WP/18/269

IMF Working Paper

International Knowledge Spillovers

by Johannes L. Eugster, Giang Ho, Florence Jaumotte, and Roberto Piazza

INTERNATIONAL MONETARY FUND

 WP/18/269

IMF Working Paper

Research Department (RES)

International Knowledge Spillovers

Prepared by Johannes L. Eugster, Giang Ho, Florence Jaumotte, and Roberto Piazza

Authorized for distribution by Helge Berger

December 2018

Abstract

How important is foreign knowledge for domestic innovation outcomes? How is this relation shaped by globalization and the attendant intensification of international competition? Our empirical approach extends the previous literature by analyzing a large panel comprising industries in both advanced and emerging economies over the past two decades. We find that barriers to the domestic diffusion of foreign knowledge have fallen significantly for emerging economies. For all countries, and especially for emerging economies, inflows of foreign knowledge have a growing and quantitatively important impact on domestic innovation. Controlling for the amount of domestic R&D, we find evidence that increases in international competitive pressure at the industry level had a positive effect on domestic innovation outcomes.

JEL Classification Numbers: F1, F2, O3, O4

Keywords: knowledge spillovers, knowledge barriers, international competition.

Authors' E-Mail Address: JEugster@imf.org; FJaumotte@imf.org; RPiazza2@imf.org

1 Introduction[1]

Technology is the key driver of improvements in income and living standards. Since the industrial revolution, the bulk of global technological innovation has occurred in just a few countries. This is still the case nowadays. For instance, about 75 per cent of international patent families registered between 1995 and 2014 were attributed to an inventor located in one of the G5 countries, defined as France, Germany, Japan, United Kingdom and the United States.[2] In the more recent years, however, the share of patents by inventors in China and Korea has grown significantly, and in 2014 represented 16 per cent of total international patents.

The geographical concentration of technological innovation has important consequences. To increase global living standards, inventions embodied into new goods, or in the form of codified or "tacit" knowledge, must cross boundaries and diffuse to the rest of the world. For most countries, if not all, domestic technology is thus dependent on knowledge developed abroad. This relation is determined by a host of structural factors, such as the intensity of competitive forces, which shape the degree of adoption and the impact of foreign knowledge on the domestic economy.

This paper presents an attempt to quantify the strength of this nexus between domestic innovation and foreign knowledge, its evolution over time, and its potential determinants. More specifically, our empirical investigation estimates the importance of international barriers to knowledge diffusion and provides an indication of how their size has changed during the past decades. We also quantify the importance of foreign knowledge for domestic innovation outcomes and assess how the strength of this link was affected by the increased international competitive pressures that have accompanied the globalization process.

The empirical strategy is based on applying the two-step estimation in Peri (2005) to the PATSTAT database, which collects worldwide information on patents - our main proxy for innovation. In the first step, a "gravity equation" is used to assess the impact of different barriers between country-industry pairs on the international diffusion of innovations, proxied by patent citations, from the G5. A second-step regression then estimates the sensitivity of domestic innovation and productivity outcomes in non-G5 countries to the foreign knowledge

[1]This paper is dedicated to the memory of our friend and colleague Giang Ho. Her wit and smile will be sorely missed. We thank Helge Berger, Gian Maria Milesi-Ferretti and Maurice Obstfeld for insightful comments. Pankhuri Dutt, Chanpheng Fizzarotti and Menexenia Tsaroucha provided excellent research assistance.

[2]Most of the analysis is carried out on international patent families. For a definition, see Section 3.

that, overcoming international barriers, manages to flow to the domestic economy. In both steps, the sample of recipient countries comprises advanced and emerging economies, 11 manufacturing industries plus the construction and IT sectors, and spans from 1995 to 2014. The introduction of the industry dimension, the widening of the geographical scope of the analysis to emerging economies, and the focus on more recent decades represent significant extensions of Peri (2005) - all made possible by our novel use of PATSTAT. The data thus allows us to paint a picture of how knowledge flows from G5 countries, and their importance for recipient economies, have changed as globalization has progressed during the past two decades.

We find that international barriers to knowledge diffusion are sizable, but also that their effect has become less important over time, allowing for a growing integration of emerging economies into the international flow of knowledge. Consider the diffusion of an innovation originated in a G5 country to the same domestic sector in other countries. We estimate that the intensity of such international diffusion is only 15 percent the intensity of within-country diffusion. In other words, international barriers cause, on average, an 85 percent reduction in international knowledge diffusion. There is however important heterogeneity across time and space. For advanced economies, the relative intensity of diffusion remained statistically unchanged at about 20 percent over the entire sample period. For emerging economies, instead, the relative intensity of diffusion increased significantly from 10 percent in the period 1995-1999 to 16 percent in 2010-2014, so that by the end of the sample the difference between the levels of knowledge barriers in emerging and advanced economies had been significantly reduced. The estimates thus provide evidence that, over the past two decades, the world economy has undergone a process of "knowledge globalization", led by a fall of knowledge barriers towards emerging economies.

We find not only that countries are now "closer" to each other in terms of intensity of knowledge diffusion, but also that domestic technological outcomes have become more sensitive to knowledge flows from abroad. On average, controlling for country-time effects, knowledge originated in G5 economies was about 80 percent as effective in raising domestic sectoral patenting as domestically generated knowledge. Relative effects of roughly the same magnitude are found when sectoral patent flows are replaced with sectoral TFP but, due to data limitations, the sample for the TFP estimation is restricted to advanced economies. Once we break down our estimation by sub-periods and subgroups of countries, we find that the effect of foreign knowledge flows on domestic innovation has increased in a statistically significant way over time, with the increase being especially large for the sub-sample of emerging economies. These conclusions are robust to the use of alternative control variables that allow us to expand the sample of emerging economies or to the use of an alternative measure of patenting. We also confirm that the coefficients are qualitatively unchanged if they are estimated with dynamic OLS, with alternative combinations of fixed effects, and when the intensity of knowledge diffusion is proxied by the more traditional bilateral trade weights instead of the intensity of citation from the first-stage gravity model.

Motivated by these results, we proceed to explore whether changes in the strength of international competitive forces could explain the observed temporal variation in the sensitivity of domestic technological outcomes to foreign knowledge flows. The degree of product market competition is a key theoretical determinant of innovation activity, and its intensity has changed over time, shaped in part by the reduction in trade barriers that have accompanied globalization.

We construct two measures of international competitive pressure that are reasonably exogenous to developments in specific country-sectors. The first is obtained by computing the evolution of import penetration from China in U.S. industries and then using this variable to instrument import penetration in recipient advanced countries (similarly to Autor et al. (2014)). The second consists in constructing indexes of industry concentration at the global level and then excluding from our sample China which, being the largest non-G5 country, could introduce reverse causality between domestic innovation and our global concentration measure. We take each of these proxies and interact them, in the second-step regression, with the measure of foreign knowledge flow, while controlling for country-level developments with fixed effects. We find that, consistently across both measures, greater international competitive pressure increases both the level of sectoral patenting and its sensitivity to foreign knowledge flows. Although, theoretically, competition has ambiguous effects on innovation (Gilbert (2006), Akcigit et al. (2017)), our results point to a positive empirical relation in the international context. A potential explanation for our finding could be that greater international competition creates a selection effect, whereby firms that are relatively more effective in R&D activities - in particular in exploiting foreign knowledge - grow more (Aw et al. (2011)). Alternative interpretations of the positive effect of competition on innovation include the "escape competition" incentive (Aghion et al. (2005)), or the presence of "trapped" factors as in Bloom et al. (2013), where more competition reduces the opportunity cost of incumbent firms to divert internal resources towards innovative activities.

Our paper is related to two main strands of the literature. The first comprises the attempts to estimate international knowledge spillovers. Keller (2004) and Keller (2010) provide excellent reviews of this relatively large literature. For more recent contributions see also Coelli et al. (2016). As mentioned, our framework relies mainly on Peri (2005), but we also draw insights from the empirical specification in Coe et al. (2009). Our results confirm previous findings that identify the presence of important barriers to the international diffusion of knowledge, but the breadth of our sample and our focus on the evolution of the estimated parameters are novel and provide a significant extension of the literature. Our second contribution is the estimation of the impact of increased international competition on innovation activity. A small but growing number of papers has tried to empirically address this question. Autor et al. (2016) find that increasing competition from China has lowered innovation in U.S. industries. On the other hand, Bloom et al. (2016) find the opposite result for European firms. Our findings capture the conclusions of Bloom et al. (2016), as some European economies are included in our sample of recipient countries, but cannot be directly compared to Autor et al. (2016), since we consider the United States only as a source and not as a recipient economy. Moreover, differently from the aforementioned papers, our investigation does not focus on the impact of international competition on overall domestic R&D, but instead looks specifically at how competition changes the sensitivity of domestic innovation to foreign knowledge flows for given levels of R&D.

The rest of the paper proceeds as follows. Section 2 presents a simple conceptual framework that provides an interpretation of our empirical model. Section 3 provides an overview of data sources and of the construction of the variables employed in the regressions. Section 4 presents the results of the estimation of international knowledge flows, i.e. our first-step regression, while Section 5 discusses the estimated impact of foreign knowledge flows on domestic innovation outcomes, i.e. the second-step regression. Section 6 explores how international competition affected the importance of foreign knowledge for domestic innovation. Section 7 concludes.

2 Conceptual framework

The empirical analysis of this paper links knowledge stocks originated both domestically and abroad, to domestic innovation outcomes, such as patenting or productivity. A simple model can be used to formalize, under a common overarching framework, the empirical specifications typically studied in the previous literature. We also extend this basic framework to consider the effect of specific channels - international competition, in our case - on knowledge flows.

2.1 The basic framework

The derivation of the basic framework involves three steps. The first is to identify the stock of domestically generated knowledge for each country c at time t. Since knowledge stocks are not directly observable, a typical proxy employed in the literature is the cumulative discounted flow R_{ct} of domestic R&D up to time t.

The second step is to define the fraction ϕ_{cl} of knowledge generated in a foreign country l that is available for use in the domestic economy c, with the fraction $1 - \phi_{cl}$ interpreted as the size of the barriers that prevent knowledge from flowing from l to c (Eaton and Kortum (2002)). The total amount R^F_{ct} of foreign-generated knowledge available to country c can then be defined as,

$$R^F_{ct} = \sum_l \phi_{cl} R_{lt} \tag{1}$$

The third and final step is to define how the available knowledge is turned into innovation, either measured as the flow of patents P_{ct} or as productivity levels A_{ct}. First, domestic and available foreign knowledge are combined, at the country level, into an overall stock of usable knowledge K_{ct},

$$K_{ct} = (R_{ct})^{\beta} (R^F_{ct})^{\chi}$$

We then assume that

$$P_{ct} = X^P_{ct} (K_{ct})^{\theta}$$
$$A_{ct} = X^A_{ct} (K_{ct})^{\psi}$$

The variables X^P_{ct} and X^A_{ct} are aggregate factors that influence the process of turning knowledge into observable innovation measures. For instance, X^P_{ct} could be a country-time specific propensity to patent innovations, influenced by legal requirements and institutions. Similarly, standard growth models typically feature an equilibrium equation where $X^A_{ct} = A^{\rho}_{ct-1}$, with $\rho \in [0, 1]$.[3] By combining the equations above, it is straightforward to obtain two

[3] In this way, it is possible to nest several relations between knowledge and long-run productivity growth. For instance, if $\rho = 1$ and the R&D flow is calculated in terms of the number of employed research scientists, then exponential growth in productivity can be achieved even with a constant amount of knowledge. On the other hand, if $\rho = 0$ then exponential growth is possible only if the stock of knowledge grows exponentially which, by equation (1), would require exponential growth in at least one of the two R&D stocks. For an in-depth discussion, see Kortum (1997) and Bloom et al. (2017).

testable relations between innovation measures and R&D,

$$\log Y_{ct} = X_{ct}^{Y} + \gamma_Y \log R_{ct} + \mu_Y \log R_{ct}^{F} \tag{2}$$

where $Y_{ct} = \{P_{ct}, A_{ct}\}$ and γ_Y, μ_Y are combinations of the various elasticities introduced in the model. Equation (2) forms the basis of much of the empirical investigations of international technological spillovers in the literature, with the coefficient μ_Y indicating the importance of usable foreign knowledge for the determination of domestic innovation. For instance, Coe and Helpman (1995)) estimate equation (2) with $Y_{ct} = A_{ct}$ on a panel of countries by assuming that $X_{ct}^{A} = X_c$ is a country specific fixed effect and that the weights ϕ_{cl} are well captured by the intensity of bilateral trade links. Peri (2005), instead, estimates (2) with $Y_{ct} = P_{ct}$ on a dataset where the index c represents different regions in advanced countries, with the factors X_{ct}^{Y} assumed to be the same for all regions within a given country and at a given time (country-time fixed effects). He derived the weights ϕ_{cl} from a gravity model of the relative frequency of bilateral patent citations.

This paper broadens the analysis in Peri (2005) by augmenting the model with a sectoral dimension, by focusing on a more recent time period, and by letting the index c represent a wide set of non-G5 countries, both advanced or emerging. In parallel, source countries (i.e. countries l) are restricted to be in the G5 group. The new empirical specification allows us to study how R&D stocks R_{cit}^{F} in a foreign sector i influences the innovation outcome Y_{cit} in the corresponding domestic sector, after controlling for country-time effects X_{ct}^{Y},

$$R_{cit}^{F} = \sum_{l \in G5} \phi_{cli} R_{lit} \tag{3}$$

$$\log Y_{cit} = X_{ct}^{Y} + \gamma_Y \log R_{clit} + \mu_Y \log R_{cit}^{F} \tag{4}$$

Following Peri (2005), the weights ϕ_{clit} in (3) are obtained from a first-stage gravity model of bilateral patent citations regressed on exogenous variables including geographic, linguistic and initial technological differences. The weights are defined as the predicted citation frequencies from a gravity equation, excluding fixed effects (see Section 4). As a robustness check, we also use an alternative and simpler weighting scheme based on the traditional time-varying bilateral trade links between each country-industry pair.

The paper's main focus is on testing equation (4) with the dependent variable equal to the sectoral patent flows P_{cit}. However, we also recognize that the effect of foreign knowledge does not need to be limited to spurring patented innovations in the recipient country. Foreign technology can also affect productivity more directly, either through the purchase of imported equipment, licensing of technologies or other processes not involving a measurable economic transaction, including non-patented innovations. For these reasons, for a restricted group of countries for which data is available, we also tests (4) with the dependent variable equal to sectoral productivity A_{cit}.

The main benefit of the sectoral dimension is that it allows us to use country-time fixed effects to control for time-varying aggregate unobservable factors, including institutional characteristics that determine the propensity to patent or cite and which could be correlated with foreign knowledge flows. Moreover, adding the sector dimension increases significantly the number of observations for a given year, allowing us to obtain precise estimates of the parameters even when we break our sample into different sub-periods. The drawback is that macroeconomic variability, including general equilibrium effects, are also absorbed in these country-time fixed effects, which complicates the extent to which conclusions can be drawn about aggregate effects. Note also that the choice to restrict the source countries to the G5 group is motivated by the desire to reduce endogeneity concerns - as already pointed out, G5 countries represent, by and large, the technological frontier of the world economy. Finally, equation (4) could also be estimated in a panel where we the source sector is allowed to differ from the recipient sector. Here we focus on the empirical formulation where the two sectors coincide because this is the case where knowledge flows are found to be the largest.[4] Due to data availability, our analysis considers mostly the manufacturing industries (plus the construction and IT sectors), which are the ones where patenting activity is largely concentrated.

2.2 The effect of international competition

The degree of product market competition is among the key determinants of innovation. The entry of emerging market firms into the global stage can be thought, in particular, as a competition shock with potential repercussions on innovation activities and international knowledge flows. To delve deeper into this issue, we extend the empirical specification (4). We construct proxies Z_{cit} for the degree of international competition in sector i and in a non-G5 country c, and then add this proxy to our empirical model. We thus estimate the

[4] See Appendix A.2 for the inter-sectoral estimation of knowledge flows. Note that in our dataset sectors are already defined in broad way, and hence production activities are very different across sectors. For this reason, it should not be surprising that inter-sectoral knowledge flows are found to be small.

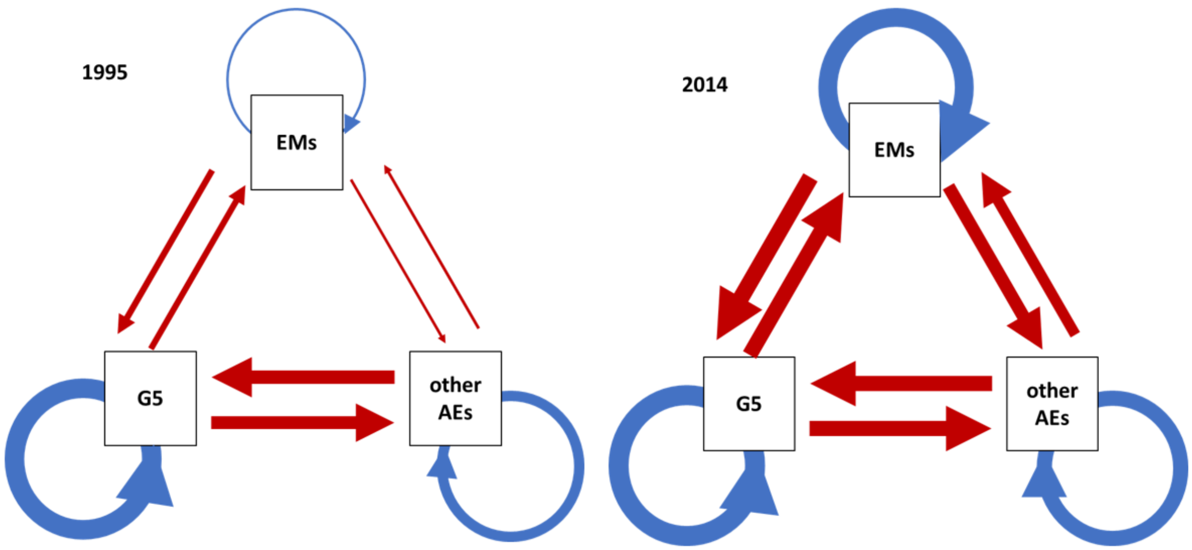


Figure 1: International knowledge flows proxied by patents citations. The width of the arrows is proportional to log of citations.

following equation for non-G5 countries:[5]

$$\log Y_{cit} = X^Y_{ct} + \theta Z_{cit} + \tilde{\gamma}_Y \log R_{cit} + \tilde{\mu}_Y \log R^F_{cit} + \delta Z_{cit} R^F_{cit} \tag{5}$$

To the extent that our proxies for competition are reasonably exogenous, the coefficient on the main effect (θ) captures the direct impact of international competition on innovation. The total impact of the weighted foreign knowledge stock on innovation is now given by $\tilde{\mu} + \delta Z$, and thus the coefficient on the interaction term (δ) reflects the marginal change to knowledge diffusion coming from a change in Z.

3 Data

This section introduces the key variables of the analysis, namely patent counts and cross-patent citations, which are all constructed from the PATSTAT database. Appendix A.1 provides a more detailed overview as well as descriptive statistics for key variables not shown here.

[5] See also Coe et al. (2009).

Data on patent applications are available internationally at disaggregated levels and allows us to precisely attribute an idea to its creator, to its time of creation, and to its areas of industrial use. Despite the precision of the information, patenting can be a noisy measure of innovation capacity as incentives to patent an innovation can differ across countries and over time, e.g. with changing procedures, requirements and fee structures of patent offices. To address this issue and improve comparability, the paper, in addition to using country-time fixed effects in all regressions, follows much of the literature in constructing quality-adjusted patent measures. We apply the concept of international patent family, which is a group of patents based on the same underlying innovation and featuring one application in at least two distinct patent offices. This tends to exclude patents whose lower expected payoffs does not warrant the extra cost for application, examination and maintenance in a foreign country.[6]

The patent count attributes a given patent family to the country of residence of the first inventor, the earliest publication year and the main industrial sector of applicability. Comprehensive citations to prior knowledge are a necessary component of a successful patent application. Patent citations included in new patent applications thus provide reliable and detailed information to trace cross-country and cross-sectoral linkages between innovation activities, which can be used, as in Peri (2005), to estimate the weights ϕ_{clit} in (3). The construction of the citation count largely follows the same principles as the one for patents.[7]

The raw data on bilateral citations, illustrated in Figure 1 by the thickness of the arrows and circles, provides an indication of the dramatic increase of cross-country patent citations, which over the past two decades accompanied the progressive integration of emerging markets into global knowledge flows, with China representing a key player among emerging markets.

Besides patent-related measures, our analysis relies also on sector-level R&D data. The stock of R&D is constructed using the perpetual inventory method and the business R&D spending in constant PPP USD provided by the OECDs ANBERD database. To test for a broader impact of foreign knowledge on total factor productivity, we complement the patent analysis with regressions using real value added as a dependent variable and controlling for

[6]Our use of international patent families as a quality-adjusted patent measure is different from Peri (2005), who weights the simple patent count by citation in the first 4 years after the patent was granted. While citation weights are a standard quality adjustment, they assume comparability of the underlying patenting and citation behavior. In our sample, which includes vastly different patenting cultures (e.g., Japan and China), a patent count based on the higher-level international patent families seems more appropriate, as it reduces the influence of country specificity.

[7]We attribute a patent to the publication year for the patent count, but to the application year for the citation count. Citations are counted only if they occur within four years of the publication year (see Table 8 in Appendix A.1 for discussion). Self-citations, which are defined as citations between patents with identical inventor, are excluded.

labor and capital inputs (as in Acharya and Keller (2009)). Data on industry value added, employment, and capital stock come from the 2017 EU KLEMS database, which shrinks our sample to mainly advanced economies.

4 Determinants of knowledge flows

This section describes the gravity model used to estimate the relative intensity of knowledge flows, i.e. the weights ϕ_{clit} in equation (3). Following Peri (2005), patent citations are modelled as an exponential function of a set of dummy variables that indicate whether citations involve two distinct countries ($diff.country$), which share a common border ($diff.border$) or an official language ($diff.lang$). In addition, the regression includes a measure of the geographical distance between the countries' capitals ($dist.int$), and differences in technological specialization ($tech.spec$) and development ($tech.dev$). While the latter technology variable captures the absolute difference in technological intensity (measured as the log-difference in R&D or value added per worker), the former captures compositional differences in the types of technology that is used.[8] The model can be written as follows:[9]

$$\begin{aligned}\phi_{cli} = \exp[a + f_{ci} + \tilde{f}_{li} + b_1(diff.country_{cl}) + b_2(diff.border_{cl}) + b_3(diff.lang_{cl}) \\ + b_4(dist.int_{cl}) + b_5(tech.controls_{cli}) + \epsilon_{cli}]\end{aligned}$$

where c and l respectively denote the citing- and cited country, and i is the common industry. The empirical estimation includes country-sector fixed effects for both the citing (f_{ci}) and cited country-sector ($\tilde{f}_{li}$) to control for the quantity of patenting and for institutional or cultural factors that influence the propensity to patent and cite. The variables are defined such that if the cited and citing country-sectors are the same, the values of all regressors (except the fixed effects) are zero: no geographic or linguistic border needs to be crossed and no technological differences exist. Given the exponential function, the zero value of the regressors assures that the predicted flow of information excluding fixed effects is always equal to 1 within a given country-sector ($\hat{\phi}_{ccit}$).[10] The variable $\hat{\phi}_{clit}$ is thus a measure of the

[8] The difference in technological specialization is based on compositional differences in patent application. Similar to Peri (2005), for each country-sector a vector is produced where the cells are the proportions of all patent applications that relate to each of the 23 IPC subsections. The variable is then defined as 1 minus the uncentered correlation between the two country-sectors' proportion vectors. Both technological variables use absolute values in order to capture the technological proximity, independently of the sign.

[9] The empirical equation is based on the assumption that the probability that an idea created in country-industry (l, i) is cited by (c, i) during an interval τ, is $\phi_{cli} = e^{f(c,l,i)}(1 - e^{-b\tau})$.

[10] The removal of the fixed effects from the predicted values allow us to exclude the country-time drivers of patent citations that, by themselves, do not affect the intensity of bilateral international knowledge flows.

	1 1995-2014	2 1995-99	3 2000-04	4 2005-09	5 2010-14
diff_country	-0.457***	-0.595***	-0.424***	-0.405***	-0.554***
	(-3.69)	(-7.45)	(-5.56)	(-4.53)	(-4.47)
diff_next	-0.124	-0.333***	-0.0543	0.0762	-0.381**
	(-0.93)	(-4.89)	(-0.64)	(0.71)	(-2.19)
diff_lang	-0.810***	-0.539***	-0.679***	-0.913***	-0.930***
	(-11.96)	(-10.42)	(-11.29)	(-12.31)	(-9.02)
dist_int	-24.93	17.28**	-26.19***	-45.08***	2.041
	(-1.51)	(1.96)	(-2.81)	(-4.28)	(0.11)
tech_spec	-2.214***	-3.779***	-3.991***	-3.440***	-3.009***
	(-3.30)	(-8.32)	(-9.58)	(-5.46)	(-3.74)
tech_dev_RnD	-0.0655	-0.143***	-0.140***	-0.133**	0.0391
	(-0.68)	(-3.89)	(-2.85)	(-2.03)	(0.33)
_cons	4.279***	5.858***	6.348***	4.574***	4.311***
	(19.58)	(52.98)	(60.49)	(29.19)	(12.81)
Citing-Cou-Ind. FE	Yes	Yes	Yes	Yes	Yes
Cited-Cou-Ind. FE	Yes	Yes	Yes	Yes	Yes
N	1759	1134	1263	1710	1717
r2	0.997	1.000	1.000	0.998	0.993

Table 1: Gravity regression with G5 as cited countries. Standard errors are robust and clustered at the citing country-industry level. T-values in parentheses and significance levels are indicated with $^*p < 0.10$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Constant not reported.

frequency of citation relative to the "frictionless" frequency of citation within the originating country-sector and can be interpreted as the relative share of knowledge that diffuses from the cited to the citing country-sector. The advantage of this measure is that it is reasonably exogenous to our dependent variables and that it implicitly captures, in the gravity coefficients, globalization-related channels of knowledge transmission, including trade, FDI, and migration.

The model is estimated using averages of the variables over different time periods and via the Pseudo-Poisson-Maximum Likelihood estimator (PPML), a natural choice for a gravity-type model with significant heteroskedasticity, many zero entries and a large number of dummies (Santos Silva and Tenreyro (2006) and Santos Silva and Tenreyro (2011)). The estimation restricts cited countries to be members of the G5. The sample of citing countries includes 23 advanced economies and 9 emerging economies, reflecting in part the more limited

For instance, changing legal and cultural factors may affect the level of patent applications and citations at the country-level, without implying any variation in international knowledge flows.

availability of sectoral R&D data for emerging economies.[11] Table 1 shows the main results. Column (1) estimates the model estimated over 1995-2014 averages; columns (2) to (5) show the results for the model estimated over each of the 5-year sub-periods.

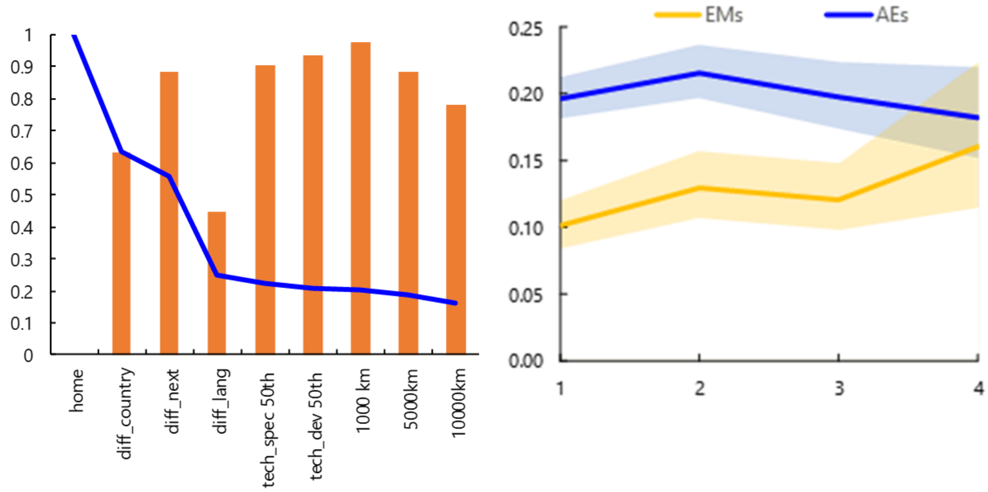
Coefficients are generally negative, which is in line with our prior that differences work as a barrier to information diffusion. In all subperiods, the coefficients for the national border ($diff.country$), a different official language ($diff.lang$) and the difference in technological specialization ($tech.spec$) are of the expected sign and statistically significant. The coefficients for contiguity ($diff.next$) and distance ($dist.int$) however move around somewhat. This is partly related to the close collinearity between the two variables. In the columns 2 and 3, dropping $dist.int$ would make the coefficient on diff.next statistically significant.[12] The reverse is however not the case for $dist.int$ in the last period, which stays insignificant even if contiguity is omitted. Another variable whose estimated coefficient is not consistent over the different estimation periods is the difference in technological development ($tech.dev.RnD$), which becomes statistically indistinguishable from zero in the last period. The pattern of the coefficients on $tech.dev.RnD$ and $dist.int$ appears consistent with the idea that both physical and technical distance have been playing a smaller role in more recent years and could reflect the deepening integration of emerging markets in knowledge flows. Exploring this possibility in more detail is however beyond the scope of this paper.

Based on the coefficients of Table 1, we construct the predicted frequencies ($\hat{\phi}_{clit}$) of citations for each country-sector pair. The left panel of Figure 2 shows how the intensity of knowledge diffusion $\hat{\phi}_{clit}$ changes as different barriers are crossed. While naturally at 100 percent in the home country-sector, knowledge diffusion declines by roughly 1/2 when information crosses a national border. While the effect of contiguity ($diff.border$) is more moderate, a different language ($diff.lang$) again significantly decreases this share by an even larger amount.

Estimating the baseline for different 5-year periods allows us to assess how the intensity of knowledge diffusion ($\hat{\phi}_{clit}$) has changed, both over time and across the groups g of either advanced or emerging economies. Specifically, for each sub-period we construct average knowledge flow intensities $\bar{\phi}_{gG5}$ from G5 countries to the average country-sector in the group g. The measure is calculated by applying the estimated coefficients in Table 1 to the average value of each regressor, with the average computed across the country-sectors belonging to the group g. The right panel of Figure 2 reports these average knowledge flow intensities and associated 95-percent confidence intervals for advanced and emerging economies. Three

[11] For a detailed list of countries included in the sample see Appendix A.1.

[12] Results not reported, but available from the authors on request.



(a) (b)

Figure 2: Panel (a): Reduction of knowledge flows with additional barriers. The bars show the fraction of knowledge that diffuses across the individual barrier. The line is the cumulative effect of the barriers. Panel (b): Average predicted knowledge diffusion $\bar{\phi}_{gG5}$ from G5 countries. Estimates are in percentage points and are computed as linear combinations of the estimated gravity coefficients and the average values of the regressors for the group of countries of interest. Figures in brackets represent 95-percent confidence bands around the linear combination of coefficients.

results stand out.

First, the average intensity of knowledge flows towards emerging economies $\bar{\phi}_{gG5}$ has increased over time from a point estimate of about 10 percent in 1995-1999 to about 16 percent in 2010-2014. As indicated by the 95 percent bands, this change is not only economically but also statistically significant[13]. Second, the intensity $\bar{\phi}_{AEG5}$ of knowledge flows to the average (non G5) advanced country hovered around 20 percent, with no statistically significant change over the sample period. Third, while in the first sub-period $\bar{\phi}_{EMG5}$ was statistically much smaller than $\bar{\phi}_{AEG5}$, during the last period we cannot reject the hypothesis that the two estimates are identical (notice that, in the last period, the confidence band for $\bar{\phi}_{AEG5}$ is contained in the band for $\bar{\phi}_{EMG5}$). This latter conclusion is the result of the stability of $\bar{\phi}_{AEG5}$, the contemporaneous rise in $\bar{\phi}_{EMG5}$, but also of the increase in the standard deviation of the estimates during the last period. Overall, these results suggest that over the last two decades the world has experienced a process of knowledge globalization, whereby international barriers to knowledge flows have weakened, allowing emerging economies to deepen their knowledge integration with G5 economies.

[13]The 95 percent confidence band for the period 1995-199 is $[8.4, 12.0]$ and for 2010-2014 is $[11.5, 22.4]$

We performed a series of robustness tests to check how these results change with different variable definitions, samples and specifications (see Appendix A.2). First we re-estimate the model excluding China from the sample, and we find that the results are largely unaffected. Second, we examine two additional extensions of the model, one where we test for technology diffusion from all countries (not only the G5) and another one where we include cross-sectoral knowledge diffusion. We find, not surprisingly, that technology diffusion is weaker across than within sectors (the collective effect of the barriers strengthens) as well as from non-G5 countries than from G5 countries. In all exercises, the relative importance of the individual barriers remains qualitatively unchanged. Finally, we calculate the estimates $\hat{\phi}_{clit}$ by proxing the *tech.dev* regressor with differences in value-added (instead of R&D) per employee. This allows us to significantly increase the number of emerging economies in our sample. We find that the evolution of the point estimates for $\bar{\phi}_{AEG5}$ and $\bar{\phi}_{EMG5}$ change little, but with the new proxy the estimates become more imprecise.[14]

Our 20 percent value for $\bar{\phi}_{AEG5}$ can be roughly compared to Peri (2005)'s 20-25 percent average share of information that diffuses from technological leaders. However, it is important to bear in mind that the sample in Peri (2005) greatly differs from ours, since it comprises of regions within advanced countries, covers an earlier period, and does not include the industry dimension.

5 Impact on innovation and productivity

Armed with an estimate of bilateral knowledge flows from the gravity model, we proceed to assess the effect of foreign knowledge flows on domestic innovation activity and productivity from the second-step regressions of type (4). Using equation (3), for each country-sector we first construct foreign knowledge flows by weighting the G5 countries R&D stocks with the corresponding $\hat{\phi}_{clit}$ obtained from the previous section.

5.1 Impact on innovation

We first focus on innovation, as measured by the (log) count of international patent families in a given country-sector, i.e. we estimate (4) with $Y_{cit} = P_{cit}$. The estimation is performed by OLS on the of sample of eleven manufacturing sectors in 27 countries over 1995-2014. As discussed in Section 2, the regression includes country-year fixed effects to control for any

[14] Results available upon request.

time-varying factors that may drive innovation trends at the country level.

Table 2 shows that, along with the impact of own R&D (coefficient γ in equation (4)), foreign knowledge flows play an important role (coefficient μ in (4)) in stimulating domestic innovation, indicating significant learning from the technological frontier. The estimated elasticity of innovation to foreign R&D is 0.35, against an elasticity to domestic R&D of about 0.45 (column 1) - both are notably smaller than those estimated in Peri (2005).[15] We also split the estimation sample into sub-samples of 15 AEs and 12 EMs recipients (columns 2 and 3).[16] We find that the foreign R&D coefficient is similar between the two groups, but its value relative to the coefficient on domestic R&D is much higher for emerging economies, where foreign R&D has roughly the same impact on domestic patenting as domestic R&D.

Similarly to what we did in Section 4, in columns 4 and 5 we look at the evolution over time of the main coefficient of interest. Specifically, we allow the coefficient on foreign R&D to vary over time by interacting it with dummy variables indicating each five-year period (with 1995-99 being the excluded period). We find a steady and statistically significant increase in the coefficient, suggesting that the impact of international knowledge flows on domestic innovation has intensified over time. Moreover, while this result holds for both advanced and emerging economies, the rise is larger for the latter. The increase in the coefficient on foreign R&D is robust to restricting the sample to be roughly balanced to avoid sample composition effects, as well as to allowing all coefficients to vary over each sub-period.[17]

As our series of patent flows, domestic R&D, and foreign R&D tend to exhibit some trends over the 1995-2014 period, we also check our results against an estimation procedure that takes into account non-stationarity. The standard approach is to test for the presence of unit roots and cointegration, and if there are, to estimate a panel cointegrating equation (e.g., see Coe et al. (2009)). Although our panel unit root test finds little evidence that the series have unit roots, the panel cointegration test generally points to cointegration between patent flows and domestic and foreign R&D series (see Appendix A.3 for details). Therefore, in column 6 of Table 2 we also report results from an alternative estimation method, namely Dynamic OLS (Kao and Chiang (2001)), which provides consistent estimates with cointegrated panel data. The procedure essentially involves adding several lags and leads of the change in the regressors and requires a strongly balanced sample (thus the sample of country-sectors becomes smaller). The number of lags we choose is two, and the number of leads is one

[15] Peri (2005) estimates an elasticity to foreign R&D of 0.4-0.47, and elasticity to domestic R&D of 0.74-0.81. As already mentioned, the many differences in the estimation sample and in the level of analysis (ours is at the industry level), make our results not easily comparable with his.

[16] See Appendix A.1 for the list of advanced economies and emerging economies in the sample.

[17] These results are available upon request.

VARIABLES	(1) Full sample	Baseline (2) Advanced economies	(3) Emerging markets	Changing diffusion (4) Advanced economies	(5) Emerging markets	Dynamic OLS (6) Full sample
Foreign R&D stock	0.350	0.353	0.342	0.232	0.115	0.298
	[0.055]***	[0.070]***	[0.088]***	[0.078]***	[0.085]	[0.070]***
Own R&D stock	0.448	0.477	0.361	0.440	0.346	0.410
	[0.061]***	[0.077]***	[0.089]***	[0.091]***	[0.107]***	[0.042]***
Foreign R&D stock * 2000-04				0.125	0.239	
				[0.034]***	[0.064]***	
Foreign R&D stock * 2005-09				0.184	0.280	
				[0.044]***	[0.076]***	
Foreign R&D stock * 2010-14				0.249	0.353	
				[0.056]***	[0.083]***	
Observations	3,487	2,345	1,142	2,132	940	1,605
R-squared	0.779	0.750	0.707	0.747	0.723	0.323
Country-Year fixed effects	YES	YES	YES	YES	YES	YES

Table 2: R&D spillovers and innovation. Dependent variable is (log) patent flows. The full sample covers 11 manufacturing industries in 27 countries over 1995-2014. The foreign R&D stock is weighted by the estimated bilateral knowledge flows between the G5 and recipient country-sector. Robust standard errors clustered at country-sector level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

(results are little affected by these choices). We obtain estimates that are similar to those in the baseline, and the coefficient on the weighted foreign R&D stock remains sizable and statistically significant. In light of this, and the fact that our panel unit root test gives inconclusive results, we keep focusing on OLS in the remainder of the analysis.

The results are also robust to a number of additional sensitivity checks, reported in Table 3. We consider, in particular: (1) *Alternative patent measures*. While the baseline uses international patent families as a measure of innovation, results are very similar using patent families with at least one application at one of the top 3 patent offices (United States Patent and Trademark Office, European Patent Office, and Japanese Patent Office), which is another measure of quality-adjusted patent counts; (2) *Alternative weighting schemes*. The baseline results are robust to using, in place of the predicted share of knowledge flow $\hat{\phi}$ based on cross-patent citations, the (time-varying) bilateral trade links between country-sectors. For each receiving country-sector, the trade weights are constructed as imports of goods from the originating country-sector as a share of the gross output of the importing country-sector;[18] (3) *Alternative fixed effects*. While the baseline specifications use country-year fixed effects, in line with Peri (2005), the results are robust to using sector-year fixed effects instead, which can capture sector-specific developments that are common across countries.[19] The coefficients

[18]Data on bilateral goods imports are from the UN COMTRADE database (we aggregate the six-digit product level data into the two-digit ISIC industry level), and data on sectoral gross output are from the World Input-Output Table (WIOT).

[19]The inclusion of both country-year and sector fixed effects removes most of the variation in the data,

VARIABLES	(1) Alternative patent measure	(2) Trade weight	(3) Sector-Year fixed effects	(4) Expanded EM sample
Foreign R&D stock	0.359	0.240	0.508	0.300
	[0.057]***	[0.033]***	[0.113]***	[0.098]***
Own R&D stock	0.464	0.468	0.724	
	[0.064]***	[0.066]***	[0.039]***	
Aggregate R&D stock * Sector R&D intensity				0.145
				[0.046]***
Observations	3,468	3,021	3,487	2,570
R-squared	0.790	0.794	0.758	0.659
Country-Year fixed effects	YES	YES	NO	YES

Table 3: R&D spillovers and innovation - Robustness. Dependent variable is (log) patent flows. In column 1, the patent measure is patent families with at least one application at one of the top 3 patent offices. Robust standard errors clustered at country-sector level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

on both foreign and domestic R&D become significantly larger under the specification with sector-year fixed effects; (4) *Expanded sample for emerging economies*. Since the availability of sector-level R&D data limits the sample to a small number of emerging economies, we also estimate an alternative specification, where the domestic sector-level R&D stock is replaced by a variable constructed from interacting the domestic aggregate R&D stock with a sector's "representative" R&D intensity. The sector's representative R&D intensity is based on U.S. data and calculated as R&D spending per employee (average for the 1995-2014 period).[20] This allows us to include a significantly larger number of EMs in the sample.[21] The coefficient on the weighted foreign R&D stock remains statistically significant with this larger sample, and although not shown here, the results regarding the time evolution of the point estimates in knowledge diffusion also hold.

5.2 Impact on productivity

To explore the possibility that foreign knowledge flows not only stimulate domestic innovation, but also directly impact the efficiency of domestic production, we estimate an augmented production function, in which domestic and foreign R&D capital enter the production

and thus the results are not discussed here.

[20]The correlation between sector-level R&D stock and this interacted variable is about 0.49 (calculated over country-sectors for which both are available).

[21]See Appendix 1 for the list of emerging economies in this larger sample.

VARIABLES	(1) Baseline	(2) Changing diffusion
Foreign R&D stock (lagged)	0.053	0.018
	[0.021]**	[0.037]
Own R&D stock (lagged)	0.060	0.058
	[0.023]**	[0.030]*
Employment	0.723	0.734
	[0.055]***	[0.065]***
Capital stock	0.337	0.329
	[0.045]***	[0.050]***
Foreign R&D stock * 2000-04		0.026
		[0.014]*
Foreign R&D stock * 2005-09		0.052
		[0.024]**
Foreign R&D stock * 2010-14		0.072
		[0.030]**
Observations	1,192	959
R-squared	0.958	0.955
Country-Year fixed effects	YES	YES

Table 4: R&D spillovers and productivity. Dependent variable is (log) real value added. The sample covers 11 manufacturing industries in 9 (mostly advanced) countries over 1995-2014. Robust standard errors clustered at country-sector level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

in addition to the traditional factors of labor and physical capital (see Acharya and Keller (2009)). In practice, instead of estimating directly (4) with a Solow residual A_{cit} as the dependent variable, we use instead sectoral output as dependent variable and include capital and labor input measures as regressors. Unlike in the patent specification, the domestic and foreign R&D variables are lagged by one year to allow for the possibility that it may take time for investment in R&D to positively impact productivity. The estimation sample is reduced to nine (mostly advanced) countries due to limited availability of sector-level capital stock data.[22]

Results are reported in Table 4. Foreign knowledge plays a role - albeit modest compared with the case of innovation - in boosting domestic productivity (column 1). This result is consistent with those in Coe et al. (2009) and Acharya and Keller (2009), who found significant international R&D spillovers on productivity, despite the differences in methodology, sample and data.[23] Specifically, our estimates indicate that a one percent increase in the

[22] See Appendix A.1 for the list of countries in this sample.

[23] For example, Coe et al. (2009) use aggregate data at the country level and estimate a panel cointegrating equation. Acharya and Keller (2009) use industry-level data as in our study, but estimate R&D spillovers from G6 countries (United States, Japan, Germany, France, United Kingdom, Canada) separately and either do not weight the foreign R&D variables or weight by import shares.

weighted foreign knowledge stock is associated with about 0.05 percent increase in the TFP of the receiving country-sector, just slightly smaller than the TFP response to a similar increase in the domestic R&D stock (0.06 percent). Again, allowing the coefficient on the weighted foreign R&D stock to vary over five-year periods shows that diffusion to TFP has strengthened over the past two decades (column 2), and the increases in the coefficient are statistically significant. This result also extends the finding in Acharya and Keller (2009) that R&D spillovers increased between the 1980s and the 1990s to the more recent time period.

6 The role of international competition

The previous sections have documented that knowledge at the frontier is an important input into the innovation process of many countries, and that this diffusion of foreign knowledge has generally strengthened over the past decades. To delve deeper into possible factors that may explain these trends, this section uses the empirical extension to the basic model in equation (5) to provide an analysis of the role of international competition in influencing the extent to which foreign knowledge is turned into domestic innovation. As discussed above, competition is among the key determinants of innovation, and the entry of emerging market firms onto the global stage has transformed the international competition landscape, with potential implications for innovation and technology transfer.

The theoretical link between competition and innovation is complex. The early literature on endogenous growth emphasized a Schumpeterian "rent effect", according to which less product market competition increases post-innovation rents for the new incumbent, thus increasing the incentives to innovate. Subsequent literature has highlighted the importance of an additional force, the "escape competition" effect: if competitive pressure is too low and profits are already large, a firms incentive to exert effort on innovation to get ahead of competitors will be low.[24] The empirical literature reflects some of these conflicting forces. For instance, policies that increase product market competition have been found to spur innovation, but only up to a certain point, after which innovation decreases (Aghion et al. (2005)). Several recent papers have examined how innovation rates in advanced economies have been affected by the increased competitive pressure stemming from globalization and the entry of China into world trade. The effect on innovation is found to be positive in

[24] In the international context, the rent and escape competition effects have a wider interpretation (Akcigit et al. (2017).

Europe and negative in the United States (Autor et al. (2016); Bloom et al. (2016)).[25]

A related discussion investigates the relationship between market concentration and competition. Theoretically, higher concentration could be consistent with higher competitive pressure - and possibly also greater innovation - for example, if innovative "superstar" firms were more likely to appear in more competitive markets (Autor et al. (2017)). However, there is empirical evidence that suggests that increased concentration in the United States is at least in part linked to reduced competition (Grullon et al. (2017); Gutiérrez and Philippon (2017)). A final crucial observation is that trends in concentration are sensitive to the definition of the relevant market. For instance, while concentration within some large countries is rising, global concentration appears to be falling, thanks to the increased role in international markets of firms from emerging market economies (Freund and Sidhu (2017)).

To estimate (5), we construct two measures of international competition and sectoral concentration. The first is import penetration from China, which captures the increased competitive pressure coming from the entering of China into global trade. It is computed as goods imports from China as a share of the receiving country-sectors gross output.[26] The second is global market concentration, measured for each sector as the global market share of the four largest firms based on sales, calculated by Diez et al. (2018) from an extensive firm-level Orbis dataset covering both advanced and emerging countries.[27] There is evidence that trade competition with China has intensified over the past two decades for all countries: import penetration from China has risen markedly, not only in the textile industry but also in innovation-intensive industries such as electrical and optical equipment and transport equipment. In addition, global market concentration in most industries appears to have declined according to our measure, in line with findings by Freund and Sidhu (2017).

VARIABLES	(1) AEs OLS	(2) EMs OLS	(3) EMs expanded	(4) AEs 2SLS
Foreign R&D stock	0.375	0.298	0.343	0.394
	[0.066]***	[0.070]***	[0.110]***	[0.073]***
Own R&D stock	0.523	0.403		0.541
	[0.096]***	[0.061]***		[0.095]***
Trade with China	2.032	2.716	1.839	5.969
	[0.488]***	[0.530]***	[0.636]***	[1.200]***
Foreign R&D stock * Trade with China	1.022	1.126	0.802	2.870
	[0.255]***	[0.426]***	[0.387]**	[0.590]***
Aggregate R&D * R&D intensity			0.147	
			[0.054]***	
Observations	1.593	922	1.612	1.593
R-squared	0.772	0.616	0.612	0.699
Country-Year FE	YES	YES	YES	YES

Table 5: Innovation and China shock. Dependent variable is (log) patent flows. In column 4, trade with China is instrumented using Chinese import penetration for the US. Robust standard errors clustered at country-sector level. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

6.1 Impact on innovation

Results on the impact of Chinese trade competition on domestic innovation, measured by patenting activity, and diffusion of foreign knowledge are reported in Table 5. In both advanced and emerging economies, stronger trade competition with China is found to boost domestic innovation and technology diffusion, as indicated by positive and statistically significant main and interaction effects (columns 1 and 2). Thus, industries that are more exposed to Chinese import competition are not only more innovative, but they also use foreign knowledge more efficiently. The magnitude of the estimates is broadly similar across the two country groups. We also try to expand the sample of EMs by replacing the domestic sectoral R&D variable with the interaction between aggregate R&D and sectoral R&D intensity (similar to the analysis in the previous section). Results broadly hold for this larger sample as well (column 3).

There is a possibility that the OLS coefficient on the China trade variable is biased. One

[25]To complicate matters further, product market competition appears to interact in important ways with the degree of intellectual property rights protection - another determinant of innovators' rents. For instance, some evidence suggests that stronger product market competition is associated with more innovation only when intellectual property rights protection is strong (Aghion et al. (2015)). However, while strong protection motivates multinational companies to transfer technology across countries, it reduces innovation in other contexts (Williams (2013); Bilir (2014)). See Also, Boldrin and Levine (2008).

[26]This measure is available for the period 1998-2014. Alternative measures using final instead of total goods trade yield similar results.

[27]The calculation of market shares uses the largest 1000 firms in each four-digit NACE sector from 28 countries. The shares of the largest four firms are then aggregated to the two-digit ISIC industry level using sectoral revenues as weights. This measure is available for the period 2000-15. An alternative measure using the Herfindahl index produces similar results. We thank Federico Diez for providing us with these data.

potential source of bias is reversed causality, by which less innovative industries are likely to be less competitive and hence to experience stronger import competition from China - this would tend to bias the coefficient downward. Another possible source of bias, stressed by Autor et al. (2017), comes from the simultaneous determination of innovation and China import penetration by omitted variables such as a domestic demand shock. For example, a positive domestic demand shock would increase the demand for Chinese imports and at the same time may increase or reduce domestic innovation - the former if domestic firms, whose profits rise with greater demand, direct more resources towards innovating, the latter if a rise in demand diminishes the need for innovation. The direction of the bias from this source is, therefore, a priori unclear.

To address these issues, we follow the approach in Autor et al. (2014) to instrument import penetration from China in the non-G5 advanced economies with import penetration in the US.[28] When instrumenting, the coefficients on trade with China and on its interaction with the weighted foreign R&D stock remain strongly statistically significant but with larger magnitude, possibly indicating a downward bias in the OLS estimates (column 4).[29] We do not perform the two-stage least square estimation for the group of emerging economies sample, since in this case the validity of the US import penetration as an instrument is highly dubious. Finally, we also check for non-linear effects by running the OLS regression on a full sample including quadratic terms.[30] Similarly to Aghion et al. (2005) we find an "inverted-U" relationship (Figure 3) between competition and innovation, with stronger international competition initially associated with more innovation and the relation becoming negative for higher levels of competitive pressure. Note, however, that the vast majority of observations in our sample lie in the upward-sloping part of the inverted U-curve, where more competition leads to more innovation.

We repeat the analysis presented above using the global market concentration measure instead of the China shock. The risk of endogeneity, e.g. due to reversed causality, with this measure is arguably lower compared with the import penetration measure, as it is unlikely that developments in a single country can drive the global concentration trends. An exception, though, is China, given its sheer size and the evidence that the rise of Chinese firms

[28] Autor et al. (2014) did the opposite in their U.S.-focused study, instrumenting import penetration in the United States with that in other advanced economies.

[29] 2SLS coefficients can also be larger than the ones from OLS in the presence of weak instruments. However, our confidence in the instrumentation is supported by first-stage test statistics, which strongly reject the hypotheses of under- or weak identification, suggesting that the instruments are relevant. In particular, the Kleibergen-Paap rk LM statistic (under-identification test) is 22.4, with a p-value of 0. The Kleibergen-Paap rk Wald F statistic (weak identification test) is 23.702, also comfortably larger than any Stock-Yogo weak ID test critical values.

[30] Detailed results available upon request.

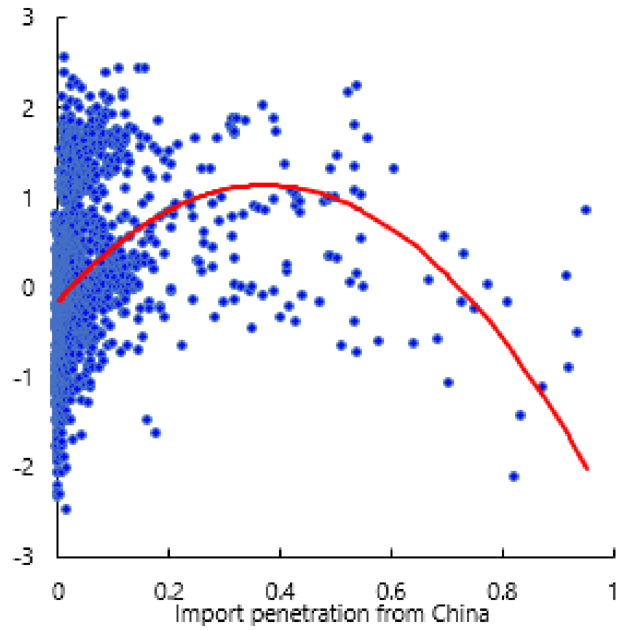


Figure 3: Relationship between innovation and China trade competition. Predicted values and fitted quadratic relation.

has contributed to the observed decline in global concentration in several sectors (Freund and Sidhu (2017)). Thus, we exclude China from our sample of emerging markets. Table 6 shows that higher global concentration for an industry, as indicated by the share of the top four firms in global markets, has a negative and statistically significant effect on both innovation and knowledge diffusion from G5 leaders, and this is true for both non-G5 advanced economies and emerging markets (columns 1, 2 and 3). Once we add quadratic terms to the regressions, we find again some evidence of an inverted-U shape. The non-linearity is statistically significant but quantitatively very small (Figure 4). Most observations lie on the downward-sloping part of the inverted U, i.e. where less concentration is associated with higher productivity.[31]

In sum, our analysis provides some evidence that, for a broad sample of both advanced and emerging market economies, stronger international competition is associated with greater innovation and technology diffusion at the sector level. This supports findings by Bloom et al. (2016), who examine the effect of the China trade shock on firms in twelve European countries. On the other hand, our findings are not necessarily inconsistent with those in Autor et al. (2016), who find a negative impact of Chinese trade competition on innovation in U.S. firms, as the United States is not among our recipient countries.

[31] Detailed results are available upon request.

VARIABLES	(1) AEs baseline	(2) EMs baseline	(3) EMs expanded
Foreign R&D stock	0.657	0.621	0.635
	[0.063]***	[0.066]***	[0.075]***
Own R&D stock	0.469	0.334	
	[0.067]***	[0.063]***	
Top 4 global market share	-4.221	-2.714	-3.027
	[0.634]***	[0.906]***	[0.548]***
Foreign R&D stock * Top 4 global market share	-1.580	-2.563	-2.320
	[0.344]***	[0.484]***	[0.313]***
Aggregate R&D * R&D intensity			0.107
			[0.041]***
Observations	1,712	790	1,819
R-squared	0.819	0.621	0.648
Country-Year FE	YES	YES	YES

Table 6: Innovation and global concentration. Dependent variable is (log) patent flows. Robust standard errors clustered at country-sector level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

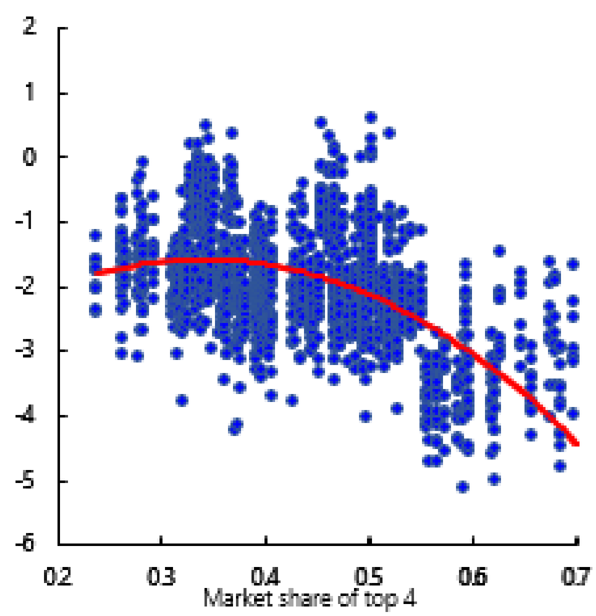


Figure 4: Relationship between innovation and concentration. Predicted values and fitted quadratic relation.

VARIABLES	(1) China shock, OLS	(2) China shock, 2SLS	(3) Concentration
Foreign R&D stock (lagged)	0.064	0.069	0.090
	[0.021]**	[0.019]***	[0.019]***
Own R&D stock (lagged)	0.066	0.063	0.069
	[0.028]*	[0.025]**	[0.026]**
Employment	0.689	0.683	0.728
	[0.060]***	[0.054]***	[0.046]***
Capital stock	0.325	0.336	0.287
	[0.034]***	[0.036]***	[0.033]***
Trade with China	0.108	0.427	
	[0.111]	[0.205]**	
Foreign R&D stock * Trade with China	0.067	0.215	
	[0.045]	[0.099]**	
Top 4 global market share			-0.008
			[0.144]
Foreign R&D stock * Top 4 global market share			-0.418
			[0.156]**
Observations	823	823	783
R-squared	0.963	0.962	0.964
Country-Year FE	Y	Y	Y

Table 7: TFP, China shock, and global concentration. Dependent variable is (log) real value added. Robust standard errors clustered at country-sector level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

6.2 Impact on productivity

As in Section 5, we complete the analysis with an investigation of the impact on industry productivity (Table 7). Again, the sample is reduced due to data availability and consists of mainly advanced economies. While the OLS estimation reveals no statistically significant effect (column 1), results from instrumenting Chinese import penetration suggest that increased trade competition from China also helps boost production efficiency and diffusion to TFP (column 2). Instrumentation seems critical in this specification as the risk of reversed causality is clear: lower productivity sectors are more likely to have higher Chinese import penetration, leading to a downward OLS bias to the estimated coefficients. In column 3, trade competition with China is replaced by the global concentration measure. We find that, while the degree of sectoral concentration does not seem to affect productivity directly (the main effect is not statistically significant), higher concentration is associated with reduced diffusion of foreign knowledge to productivity (the interaction effect is negative and significant). Finally, we run once again the regressions including a quadratic term, which in this turns out to be statistically non-significant.

7 Conclusions

In this paper we investigated the empirical importance of foreign knowledge for domestic innovation. Our methodology builds on the work of Peri (2005), which we extend along several directions. First, we estimate the intensity of cross-country knowledge diffusion and its importance for domestic innovation at a disaggregated industry level, including a number of emerging economies and focusing on a more recent sample period, which largely overlaps with the globalization process of the last two decades. Second, our rich dataset allows us to estimate with statistical precision how the intensity of knowledge diffusion and its impact on domestic innovation has evolved over time. Third, we explore one possible channel that can affect knowledge diffusion, namely changes in international competitive pressure, which we proxy, alternatively, with a “China shock” and with changes in global concentration. We find that, over the entire sample, only 15 percent of the G5s domestic knowledge diffuses internationally. This overall figure masks important source of geographical and time heterogeneity. Specifically, the intensity of knowledge diffusion has remained roughly stable at around 20 percent in advanced economies, while the diffusion towards emerging economies progressively increased from 10 percent in the period 1995-1999 to 16 percent in the period 2010-2014. We also estimate to what extent foreign knowledge, once it diffuses to the domestic economy, contributes to increasing domestic innovation outcomes. We find that foreign knowledge is about 80 percent as effective as domestically generated knowledge in raising domestic innovation. Looking at sub-samples we find again a statistically significant increase in the estimated coefficients, with the change being especially large for emerging economies. Finally, our results indicate that heightened international competitive forces had a positive impact on domestic innovation and on its sensitivity to foreign-generated knowledge.

References

Acharya, R. C. and Keller, W. (2009). Technology transfer through imports. *Canadian Journal of Economics/Revue canadienne dconomique*, 42:1411–48.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and innovation: An inverted-U relationship. *Quarterly Journal of Economics*, 120:701–28.

Aghion, P., Howitt, P., and Prantl, S. (2015). Patent rights, product market reforms, and innovation. *Journal of Economic Growth*, 20:223–62.

Akcigit, U., Ates, S., and Impullitti, G. (2017). Innovation and trade policy in a globalized world. *Unpublished*.

Autor, D., Dorn, D., Hanson, G. H., Pisano, G., and Shu, P. (2016). Foreign competition and domestic innovation: Evidence from U.S. patents. *NBER Working Paper, National Bureau of Economic Research, Cambridge, MA*, 22879.

Autor, D., Dorn, D., Hanson, G. H., and Song, J. (2014). Trade adjustment: Worker level evidence. *Quarterly Journal of Economics*, 129:1799–1860.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Reenen, J. V. (2017). The fall of the labor share and the rise of superstar firms. *NBER Working Paper, National Bureau of Economic Research, Cambridge, MA*, 22879.

Aw, B. Y., Roberts, M. J., , and Xu, D. Y. (2011). R&D investment, exporting, and productivity dynamics. *Review of Economic Studies*, 101:1312–44.

Bilir, L. K. (2014). Patent laws, product life-cycle lengths, and multination activity. *The American Economic Review*, 104:1979–2013.

Bloom, N., Draca, M., and Reenen, J. V. (2016). Trade induced technical change? the impact of chinese imports on innovation, IT and productivity. *Review of Economic Studies*, 83:87–117.

Bloom, N., Jones, C. I., and John Van Reenen, M. W. (2017). Are ideas getting harder to find? *NBER Working Paper*, 23782.

Bloom, N., Romer, P. M., Terry, S. J., and Reenen, J. V. (2013). A trapped-factors model of innovation. *American Economic Review*, 103:208–213.

Boldrin, M. and Levine, D. K. (2008). *Against Intellectual Monopoly*. Cambridge University Press, NY.

Coe, D. T. and Helpman, E. (1995). International R&D spillovers. *European Economic Review*, 39:859–87.

Coe, D. T., Helpman, E., and Hoffmaister, A. W. (2009). International R&D spillovers and institutions. *European Economic Review*, 53:723–41.

Coelli, F., Moxnes, A., and Ulltveit-Moe, K. H. (2016). Better, faster, stronger: Global innovation and trade liberalization. *NBER Working Paper*, 22647.

Diez, F., Fan, J., and Villegas-Sanchez, C. (2018). Global declining competition. *forthcoming IMF Working Paper*.

Eaton, J. and Kortum, S. (2002). Technology, geography and trade. *Econometrica*, 6:1741–79.

Freund, C. and Sidhu, D. (2017). Global competition and the rise of China. *PIIE Working Paper*.

Gilbert, R. (2006). Looking for mr. schumpeter: Where are we in the competition innovation debate? In *Innovation Policy and the Economy*, volume 6. The MIT Press.

Grullon, G., Larkin, Y., and Michaely, R. (2017). Are us industries becoming more concentrated? *Unpublished Working Paper*.

Gutiérrez, G. and Philippon, T. (2017). Declining competition and investment in the U.S. *NBER Working Paper, National Bureau of Economic Research, Cambridge, MA.*, 23583.

Kao, C. and Chiang, M.-H. (2001). On the estimation and inference of a cointegrated regression in panel data. In *Nonstationary Panels, Panel Cointegration, and Dynamic Panels (Advances in Econometrics)*, volume 15. Elsevier.

Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42:752–82.

Keller, W. (2010). International trade, foreign direct investment, and technology spillovers. In *Handbook of the Economics of Innovation*, volume 2, pages 793–829. Elsevier.

Kortum, S. (1997). Research, patenting and technological change. *Econometrica*, 6:1389–1419.

Long, C. X. and Wang, J. (2016). Global declining competition. *Evaluating Patent Promotion Policies in China: Consequences for Patent Quantity and Quality.*

Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics*, 61:653–670.

Pedroni, P. (2004). Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory*, 20:597–625.

Peri, G. (2005). Determinants of knowledge flows and their effect on innovation. *Review of Economics and Statistics*, 87 (2):308–22.

Santos Silva, J. and Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, 88:641–58.

Santos Silva, J. and Tenreyro, S. (2011). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters*, 112:220–22.

Williams, H. L. (2013). Intellectual property rights and innovation: Evidence from the human genome. *Journal of Political Economy*, 121:1–27.

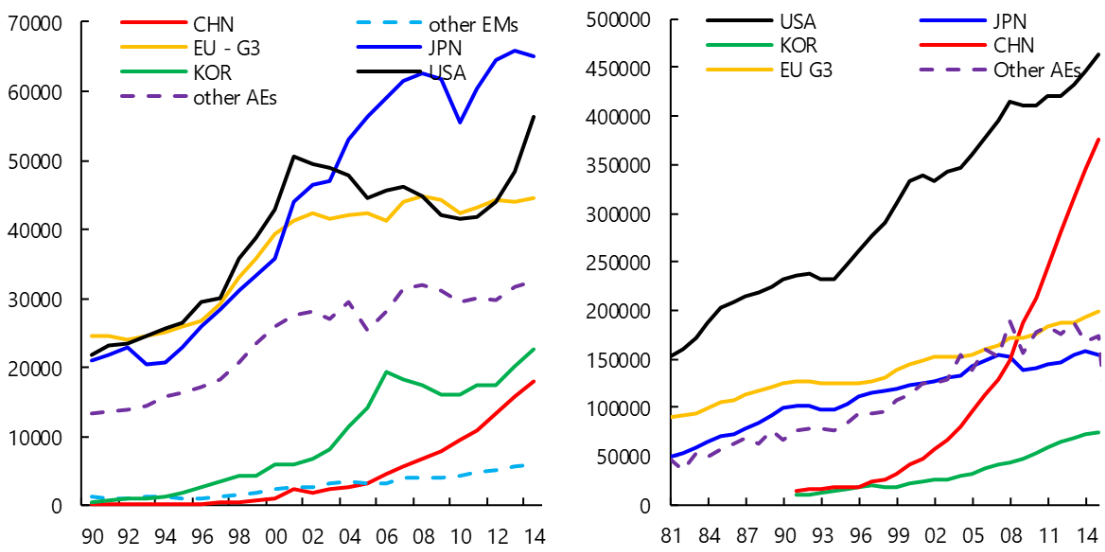


Figure 5: Patent count (left panel): international patent families by publication year and residence of first inventor of first application. Other AEs include: AUT, AUS, BEL, CAN, CHE, DNK, ESP, FIN, IRL, ISR, ITA, NLD, NOR, NZL, PRT, SWE, SGP. Other EMs include: ARG, BGR, BRA, CHL, COL, CZE, EST, HUN, IDN, IND, MEX, MYS, PER, PHL, POL, ROU, RUS, SVN, SVK, THA, TUR, UKR, URY, VNM, ZAF. R&D spending (right panel): million USD. Given low data quality, the line for other EMs is omitted.

A Appendices

A.1 Data and regression samples

This appendix shows descriptive statistics of our patent count and R&D expenditure data, lists the definitions and sources for the all the variables used in the empirical exercise, and provides the composition of country samples.

Figure 5 plots international patent families counts, aggregated across sectors and for the main countries and regions in our sample. EU-G3 comprises Germany, France and the United Kingdom. On average, 3/4 of all international patent families can be attributed to a first inventor in a G5 country. Korea and China have emerged as new global innovators over the course of our sample.[32]

Table 8 provides data sources. Table 9 and Table 10 list countries and sectors used in the estimations.

[32]The increase in Chinese patenting would be much more dramatic if simple patent applications would be used. Their representativeness is however questionable, as the Chinese Patent Promoting Policies (PPP) have contributed not only to an explosion in patent applications, but also to a deterioration in the average patent quality (e.g. Long and Wang (2016)).

Variable	Definition	Source
Patent Flows (international)	Patent families with an application in at least two distinct patent offices	Constructed from PATSTAT
Patent Flows (top three)	Patent families with an application in at least one of top three patent offices (EPO, JPO, USPTO)	Constructed from PATSTAT
R&D Expenditure	Spending on research and development, in constant price PPP US dollar	OECD ANBERD database
R&D Stock	Cumulated R&D expenditure constructed using perpetual inventory method (with discount rate = 10 percent)[1]	Constructed from OECD ANBERD
Real value added	Industry value added in 2010 US dollars	KLEMS database
Employment	Number of employees in a given industry	KLEMS database
Capital stock	Industry capital stock	KLEMS database
Trade with China	Imports of goods from China as a share of sector gross output	UN Comtrade
Global Concentration	Revenue share of top four firms globally	Diez et al. (2018)
Aggregate R&D Stock	Cumulated gross domestic expenditure on R&D (in constant price PPP US dollar), constructed using perpetual inventory method (with discount rate = 10 percent)	Constructed from OECD data
Sector R&D Intensity	R&D spending per worker	Constructed from OECD and KLEMS
Technological Specialization	Uncentered bilateral correlation between two country-sectors' vectors of patent applications in the 23 IPC subsection	Constructed based on PATSTAT
Technological Distance	Absolute ln-difference between two country-sectors in the ratio of R&D (in constant PPP terms) per number of person engaged	Constructed from OECD and KLEMS
Different Country	Dummy for an international country pair	Mayer and Zignago (2011)
Different Border	Dummy for a country pair sharing no common border	Mayer and Zignago (2011)
Different Language	Dummy for a country pair sharing no common official language	Mayer and Zignago (2011)
International Distance	Distance between the capital cities of two countries, zero for the same country pair	Mayer and Zignago (2011)
Bilateral Citations	Sum of citations between two country-industry pairs, within four years of the publication of the cited patent[2]	Constructed based on PATSTAT

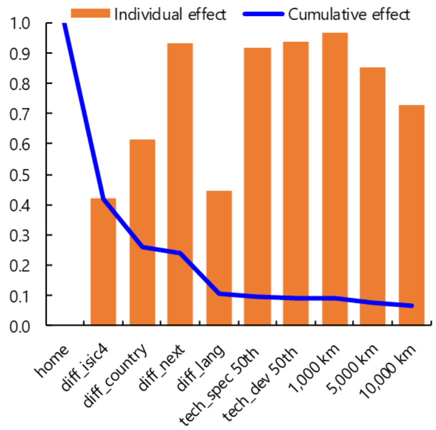
Table 8: For the CEPIIs distances measures see the GeoDist Database. Acronyms: EPO = European Patent Office; IPC = International Patent Classification; JPO = Japan Patent Office; USPTO = United States Patent and Trademark Office; UNIDO = United Nations Industrial Development organisation; WIOT = World InputOutput Tables. OECD ANBERD database: for each country-sector, the initial R&D stock is assumed to equal $R\&D_0/(g+\delta)$, where $R\&D_0$ is the initial R&D spending, δ is the annual depreciation rate (assumed to be 10 percent), and g is the average growth rate of R&D spending in that country-sector during the first five years (for most country-sectors, R&D data start in 1986). Patents attribution: We attribute a patent to the publication year for the patent count, but to the application year for the citation count Self-citations, which are defined as citations between patents with identical inventor, are excluded from the count). Citations that occur more than four years after the publication of the cited patent are also excluded (the bulk of citations happens within very few years after publication and restricting the window to either 3 or 5 years hardly affects relative citation counts).

ISIC4 Code	Sector Description
10–12	Food products, beverages, and tobacco
13–15	Textiles, wearing apparel, leather, and related products
16–18	Wood and paper products, printing, and reproduction of recorded media
19	Coke and refined petroleum products
20–21	Chemicals and chemical products
22–23	Rubber and plastics products, and other non-metallic mineral products
24–25	Basic metals and fabricated metal products, except machinery and equipment
26–27	Electrical and optical equipment
28	Machinery and equipment, not elsewhere classified
29–30	Transport equipment
31–33	Other manufacturing, repair and installation of machinery and equipment
F	Construction
62–63	Information technology and other information services

Table 9: List of sectors in estimation samples. Due to R&D data availability, the construction and IT services sectors are only included in the first-stage regression (gravity).

Regression	Advanced Economies	Emerging Market Economies
Gravity model of knowledge diffusion sample (with technological distance based on research and development)	Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Israel, Italy, Japan, Korea, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, United Kingdom, United States	China, Czech Republic, Estonia, Hungary, Mexico, Poland, Slovenia, Slovakia, Turkey
Alternative gravity model of knowledge diffusion sample (with technological distance based on value added)	Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Israel, Italy, Japan, Korea, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, United Kingdom, United States	Argentina Brazil, Bulgaria, Chile, China, Colombia, Czech Republic, Estonia, Hungary, India, Indonesia, Malaysia, Mexico, Poland, Russia, Slovakia, Slovenia, South Africa, Thailand, Turkey, Uruguay, Vietnam
Patent sample	Australia, Austria, Belgium, Canada, Denmark, Finland, Ireland, Israel, Italy, Korea, Netherlands, New Zealand, Norway, Portugal, Singapore, Spain, Sweden, Switzerland	China, Czech Republic, Estonia, Hungary, Mexico, Poland, Slovakia, Slovenia, Turkey
Patent sample, expanded emerging market economy sample		Argentina, Bulgaria, Brazil, Chile, China, Colombia, Czech Republic, Estonia, Hungary, India, Malaysia, Mexico, Poland, Russia, Slovakia, Slovenia, South Africa, Turkey, Uruguay
Total factor productivity sample	Austria, Denmark, Finland, Italy, Netherlands, Spain, Sweden	Czech Republic, Slovakia

Table 10: List of countries in estimation samples. The classification of countries into advanced economies and emerging economies is as of the beginning of the sample period, that is, around 1995. Israel, Korea, and Singapore all became advanced economies around 1997 and thus are classified as advanced economies in the sample.



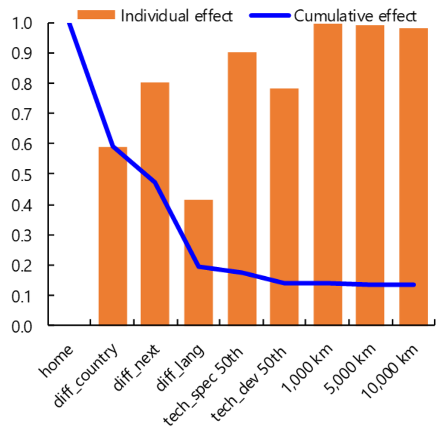


Figure 6: Effect of barriers, including cross-sectoral citations (left panel). Effect of barriers, with unrestricted source country (right panel). The bars show the fraction of knowledge that diffuses across the individual barrier. The line is the cumulative effect of the barriers.

A.2 Sensitivity analysis for gravity model

This appendix presents a series of sensitivity analyses to illustrate how results depend on the extension of the model to inter-sectoral spillovers or spillovers from all countries as well as the exclusion of China from the estimation.

Inclusion of cross-sectoral citations. In the baseline we restrict ourselves to citations between same-sector pairs, as this is where technology diffusion is likely to be most pronounced. Here we expand the sample to include cross-sectoral patent citations, which allows us to test the effect of an additional dummy $diff.isic$ indicating whether the citing and cited sectors differ. The left panel of Figure 6 presents the regression result for the share of knowledge that flows from a given country-sector. As can be expected, crossing a sectoral barrier entails a significant reduction in knowledge diffusion. Accordingly, the average $\hat{\phi}$ now converges to levels just below 10 percent, roughly half compared to the same-sector setup.

Unrestricted source country sample. Instead of expanding the sample to cross-sectoral citations, we can relax the restriction that the cited patents are from a G5 inventor. In this specification all countries enter on the cited as well as citing side, meaning they are both source and recipient of technology diffusion. The differences with the baseline estimation are small, shown in the right panel of Figure 6. The effect of most barriers is slightly higher than in the baseline, which is consistent with findings in Peri (2005) that information from non-leaders tends to diffuse less.

Excluding China from the baseline regression. The exclusion of China from the baseline estimation seems justified as its size and rapid evolution could be the main driver of the results. We find that the importance of the national border is reduced by about one half, partly compensated by an increased importance of technology variables. Moreover, a shift is observed between sharing a border (getting weaker) and international distance (getting stronger). Overall, point estimates and the average $\hat{\phi}$ are largely comparable to our baseline.

	Patenting	Domestic R&D	Foreign R&D
Inverse chi-squared	0.00	0.00	0.00
Inverse normal	0.00	0.48	0.81
Inverse logit	0.00	0.00	0.00
Modified inv. chi-2	0.00	0.00	0.00

Table 11: Fisher-type panel unit root test. The table presents p-values for rejecting the Null hypothesis that all panels contain unit root in favor of the alternative that at least one panel is stationary, as a function of the assumed asymptotic distribution. The test is based on augmented Dickey-Full tests, allow for panel-specific AR parameters, but not trend or drift.

A.3 Unit root and co-integration tests

In one of the robustness checks we use dynamic OLS to estimate the impact of foreign knowledge on innovation. This appendix provides the results of the stationarity and cointegration tests supporting it. Coe et al. (2009) have found that TFP and R&D both have unit roots and are cointegrated. It is thus not unreasonable to expect that the same applies to patenting and R&D. Our test results are however rather inconclusive, as particularly the evidence of unit roots is generally weak. To the extent that this still warrants testing for cointegration, we find some support for it, which motivates the use of dynamic OLS for a sensitivity analysis.

We start by testing the order of integration of the three series of interest, which are patenting and the stocks of foreign and domestic R&D spending. For all of them, the panel is strongly unbalanced, which restricts us to the use of Fisher type tests.[33] Table 11 shows the p-values for rejecting the Null hypothesis that all panels contain unit roots.

The test clearly rejects the Null hypothesis for patenting but provides more mixed results for domestic and foreign R&D depending on which asymptotic distribution is assumed. While the results would rather support the rejection of the Null, the tests are not conclusive. To err on the side of caution, we still proceed with a cointegration test.

Panel cointegration tests, such as the ones proposed by Kao and Chiang (2001), Pedroni (1999) and Pedroni (2004)) and Williams (2013), generally allow for unbalanced panels but require long enough time series to run time-series regressions on the individual panels.[34] The results of the various tests in Table 12 all clearly reject the null hypothesis of "no cointegration" in favor of the alternative that either all or at least some panels are cointegrated. However, such result should be read in conjunction with our previous finding that there is only weak evidence of unit roots in the panel suggests, a fact that by itself would not call for the use of dynamic OLS. For this reason, we present the results of dynamic OLS simply as a sensitivity analysis. Finally, notice also that due to data limitation the cointegration tests as well as the sensitivity analysis using dynamic OLS is not presented for TFP.

[33] The Im-Pesaran-Shin test which generally works with unbalanced panels had insufficient observations even when using interpolation to fill in missing data in the series.

[34] In order to satisfy this condition for the three variables, we linearly interpolate missing years in the data and drop any country-sector with less than 5 observations.

	Stat	p-value
Kao test for cointegration (H_0 = no cointegration; H_a : all panels are cointegrated)		
Modified Dickey-Fuller t	-3.75	0.00
Dickey-Fuller t	-15.30	0.00
Augmented Dickey-Fuller t	-2.39	0.01
Unadjusted modified Dickey	-37.52	0.00
Unadjusted Dickey-Fuller t	-32.85	0.00
Pedroni test for cointegration (H_0 = no cointegration; H_a : all panels are cointegrated)		
Modified Phillips-Perron t	-12.58	0.00
Phillips-Perron t	-34.69	0.00
Augmented Dickey-Fuller t	-154.86	0.00
Westerlund test for cointegration (H_0 = no cointegration; H_a : some panels are cointegrated)		
Variance ratio	-11.37	0.00

Table 12: Cointegration tests. The table shows test statistics and p-values for rejecting the Null of no integration in favor of the alternative that either all or some panels are cointegrated for different types of cointegration tests.