***	813.42***	6.49**

Column 1 of the table shows results from grant lag regressions which include the key variables, domestic applications in India (ID), domestic applications in China (CD), and foreign applications in China (CF). The group foreign applications in India (IF) is the reference group. Grant lag is counted as the difference between the application date and the grant date. Column 2 includes control variables and two domain groups, Pharmaceuticals is the reference category. Column 3 and 4 includes RTA and RTA interacted with ID, CF and CD respectively. The regression also controls for publication year fixed effects. Standard errors are clustered on grant months, z-statistics are shown within the parentheses and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 5: Hazard rates for granted applications

Panel A			
Statistics	Weillbull	Exponential	Cox
AIC	8289152.7	8980547	58372948
BIC	8289303.9	8980685	58373072

Panel B			
	Hazard rate	Hazard rate	Hazard rate
ID	-0.3126	-0.3025	-0.4181
	(-7.38)***	(-7.27)***	(-6.61)***
CF	1.3097	1.355	1.6
	(27.20)***	(26.82)***	(32.53)***
CD	1.9743	2.0549	2.0668
	(35.00)***	(35.28)***	(30.18)***
Nb indep claims	-0.1557	-0.1349	-0.1344
	(-30.23)***	(-33.16)***	(-32.89)***
Tot wordcount ic	0.0003	0.0003	0.0003
	(16.22)***	(19.86)***	(19.87)***
Inv count	0.0471	0.0417	0.0417
	(43.79)***	(38.96)***	(38.57)***
Dep per ind claim	-0.0278	-0.0266	-0.0263
	(-49.85)***	(-49.08)***	(-49.08)***
Wordcount per c	0.0008	0.0009	0.0009
	(13.24)***	(16.64)***	(16.66)***
Family size	-0.0465	-0.0477	-0.0473
	(-37.37)***	(-34.29)***	(-34.54)***
Technology		-0.4067	-0.3971
		(-15.53)***	(-15.29)***
Others		0.0504	0.0586
		(2.41)**	(2.84)***
RTA		-0.167	-0.069
		(-5.68)***	(-3.47)***
RTA*ID			0.0951
			(2.12)**
RTA*CF			-0.2659
			(-9.03)***
RTA*CD			-0.0178
			(-0.35)
Constant	-22.215	-22.4074	-22.5118
	(-118.78)***	(-127.61)***	(-127.83)***
ln_p	1.0705	1.0846	1.0846
	132.38***	138.31***	138.45***
Wald test	28425.17	31116.31	40855.29
N	1989703	1989095	1989095

Panel A of the table shows the AIC and BIC defined as $AIC := -2log(L)+2(p)$ and $-2ln(L)+p\times log(n)$ where p is the number of model parameters, L is the model likelihood function and n is the number of observations.

Panel B of the table shows results from the main analyses (column 1), the analyses with domains (column 2) and analyses with domains and RTA (column 3) that only use granted patent observations. All the models include the key variables, domestic applications in India (ID), domestic applications in China (CD), and foreign applications in China (CF). The group foreign applications in India (IF) is the reference group. Column 1 includes the control variables, column 2 adds two domain groups to the model, Pharmaceuticals is the reference category and column 3 adds RTA and RTA interacted with ID, CF and CD respectively. Standard errors are clustered on grant months, z-statistics are shown within the parentheses and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 6: Logit model using probability of grant as dependent variable

Variable	(1)	(2)	(3)	(4)
ID	-0.0618	-0.0601	-0.0615	-0.0528
	(-2.44)**	(-2.36)**	(-2.43)**	(-1.59)
CF	0.2455	0.2434	0.2464	0.439
	(6.77)***	(6.67)***	(6.74)***	(11.80)***
CD	0.2325	0.2325	0.2333	0.0065
	(5.82)***	(5.79)***	(5.79)***	(0.13)
Nb indep claims	-0.0044	-0.0036	-0.0037	-0.0022
	(-3.25)***	(-3.31)***	(-3.36)***	(-2.03)**
Tot wordcount ic	0.0001	0.0001	0.0001	0.0001
	(5.08)***	(4.74)***	(4.88)***	(4.76)***
Inv count	0.0149	0.0148	0.0148	0.0148
	(10.10)***	(10.30)***	(10.26)***	(10.34)***
Dep per ind claim	0.0071	0.0071	0.0072	0.0078
	(7.12)***	(7.00)***	(7.08)***	(7.89)***
Wordcount per c	0.0003	0.0003	0.0003	0.0003
	(7.99)***	(7.90)***	(7.76)***	(7.97)***
Family size	0.015	0.0152	0.0152	0.0158
	(12.00)***	(12.10)***	(12.02)***	(12.03)***
Technology		0.0328	0.0388	0.0472
		(4.07)***	(5.01)***	(5.99)***
Others		0.0394	0.0436	0.0475
		(7.19)***	(8.28)***	(9.31)***
RTA			0.0294	-0.0198
			(3.28)***	(-3.25)***
RTA*ID				0.0024
				(0.15)
RTA*CF				-0.2139
				(-15.63)***
RTA*CD				0.2288
				(14.05)***
R2/Pseudo R2	0.064	0.064	0.064	0.068
Wald test	942.49	1,566.36	2,209.68	6,229.70
N	6,855,770	6,855,770	6,850,990	6,850,990

Column 1 shows the marginal effects from the logit regression in which key variables together with the control variables are included. Column 2 includes technology domain group dummy. Column 3 includes RTA and RTA interacted with *ID*, *CF* and *CD* respectively Standard errors are clustered on grant months, z-statistics are shown within the parentheses and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 7: Hazard rate of grant

	(1)	(2)	(3)	(4)
ID	-0.4505	-0.1711	-0.1271	-0.1123
	(-7.40)***	(-2.83)***	(-2.12)**	(-1.04)
CF	1.4677	1.5216	1.517	1.887
	(17.03)***	(16.20)***	(16.13)***	(19.61)***
CD	1.6024	1.8248	1.8458	1.8707
	(15.94)***	(16.85)***	(17.14)***	(16.85)***
Nb indep claims		-0.0503	-0.0489	-0.0475
		(-15.46)***	(-14.87)***	(-14.42)***
Tot wordcount ic		0.0005	0.0005	0.0005
		(21.97)***	(20.66)***	(20.34)***
Inv count		0.0916	0.0917	0.0917
		(38.08)***	(38.03)***	(38.06)***
Dep per ind claim		0.0347	0.0343	0.035
		(16.72)***	(16.74)***	(17.17)***
Wordcount per c		0.0022	0.0023	0.0023
		(45.41)***	(46.54)***	(46.72)***
Family size		0.0115	0.0119	0.0123
		(12.88)***	(13.19)***	(13.61)***
Technology		0.286	0.2032	0.2038
		(8.69)***	(6.28)***	(6.27)***
Others		0.5045	0.4464	0.4447
		(24.14)***	(21.00)***	(20.76)***
RTA			-0.2732	-0.1132
			(-18.29)***	(-4.65)***
RTA*ID				-0.0325
				(-0.45)
RTA*CF				-0.3996
				(-13.53)***
RTA*CD				-0.0323
				(-1.00)
Constant	-14.1713	-15.9172	-15.6206	-15.7708
	(-50.40)***	(-51.34)***	(-49.99)***	(-51.40)***
ln_p	0.4571	0.4916	0.4934	0.4926
	(24.09)***	(25.31)***	(25.40)***	(25.40)***
Wald test	442.50	11702.34	11456.45	15941.78
N	6,852,983	6,852,983	6,848,203	6,848,203
Chi2 ID vs CF	421.71***	292.47***	280.14***	290.31***
Chi2 ID vs CD	343.07***	305.93***	305.73***	216.29***
Chi2 CD vs CF	16.07***	107.01***	127.81***	0.14

Table shows the results from hazard rate of grant. The main variables of interest is domestic applications in India (ID), domestic applications in China (CD), foreign applications in China (CF) and their interactions with RTA. The group foreign applications in India (IF) is the reference group. Column 1 shows only the results from the main variables. Column 2 includes the control variables in the model. Column 3 includes RTA and RTA interacted with ID, CF and CD respectively. Standard errors are clustered on grant months, z-statistics are shown within the parentheses and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 8: Results from domain groups

	All	Pharmaceuticals	Technology	Others
ID	-0.1711	-0.4106	-0.6866	-0.0323
	(-2.83)***	(-6.01)***	(-4.84)***	(-0.51)
CF	1.5216	0.1808	1.4118	1.6032
	(16.20)***	(2.12)**	(15.01)***	(16.66)***
CD	1.8248	1.6399	1.6359	1.915
	(16.85)***	(18.35)***	(16.43)***	(16.51)***
Nb indep claims	-0.0503	0.015	-0.0398	-0.0648
	(-15.46)***	(2.87)***	(-10.95)***	(-16.04)***
Tot wordcount ic	0.0005	-0.0004	0.0005	0.0006
	(21.97)***	(-12.21)***	(20.61)***	(23.51)***
Inv count	0.0916	0.0892	0.061	0.1005
	(38.08)***	(26.16)***	(20.53)***	(40.24)***
Dep per ind claim	0.0347	0.0236	0.0267	0.0399
	(16.72)***	(10.27)***	(19.58)***	(17.23)***
Wordcount per c	0.0022	0.0019	0.0024	0.0022
	(45.41)***	(11.72)***	(32.16)***	(42.97)***
Family size	0.0115	0.0299	0.0149	0.0128
	(12.88)***	(20.08)***	(14.34)***	(15.70)***
Technology	0.286			
	(8.69)***			
Others	0.5045			
	(24.14)***			
Cons	-15.9172	-13.1517	-16.6165	-15.5064
	(-51.34)***	(-50.33)***	(-45.51)***	(-48.21)***
ln_p	0.4916	0.2977	0.5834	0.4867
	(25.31)***	(15.01)***	(26.39)***	(24.82)***
Wald test	11702.34	2257.19	5269.53	9250.14
N	6,852,983	373,000	1,239,450	5,240,533

Table shows the results from hazard rate of grant including domains. The main variables of interest is domestic applications in India (ID), domestic applications in China (CD), and foreign applications in China (CF). The group foreign applications in India (IF) is the reference group. In the analysis, 3 domain groups are formed, Pharmaceuticals, Technology, and Others. In column ALL, two domain groups are included and Pharmaceuticals is the reference category whereas in other columns each domain group is examined separately. Standard errors are clustered on grant months, z-statistics are shown within the parentheses and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 9: Results from the frailty model

	(1)	(2)	(3)	(4)
ID	-1.0179	-1.0504	-1.0001	-0.9423
	(-5.64)***	(-5.88)***	(-5.74)***	(-2.94)***
CF	4.1471	4.3636	4.3784	5.9586
	(17.54)***	(17.67)***	(17.56)***	(20.60)***
CD	6.7326	6.8153	6.841	4.7338
	(23.54)***	(23.43)***	(23.70)***	(14.90)***
Nb indep claims	-0.338	-0.3126	-0.3125	-0.3033
	(-26.10)***	(-25.57)***	(-25.76)***	(-25.77)***
Tot wordcount ic	0.0015	0.0017	0.0017	0.0017
	(21.59)***	(28.45)***	(27.97)***	(27.57)***
Inv count	0.1963	0.1874	0.1876	0.1869
	(38.65)***	(36.75)***	(37.24)***	(36.99)***
Dep per ind claim	0.0214	0.0287	0.0285	0.0332
	(3.08)***	(3.90)***	(3.74)***	(4.35)***
Wordcount per c	0.0046	0.0048	0.0049	0.005
	(25.73)***	(26.73)***	(26.85)***	(27.87)***
Family size	-0.0321	-0.0393	-0.0392	-0.0333
	(-7.43)***	(-9.49)***	(-9.89)***	(-8.49)***
Technology		-1.47	-1.4806	-1.3724
		(-9.38)***	(-10.37)***	(-9.48)***
Others		-0.2944	-0.2973	-0.2065
		(-2.19)**	(-2.28)**	(-1.54)
RTA			-0.0885	-0.5297
			(-0.71)	(-5.59)***
RTA*ID				0.0101
				(0.05)
RTA*CF				-1.771
				(-14.66)***
RTA*CD				2.1637
				(10.13)***
Constant	-53.8782	-54.2219	-54.1557	-54.4349
	(-55.40)***	(-61.46)***	(-63.76)***	(-62.15)***
ln_p	1.9154	1.9282	1.9283	1.939
	(90.86)***	(102.89)***	(103.12)***	(106.38)***
lntheta	2.2862	2.2895	2.2891	2.2982
	(66.23)***	(72.21)***	(72.56)***	(76.46)***
Wald test	5066.67	5742.67	5738.63	7272.04
N	6,852,983	6,852,983	6,848,203	6,848,203

The table shows results from the main analyses (column 1), the analyses with domains (column 2) and the analyses with RTA and RTA interacted with ID, CF and CD (column 3 and 4) in a frailty model which assumes gamma distributed frailties. All the models include the key variables, domestic applications in India (ID), domestic applications in China (CD), and foreign applications in China (CF). The group foreign applications in India (IF) is the reference group. Column 1 includes the control variables and Column 2 adds two domain groups to the model, Pharmaceuticals is the reference category. Standard errors are clustered on grant months, z-statistics are shown within the parentheses and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 10: Results from Cox proportional hazard model

	(1)	(2)	(3)
ID	-0.162	-0.1257	-0.0979
	(-2.11)**	(-1.64)	(-0.82)
CF	1.5242	1.5228	1.8872
	(15.30)***	(15.23)***	(18.02)***
CD	1.8309	1.845	1.6805
	(16.65)***	(16.88)***	(16.21)***
Nb indep claims	-0.0435	-0.0427	-0.0409
	(-13.50)***	(-13.07)***	(-12.59)***
Tot wordcount ic	0.0005	0.0005	0.0005
	(23.14)***	(21.79)***	(21.47)***
Inv count	0.0873	0.0874	0.0873
	(35.62)***	(35.68)***	(35.69)***
Dep per ind claim	0.0365	0.0363	0.0372
	(17.28)***	(17.19)***	(17.56)***
Wordcount per c	0.0023	0.0023	0.0023
	(44.15)***	(44.67)***	(44.75)***
Family size	0.0102	0.0104	0.0109
	(14.91)***	(15.06)***	(15.96)***
Technology	0.1638	0.1161	0.1209
	(4.82)***	(3.71)***	(3.85)***
Others	0.3684	0.3352	0.3348
	(16.97)***	(15.69)***	(15.44)***
RTA		-0.1785	-0.1282
		(-10.53)***	(-5.35)***
RTA*ID			-0.0283
			(-0.39)
RTA*CF			-0.3957
			(-14.86)***
RTA*CD			0.1634
			(4.35)***
Wald test	10460.8	10362.53	14082.84
N	6,852,983	6,848,203	6,848,203

The table reports results from Cox proportional-hazards model. This table reports the analyses with domains (column 1) and the analyses with RTA and RTA interacted with ID, CF and CD (column 2 and 3). All the models include the key variables, domestic applications in India (ID), domestic applications in China (CD), and foreign applications in China (CF). The group foreign applications in India (IF) is the reference group. Column 1 includes the control variables and the technology domains, Pharmaceuticals is the reference category. Standard errors are clustered on grant months, z-statistics are shown within the parentheses and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 11: Results from Harhoff and Wagner 2009 model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ID	-0.4505	-0.2371	-0.5902	-0.5454	-0.5511	-0.6862	-0.8831
	(-7.40)***	(-3.88)***	(-7.35)***	(-6.79)***	(-6.98)***	(-6.73)***	(-11.20)***
CF	1.4677	1.459	1.0997	1.0347	1.0312	1.1635	-0.0404
	(17.03)***	(15.88)***	(13.28)***	(13.33)***	(13.39)***	(10.90)***	(-0.62)
CD	1.6024	1.5369	1.6016	1.5365	1.5218	1.6982	0.5947
	(15.94)***	(14.54)***	(16.56)***	(16.45)***	(16.38)***	(12.55)***	(7.83)***
Controls		As table 7	Set 1	Set 2	Set 3	Set 3	Set 3
Pharmaceutical		-0.4329	-0.2052	-0.2278	-0.2265	-0.3894	0.0626
		(-18.88)***	(-5.41)***	(-6.77)***	(-6.83)***	(-10.84)***	(2.89)***
Technology		-0.0625	-0.0254	-0.0389	-0.0394	-0.1664	-0.0013
		(-3.43)***	(-1.22)	(-1.92)*	(-1.96)**	(-7.55)***	(-0.05)
Biotechnology		-0.2052	-0.3172	-0.2994	-0.2966	-0.3296	-0.242
		(-7.57)***	(-19.64)***	(-21.54)***	(-21.02)***	(-20.16)***	(-21.50)***
Cons	-14.1713	-15.1437	-19.7026	-19.6686	-19.6338	-19.4113	-29.1146
	(-50.40)***	(-49.93)***	(-43.61)***	(-44.22)***	(-44.33)***	(-42.21)***	(-68.45)***
ln_p	0.4571	0.5226	0.8649	0.862	0.8618	0.817	1.3393
	(24.09)***	(27.83)***	(38.07)***	(38.96)***	(38.95)***	(35.85)***	(97.03)***
Wald test	442.50	37749.62	18950.29	20,066	20927.12	13420.12	32619.39
N	6,852,983	6,852,983	1,731,572	1,284,743	1,284,743	1,274,206	553,173
Chi2 ID vs CF	421.71***	310.66***	306.62***	331.42***	328.54***	437.89***	114.57***
Chi2 ID vs CD	343.07***	256.97***	535.78***	525.51***	516.70***	522.20***	549.83***
Chi2 CD vs CF	16.07***	8.64***	107.92***	138.53***	135.28***	96.49***	155.81***

The table reports results using the model from Harhoff and Wagner 2009. Column 1 reports analysis with ID, CF and CD. Column 2 reports results with technology domains and control variables as used in column 2 of Table 7 (*nb indep claims, tot wordcount ic, inv count, dep per ind claim, wordcount per c, family size*; these variables are defined in Table A1). For the next five columns, we add additional control variables. The additional common control variables for the five columns are *number of cited patents, number of cited non-patents, number of domestic equivalents, number of equivalents, number of patents applied, IPC section, PCT application and group dummies.* For column 3 we use controls used in column 2 of 7 and the common controls. For column 4, we use controls used in column 2 of 7, *citations received within 3 years* and the common controls. For column 5 we use controls used in column 2 of 7, *number of citations* and the common controls. For column 6 we replicate column 5 but we use the values of control variables from the version of patent document which was used while applying for in a patent office. For column 7 we replicate column 5 but we use the values of control variables from the version of patent document which was granted from the patent office. A patent document if granted is republished and often has different set of claims than those in its application version. Column 6 and 7 also serves as robustness checks to ensure that there is no systematic difference between a patent application and a granted patent which affects our results. Standard errors are clustered on grant months, z-statistics are shown within the parentheses and marked with stars * for 10%, ** for 5%, and *** 1% significance levels.

Table 12: Results from patent value estimates

	(1)	(2)
CD		1.0941
		(26.73)***
maintain_10yr_hat	6.1911	4.5423
	(15.65)***	(13.60)***
Pharmaceutical	-0.6787	-0.6225
	(-20.33)***	(-17.69)***
Technology	-0.4664	-0.414
	(-17.93)***	(-16.48)***
Biotechnology	-0.3917	-0.35
	(-18.82)***	(-14.44)***
Cons	-20.8129	-21.2261
	(-43.27)***	(-43.73)***
ln_p	0.6538	0.7579
	(45.67)***	(44.92)***
Wald test	1419.33	1671.1
N	1284743	1284743

The table shows results from the analysis that replicates the patent value estimates in Harhoff and Wagner 2009. In a probit model, we first estimate whether a patent was maintained for more than or equal to 10 years by using controls variables used in column 5 Table 11. We then use these estimated probabilities together with CD and other control variables (used in column 5 of Table 11) to examine their relation to the hazard rate of grant. Since data on maintaining patent for 10 years is available only for China, results for India are unavailable in this table.

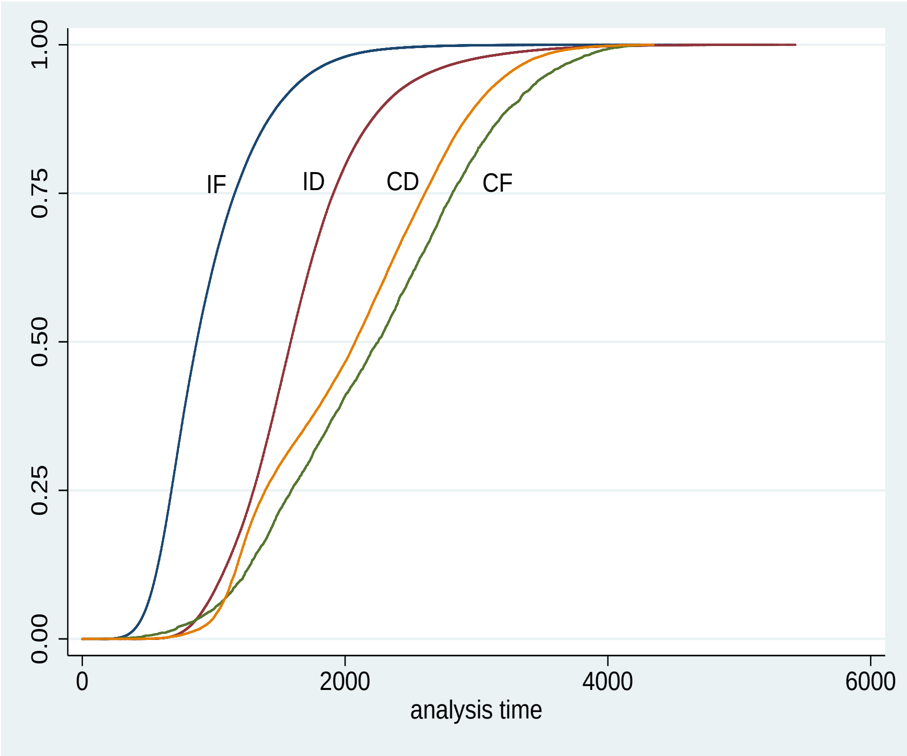

Figure 1: Kaplan-Meier failure function by groups

IF, ID, CF, CD are dummy variables for patents applied for in India by foreign applicants; patents applied for in India by domestic applicants; patents applied for in China by foreign applicants and patents applied for in China by domestic applicants respectively.

Appendices

Table A1: Definition of the variables

Panel A	
Variable	Definition
Grant lag:	number of days taken to grant
ID:	1 if applied for by a domestic entity in India; 0 otherwise
IF:	1 if applied for by a foreign entity in India; 0 otherwise
CD:	1 if applied for by a domestic entity in China; 0 otherwise
CF:	1 if applied for by a foreign entity in China; 0 otherwise
Nb indep claims:	Number of independent claims
Tot wordcount ic:	Total number of words for all independent claims
Inv count:	Total number of inventors involved
Dep per ind claim:	Ratio of dependent claims to independent claims
Wordcount per c:	Number of words per claim
Family size:	Number of patents sharing the same priority with patent i
RTA:	Revealed technological advantage for technology i for country j for year t is defined as the proportion of domestic patents in technology domain i divided by the proportion of all patents applied for in that country in technology domain i; $RTA_{ijt} = \frac{\frac{P_{ijt\|domestic=1}}{\sum_{i=1}^{35} P_{ijt\|domestic=1}}}{\frac{P_{ijt}}{\sum_{i=1}^{35} P_{ijt}}}$
Panel B	Definition of the control variables used in Harhoff and Wagner 2009 model
Citations received within 3 years:	Number of citations received within three years of application
Number of citations:	Number of citations received till 31st December 2016
Number of cited patents:	Number of patents cited. Number of patents cited by a patent application is also referred to as forward citations
Number of cited non-patents:	Number of non-patents cited. Some examples of non-patents are academic papers, technical reports, etc.
Number of domestic equivalents:	Number of times patent i or an equivalent patent to patent i has been applied for in India (China) for patents domestic to India (China). The equivalents may be a previous version or an upgrade of a patent. The equivalent patents are listed in priority information for any patent
Number of equivalents:	Number of times patent i or an equivalent patent has been applied for in any patent office. Number of equivalent patents may also be called a subset of patent family. The correlation coefficient between number of equivalents and patent family size is 0.53 and hence both may be used in one regression
Number of patents applied:	Total number of patents applied for each year in India (China) per thousand applications
IPC section:	Single digit IPC section (A, B, C, D, E, F, G or H)
PCT application:	1 if applied for through PCT route; 0 otherwise
Group:	Dummies for patents categorised as Pharmaceutical, Technology, Biotechnology and Others have been used. In Table 11, group "Others" acts as base. Our main set of results use only Pharmaceutical, Technology and Others dummies