International Knowledge Diffusion and the Comparative Advantage of Nations

HARVARD Kennedy School
JOHN F. KENNEDY SCHOOL OF GOVERNMENT

International Knowledge Diffusion and the Comparative Advantage of Nations

Faculty Research Working Paper Series

Dany Bahar

Harvard Kennedy School and Center for International Development, Harvard University

Ricardo Hausmann

Harvard Kennedy School and Center for International Development, Harvard University

César A. Hidalgo

MIT Media Lab and Center for International Development, Harvard University

May 2012
RWP12-020

International Knowledge Diffusion and the Comparative Advantage of Nations

Dany Bahar, Ricardo Hausmann, Cesar Hidalgo

CID Working Paper No. 235
April 2012

Working Papers
Center for International Development
at Harvard University

International Knowledge Diffusion and the Comparative Advantage of Nations*

Dany Bahar[1,3], Ricardo Hausmann[1,3], and César A. Hidalgo[2,3]

[1]Harvard Kennedy School

[2]The MIT Media Lab

[3]Center for International Development, Harvard University

March 26, 2012

*Working Paper Version

We are grateful to Sebastian Bustos, Elhanan Helpman, Juan Ariel Jímenez, Robert Lawrence, Marc Melitz, Dani Rodrik, Rodrigo Wagner, Muhammed Yildirim, Andrés Zahler and Richard Zeckhauser for their thougtful comments. We are also grateful to the participants of MIT Media Lab/Harvard's CID Macroconnections Seminar and Harvard's CID Lunch on Economic Policy. All errors are our own.

Abstract

In this paper we document that the probability that a product is added to a country's export basket is, on average, 65% larger if a neighboring country is a successful exporter of that same product. We interpret our result as evidence of international intra-industry knowledge diffusion. Our results are consistent with the overall consensus in the literature on technology spillovers: diffusion is stronger at shorter distances; is weaker for more knowledge-intensive products; and has become faster over time.

1 Introduction

The process through which technology diffuses across countries has been implicit in economic debates for decades since it lies at the heart of the "Great Divergence" of incomes across countries (Maddison, 1995; Pritchett, 1997). Much of the endogenous growth literature has assumed that productivity is related to the number of intermediate inputs that are used in production (e.g. Rivera-Batiz & Romer, 1990; Romer, 1990; Aghion & Howitt, 1991; Grossman & Helpman, 1991). Yet, if intermediate inputs are tradable, the world should have converged, rather than diverged,as any country could mobilize as many intermediate inputs as are globally available. If some intermediate inputs are non-tradable (Rodriguez-Clare, 1996; Rodrik, 1996), then the number of locally available inputs will affect productivity, potentially explaining divergence. How far can intermediate inputs be mobilized internationally is therefore an important question.

International technology diffusion has been tackled explicitly in a burgeoning literature that uses a variety of approaches. Some explanations assume that technology diffuses broadly across countries because it is embedded in products, such as machines or intermediate inputs, and thus enhances the production possibilities of downstream producers that may import these intermediate inputs from faraway places. This literature has tested, among other things, for the impact of import volume and variety on the productivity of the home country (e.g. Coe & Helpman, 1995; Eaton & Kortum, 2001; Coe, Helpman, & Hoffmaister, 2009; Acharya & Keller, 2009). Other explanations take into account the non-excludable and non-rivalrous character of knowledge, and suggest that innovations developed by one firm can be acquired or copied by other firms, often in the same industry, through a variety of channels including foreign direct investment (e.g. Markusen & Venables, 1999; Larrain, Lopez-Calva, & Rodriguez-Clare, 2000; Branstetter, 2006; Keller & Yeaple, 2009).

The emerging consensus in the economic literature suggests intra-industry technological diffusion occurs predominantly at a fairly short range: it is more of a local than a global phenomenon (Jaffe et al. 1993; Branstetter, 2001; Keller, 2002; Bottazzi & Peri, 2003). This is usually interpreted as the consequence of the fact that much of technology is not embodied in materials or machines, but takes the form of tacit knowledge (Polanyi ,1962), that by definition is not codified and hence cannot be easily transferred through either blueprints or instruction manuals. In other words, knowledge diffusion requires more direct forms of human interaction which limits its scope to more localized or idiosyncratic settings (Arrow, 1969).

In a world in which knowledge diffuses preferentially at short ranges, a country's economic structure, as well as its evolution, will be shaped by the tacit knowledge available in neighboring countries. As a consequence, the evolution of a country's comparative advantage, productivity and economic growth will be partially determined by the amount and kind of tacit knowledge that exists in its neighborhood. These effects could be important because, according to Keller, foreign sources of technology account for 90% of domestic productivity growth (Keller, 2002; Keller, 2004).

Measuring knowledge diffusion and technology spillovers, however, is difficult. As noted by Krugman (1992), knowledge spillovers *"... are invisible; they leave no paper trail by which they may be measured and tracked, and there is nothing to prevent the theorist from assuming anything about them that she likes"*. Despite these difficulties, researchers have found ingenious ways to measure diffusion by proxying technology, basically, through two indicators: patent citations (e.g. Jaffe et al. 1993; Bottazzi & Peri, 2003; Branstetter, 2006) and measures of total factor productivity (e.g Coe & Helpman, 1995; Keller, 2002; Keller & Yeaple, 2009).

In this paper we provide evidence that intra-industry knowledge diffusion are strongly affected by distance, confirming the predominantly local character of the intra-industry technological diffusion process. We do so by looking at evidence of diffusion in a realm that has hitherto not been used for this purpose: the entrance of new products into the export basket of countries. We posit that making a product requires varying amounts of specific tacit knowledge. If this is so, then the similarity in the composition of exports between countries at any point in time and its dynamic evolution must contain evidence of this process. Countries that can make a product that they did not invent must have acquired the requisite knowledge from somewhere. If knowledge acquisition is affected by distance, then the countries that export a given product must be much nearer to each other than we would expect, given their other characteristics. If we look at the level of the country as a whole, its export basket must have evolved to be similar to that of its neighbors. Moreover, beyond static comparisons, we should be able to observe the role of neighbors in the dynamics of new products' adoption. In this paper we document this evidence. In particular, we show that the export basket of countries is predominantly similar to that of its neighbors and that the dynamic adoption of new export products is influenced by having a neighbor that is already a successful exporter of it.

Similarity in the export basket of countries, however, can be driven by factors other than technological diffusion. Similarity in geology, climate, tastes, factor endowments, income levels and other characteristics may have caused the export basket of neighboring countries to become similar, even in the absence of technology spillovers from one to the other. We therefore need to control, to the extent possible, for other factors that could also affect our observables. We do so in a variety of ways that will be discussed throughout the paper.

If we think of each product as the result of a production function with a finite number of intermediate goods, including the specific tacit knowledge required to produce them, then only in the presence of such knowledge a country will be able to successfully export that product. Under this view, it is straightforward to see how knowledge diffusion can actually shape the productive structure of countries, since the tacit knowledge that a country will have available to it will depend on the presence of that knowledge in nearby or otherwise highly connected places. As argued by Keller (2002), *"If technology diffusion is influenced by geographic factors, then production functions and comparative advantage will also vary systematically according to location, thereby influencing international trade of countries."*

The stylized facts documented in this paper are not trivial. Gravity models have shown that, ceteris paribus, trade is more intense at short distances (Leamer & Levinsohn, 1995). Given this, most trade theories imply that trade would make neighbors specialize in different industries, in order to exploit their comparative advantage, and thus achieve gains from trade. The greater intensity of trade at short distances would force specialization and differentiation, whether the differences that cause the specialization are based on exogenous technologies (Ricardo, 1817; Dornbusch & Fischer, 1977), differential factor intensities (Heckscher & Ohlin, 1991), specific factors or demand for varieties (Helpman & Krugman 1985). Here we show, however, that in spite of trade, and after controlling for similarity in geography, institutions and factor endowments, countries are surprisingly similar to their neighbors in terms of the composition of their export baskets.

There are two novelties in this paper. First, we look at knowledge diffusion by focusing on a country's export basket and its dynamics, rather than on industry total factor productivity or patent citations. Second, we explore the evolution of the extensive margin of trade, highlighting the role of neighbors in determining which products will enter a country's export basket.

This paper is organized as follows. First, we present a set of stylized facts based on the static export similarity between countries. We show that countries are predominately similar to their geographic neighbors and that similarity decays sharply with distance, consistent with the local spread of knowledge diffusion. We then study the dynamics of this process by looking at the probability that a country will add a new product to its export basket at a later date, given the presence of that product in the export basket of neighboring countries at an earlier date. We find that, after controlling for all time-varying sources of aggregate similarity between pairs of countries, and for differences in factor endowments, income level, and the country's own predisposition to move to that particular product, having neighbors that already export the same product is associated with a 65% increase in the probability of adopting that product. This suggests that the diffusion of knowledge moves through channels that decay strongly with physical distance. Finally, we study variations of our empirical specification to study how diffusion is affected by product complexity and explore changes in the strength of diffusion over time. We find that our results of knowledge diffusion are consistent with the previous literature: intra-industry technological spillovers are stronger at short distance, they are weaker for more complex products and they have strengthened over time.

2 Data and Stylized Facts

2.1 Data

Data on exports comes from the World Trade Flows (WTF) Dataset (Feenstra et al. 2005) and extended until 2008 using data from the UN COMTRADE Website (United Nations 2010). It contains the total export value for 1005 products using the SITC 4-digit (rev. 2) classification. We exclude countries with less than 1.2 million citizens and with total trade below USD $1 billion in 2008. We also exclude other countries with poor data on exports such as Iraq, Chad and Macau. We use time varying national variables from the World Bank's World Development Indicators (World Bank, 2010). In addition, we use data on conventionally measured factors of production (stock of physical capital, stock of human capital and land) from UNCTAD (Shirotori et al. 2010). Bilateral data, such as distance between the most populated cities, common continent or region, territorial contiguity, common colonizer and colonizer-colony relationship, are from CEPII's GeoDist dataset (Mayer & Zignago, 2011).

The final sample consists of 128 countries[1] and 1005 products and it accounts for roughly 99% of World trade, 97% of World total GDP and 95% of World population (Hausmann et al. 2011). In extended specifications of our empirical models, the dataset is reduced to the 107 countries for which

[1] The sample is reduced to 124 countries since data on bilateral distances and other characteristics is limited to this subset. The number of countries is further reduced when data on factor endowments and bilateral trade is used.

data on factor endowments and bilateral trade is available. Moreover, since the data for Former Soviet Union countries is discontinuous, meaning that it is non-existent prior to 1990 and sparse and scattered until 1995, we exclude countries from the Former Soviet Union from some of our analysis, where appropriate. The distinction is made when necessary in tables and figures.

2.2 Exploring Static Similarity

We measure the intensity with which a country exports each product by computing its Revealed Comparative Advantage (RCA) (Balassa, 1965). The RCA that a country has in a product is defined as the ratio between the share of total exports that the product represents in the country's export basket and the share of global trade in that product. For example, in the year 2000, soybeans represented 4% of Brazil's exports, but accounted only for 0.2% of total world trade. Hence, Brazil's RCA in soybeans for that year was $RCA_{Brazil,Soybeans} = 4/0.2 = 20$, indicating that soybeans are 20 times more prevalent in Brazil's export basket than in that of the world. A product is over-represented in a country's export basket if its RCA is above 1. Formally, if $X_{c,p}$ is equal to the dollar exports of country c in product p, then the RCA of country c in product p ($RCA_{c,p}$) is defined as:

$$RCA_{c,p} \equiv \frac{X_{c,p} / \sum_p X_{c,p}}{\sum_c X_{c,p} / \sum_c \sum_p X_{c,p}} \tag{1}$$

To create a measure of similarity in the export structure of a pair of countries c and c' we define the *Export Similarity Index* ($S_{c,c'}$) as the Pearson correlation between the logarithm of the RCA vectors of the two countries. We take the logarithm to avoid the top export products of a country to dominate the calculation of export similarity. This is because RCA distributions exhibit fat-tails. We also add 0.1 to each element of the RCA vectors to make sure that correlations are not driven by similarities in the RCA of products that countries export very little of or not at all. After these considerations, the *Export Similarity Index* is defined as:

$$S_{c,c'} \equiv \frac{\sum_p (r_{c,p} - \bar{r_c})(r_{c',p} - \bar{r_{c'}})}{\sqrt{\sum_p (r_{c,p} - \bar{r_c})^2 \sum_p (r_{c',p} - \bar{r_{c'}})^2}} \tag{2}$$

where $r_{c,p} = log_{10}(RCA_{c,p} + 0.1)$ and $\bar{r_c}$ is the average of $r_{c,p}$ over all products for country c.

$S_{c,c'}$ is larger than zero for pairs of countries that tend to export a similar set of goods with similar intensities, and negative for pairs of countries exporting different sets of goods. This feature of our index differs from the Finger & Kreinin (F&K) Export Similarity Index (Finger & Kreinin 1979), which is calculated as the sum of the minimums of the export shares of each pair of countries. We prefer our measure as it distinguished between products that are exported by one country and not the other from those that are exported by neither. Also, we use RCA, which gives equal weights to all products while the F&K measure privileges products with large global markets. Nevertheless, our analysis is robust to using the F&K similarity

Table 1: Summary Statistics (Year 2000)

Panel A - Statistics of Country Pairs			
	Pairs	Mean	Std. Dev.
Similarity Index	7,626	0.17	0.14
Distance	7,626	7412.62	4374.70
Log Distance	7,626	8.66	0.81
Share Borders	7,626	0.02	0.15
Common Language (Official)	7,626	0.11	0.31
Share Borders and Language (Official)	7,626	0.01	0.10
(Former) Colonizer-Colony Relationship	7,626	0.02	0.12
Have/Had a Common Colonizer	7,626	0.06	0.24
Log Total Bilateral Trade Value (Imports + Exports)	3,951	7.26	1.18
Absolute Difference Ln GDP Per Capita	7,381	1.43	1.01
Absolute Difference Ln Population	7,626	1.56	1.21
Absolute Difference Ln Physical Capital Per Worker	5,671	1.64	1.20
Absolute Difference Ln Years of Education	5,671	5.06	2.77
Absolute Difference Ln Land Per Worker	5,671	0.60	0.72

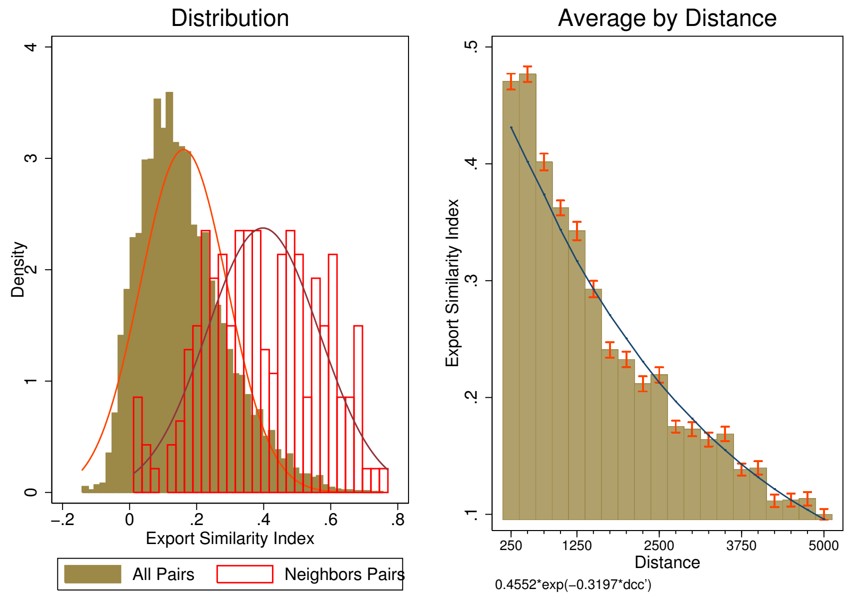
Panel B - Region of Country Pairs	Pairs	Mean	Mean within Region
Same Region	7,626	0.15	N/A
East Asia and Pacific	7,626	0.02	0.10
Eastern Europe	7,626	0.04	0.24
Western Europe	7,626	0.02	0.12
Latin America and the Caribbean	7,626	0.03	0.18
Middle East and North Africa	7,626	0.02	0.10
North America	7,626	0.0001	0.00
South Asia	7,626	0.001	0.01
Sub-Saharan Africa	7,626	0.04	0.24

index (see section A.1).

Table 1 presents summary statistics for bilateral country-level data, for the year 2000. Note that data on factor endowments and bilateral trade is limited to fewer countries.

The left panel of Figure 1 compares the distribution of Export Similarity ($S_{c,c'}$) in year 2000 for all pairs of countries (filled) and for those that share a border (unfilled), showing that countries sharing a border have productive

Figure 1: Export Similarity Index (Year 2000)



structures that are, on average, twice as similar as pairs of countries that do not share a border[2]. We show that export similarity decays exponentially with distance as $S_{c,c'} = 0.45exp(-0.32d_{c,c'})$, where $d_{c,c'}$ is the distance between country c and country c' in thousands of kilometers (right panel of Figure 1). This means that after a distance of roughly 2500 km. the average similarity index decays to half of its initial value[3].

Export similarity, however, can be the consequence of shared geology or climate, which is more likely to be the case for geographic neighbors. To

[2]The average $S_{c,c'}$ for border sharing geographic neighbors is 0.39, compared to 0.19 for non-neighbors; with t=-15.26; p-value= $7.873e^{-52}$

[3]Keller (2002) finds that spillovers halved after 1200Km. Given that in this exercise we do not control for other factors, we believe these two results are consistent.

Table 2: Lall Classification

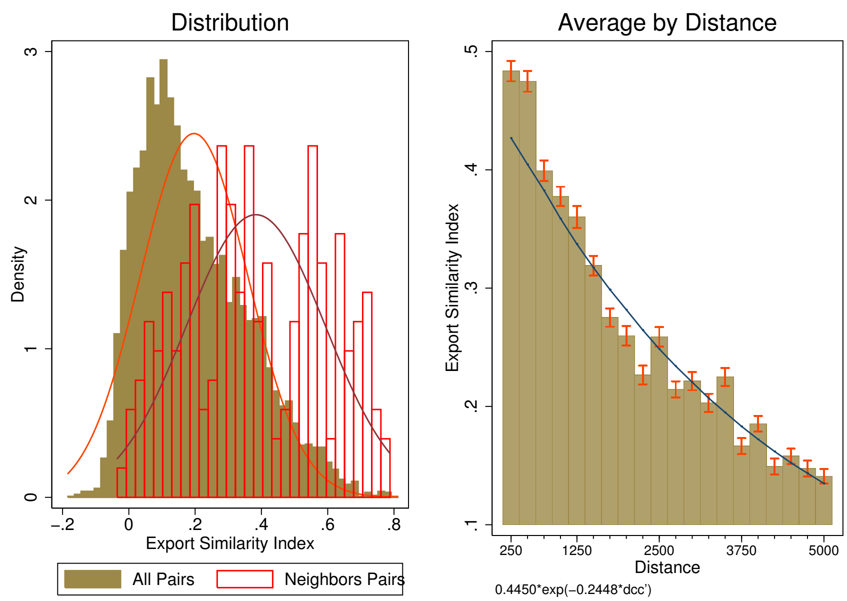
Lall Classification	# Products
Gold	1
Primary Products	193
Resource Based Manufactures 1 (agro-based products)	130
Resource Based Manufactures 2 (others non-agro based products)	108
Low Technology Manufacture 1 (textiles, garments and footwear)	100
Low Technology Manufacture 2 (others)	97
Medium Technology Manufacture 1 (automotive)	15
Medium Technology Manufacture 2 (process)	109
Medium Technology Manufacture 3 (engineering)	135
High Technology Manufacture 1 (electronic and electrical)	49
High Technology Manufacture (others)	34
Special	12
Unclassified	22

control for this fact, we exclude from the sample all products that are pinned down by geography. We do this by using the technological classification suggested by Lall (2000) that divides products in the categories presented in Table 2.

Lall's classification is used to create two categories of products: Primary and Resource Based (PRB) products and Non-Primary and Non-Resource Based (NPRB) products. We consider as PRB products those that are classified as Gold, Primary Products and Resource Based Manufactures (categories 1 thru 4 in Table 2), whereas NPRB products are the ones contained in all other categories.

Figure 2 reproduces Figure 1 using NPRB products and Figure 3 does so for PRB products only. In both cases the mean Export Similarity Index of neighboring country-pairs is significantly larger than in the overall sample

Figure 2: Export Similarity Index NPRB Products (Year 2000)



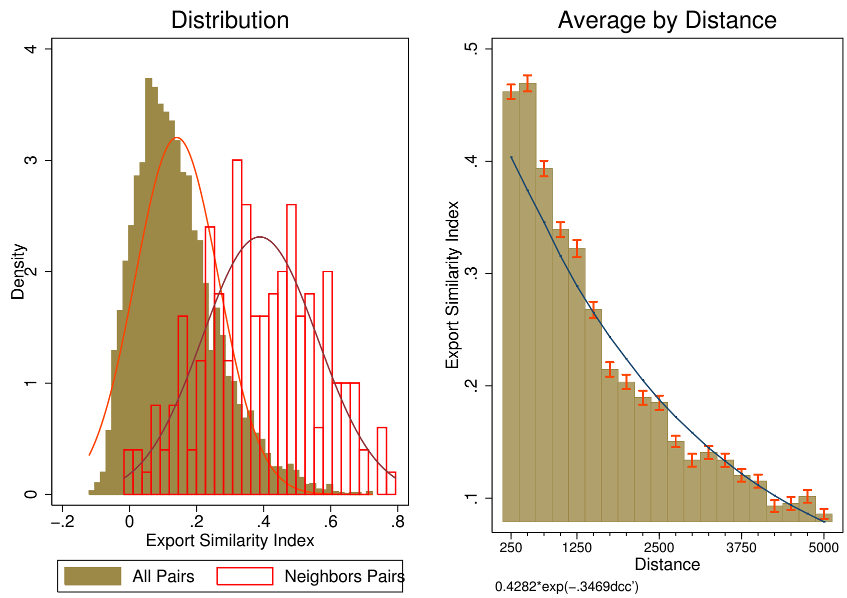
of country-pairs.[4] The graphs also show that the correlation between export similarity and physical distance is equally strong in both cases, suggesting that the observed export similarity is not driven by primary and resource based products.

2.3 Controlling for other sources of Export Similarity

The fact that, beyond geology and climate, export similarity decays with distance does not prove the presence of technological spillovers. Countries may be similar because of a shared history, culture, income levels or factor

[4]Using NPRB products the two sets are statistically different with p-value $7.873e^{-52}$. Using PRB products the two sets are statistically different with p-value $2.50e^{-147}$.

Figure 3: Export Similarity Index PRB Products (Year 2000)



endowments. To control for these factors, we formulate an adapted "gravity model" (Zipf, 1946; Tinberger, 1963) in which our index of Export Similarity ($S_{c,c'}$) is the dependent variable and where physical, cultural and economic variables are used as explanatory factors.

Our adapted gravity model follows the functional form:

$$S_{c,c'} = \alpha + \beta ln(d_{c,c'}) + z_{c,c'}\gamma + b_{c,c'}\delta + \mu_c + \mu_{c'} + \varepsilon_{c,c'} \qquad (3)$$

where $d_{c,c'}$ is the distance between countries c and c', $z_{c,c'}$ is a set of binary variables defining common characteristics between c and c', such as having a common language, sharing a border or having an historic colonial relationship. $b_{c,c'}$ is a set of continuous regressors which measures distance in quantifiable attributes between countries c and c' which can explain similarities or differences in their export baskets, such as gaps in income per capita, population and factor endowments. $b_{c,c'}$ also includes total bilateral trade between each pair of countries. Finally, μ_c and $\mu_{c'}$ are country dummies capturing any individual country characteristic for countries c and c'[5]. $\varepsilon_{c,c'}$ are the regression residuals. The results of this regression are presented in Table 3.

The six specifications presented in Table 3 sequentially introduce new correlates, and therefore differ in the total number of available observations. Column 5, on the other hand, repeats the same specification presented in column 4, but limits the dataset to those observations for which data on bilateral trade is available, so that the coefficients will be comparable to those in column 6. All regressions include country dummies that control for time invariant country characteristics, and year dummies to control for common shocks to the similarity index for all pairs of countries.

Our main result is that export similarity decreases strongly with distance in all specifications, and with a similar coefficient. In addition to the effect of distance, sharing a border increases export similarity by an amount equivalent to 0.54 to 0.78 standard deviations (since as shown in Table 1, the standard deviation of export similarity is 0.14). Note that these effects do

[5]Note this is different from one binary variable per each pair, which would make impossible to estimate the coefficient of all the invariant variables among each pair of countries, such as distance.

Table 3: Correlates of the Export Similarity Index (Year 2000)

Dependent Variable: Export Similarity Index	(1)		(2)		(3)		(4)		(5)		(6)	
Log Distance	-0.044	***	-0.045	***	-0.043	***	-0.038	***	-0.037	***	-0.038	***
	(0.003)		(0.003)		(0.003)		(0.003)		(0.004)		(0.004)	
Share Border	0.111	***	0.110	***	0.089	***	0.090	***	0.076	***	0.076	***
	(0.012)		(0.012)		(0.011)		(0.012)		(0.014)		(0.015)	
Same Region	0.055	***	0.057	***	0.026	***	0.021	***	0.018	**	0.019	**
	(0.005)		(0.005)		(0.005)		(0.006)		(0.008)		(0.008)	
Common Language			-0.021	***	-0.005		0.005		0.015	**	0.016	**
			(0.006)		(0.005)		(0.006)		(0.007)		(0.008)	
Colonial Relationship			-0.003		0.003		-0.002		0.015		0.017	
			(0.015)		(0.011)		(0.012)		(0.012)		(0.012)	
Common Colonizer			0.025	***	0.017	***	0.013	*	0.024	**	0.025	**
			(0.006)		(0.006)		(0.007)		(0.012)		(0.012)	
Absolute Difference					-0.054	***	-0.042	***	-0.046	***	-0.047	***
Ln GDPPC					(0.001)		(0.004)		(0.006)		(0.006)	
Absolute Difference					-0.012	***	-0.014	***	-0.008	***	-0.008	***
Ln Pop					(0.001)		(0.002)		(0.002)		(0.002)	
Absolute Difference							-0.018	***	-0.017	***	-0.017	***
Ln Physical Capital Per Worker							(0.003)		(0.006)		(0.006)	
Absolute Difference							0.001	**	0.001		0.000	
Ln Human Capital Per Worker							(0.001)		(0.001)		(0.001)	
Absolute Difference							-0.029	***	-0.036	***	-0.036	***
Ln Land Per Worker							(0.005)		(0.006)		(0.006)	
Log Bilateral Trade											-0.003	
											(0.003)	
Constant	0.659	***	0.676	***	0.798	***	0.594	***	0.366	***	0.387	***
	(0.036)		(0.037)		(0.036)		(0.036)		(0.036)		(0.040)	
Observations	7,626		7,626		7,381		5,460		3,175		3,175	
Adjusted R-squared	0.370		0.373		0.496		0.537		0.656		0.656	

All specifications include country dummies.
Robust standard errors in parentheses, clustered at the country-pair level.
*** p<0.01, ** p<0.05, * p<0.1

not include the added impact of sharing the same geographic region, which adds between 0.13 and 0.35 standard deviations to export similarity.

Table 4 reproduces the same specifications, using the Export Similarity Index for NPRB products only, and shows that signs of the coefficients are consistent with the ones presented in Table 3, indicating that the results are not driven by the direct effects of geology or climate on comparative advantage.

There are some other key results from these tables. Among our control variables, having a common colonizer, a (past) colonial relationship or having a common official language enhances export similarity[6]. The role of common colonizer or colony-colonizer relationship may capture past history of technological spillovers or be a proxy for similarity in institutions or government quality (La Porta et. al. 1999). As expected, differences in levels of income per capita, population, and capital and land per worker have a negative correlation with exports similarity. Interestingly, we do not find that, given the other controls, differences in human capital, measured by the difference in the average years of schooling of the labor force, is correlated with export similarity: the coefficient is small, insignificant and has the wrong sign. By contrast, bilateral trade is negatively correlated with export similarity, as it would be expected from the fact that countries trade more with economies with different productive structures. Adding it to the regression, however,

[6]Having had a colonial relationship appears only to be significant when analyzing the similarity index computed with NPRB products only. Keller (2002) also finds that having a common language enhances technology diffusion.

Table 4: Correlates of the Export Similarity Index for NPRB Products (Year 2000)

Dependent Variable: Export Similarity Index												
	(1)		**(2)**		**(3)**		**(4)**		**(5)**		**(6)**	
Log Distance	-0.024	***	-0.024	***	-0.023	***	-0.021	***	-0.019	***	-0.020	***
	(0.003)		(0.003)		(0.003)		(0.004)		(0.004)		(0.004)	
Share Border	0.090	***	0.088	***	0.065	***	0.064	***	0.056	***	0.056	***
	(0.012)		(0.012)		(0.012)		(0.013)		(0.015)		(0.015)	
Same Region	0.064	***	0.066	***	0.034	***	0.025	***	0.019	***	0.020	***
	(0.005)		(0.006)		(0.005)		(0.006)		(0.007)		(0.007)	
Common Language			-0.016	***	0.002		0.008		0.020	**	0.021	**
			(0.006)		(0.006)		(0.007)		(0.008)		(0.008)	
Colonial Relationship			0.011		0.016		0.014		0.032	**	0.032	**
			(0.016)		(0.013)		(0.014)		(0.013)		(0.013)	
Common Colonizer			0.026	***	0.018	**	0.019	**	0.046	***	0.046	***
			(0.007)		(0.007)		(0.009)		(0.015)		(0.015)	
Absolute Difference					-0.054	***	-0.055	***	-0.064	***	-0.064	***
Ln GDPPC					(0.002)		(0.005)		(0.007)		(0.007)	
Absolute Difference					-0.011	***	-0.012	***	-0.005	**	-0.005	**
Ln Pop					(0.002)		(0.002)		(0.002)		(0.002)	
Absolute Difference							-0.007	*	-0.010		-0.010	
Ln Physical Capital Per Worker							(0.004)		(0.007)		(0.007)	
Absolute Difference							0.001		0.002		0.001	
Ln Human Capital Per Worker							(0.001)		(0.001)		(0.001)	
Absolute Difference							-0.020	***	-0.026	***	-0.026	***
Ln Land Per Worker							(0.005)		(0.006)		(0.006)	
Log Bilateral Trade											-0.001	
											(0.003)	
Constant	0.526	***	0.537	***	0.701	***	0.219	***	0.224	***	0.232	***
	(0.042)		(0.042)		(0.038)		(0.045)		(0.046)		(0.051)	
Observations	7,626		7,626		7,381		5,460		3,175		3,175	
Adjusted R-squared	0.425		0.426		0.505		0.537		0.663		0.663	

All specifications include country dummies.
Robust standard errors in parentheses, clustered at the country-pair level.
*** p<0.01, ** p<0.05, * p<0.1

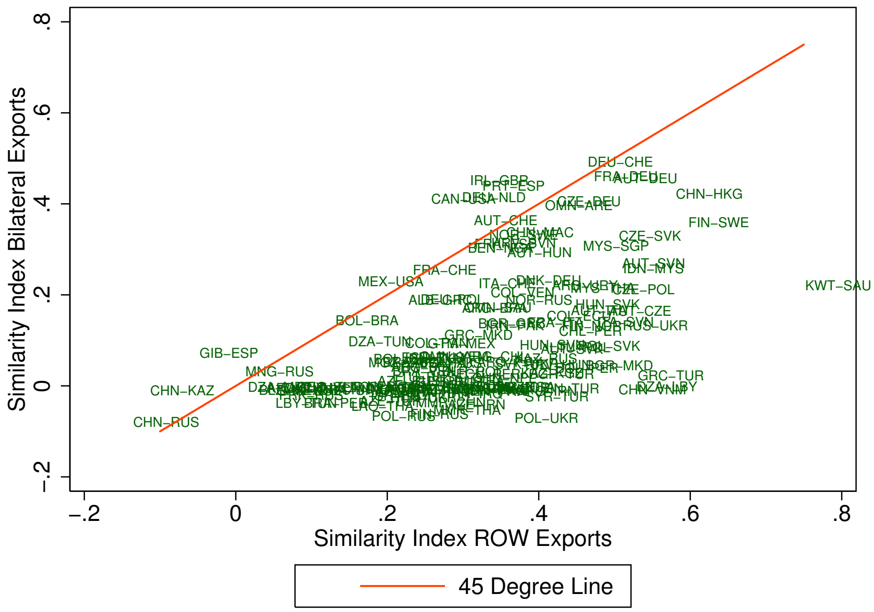
does not reverse any of the previous results.

The main takeaway from the results presented so far is that distance, borders and shared regions are important correlates of export similarity, even after controlling for a list of other variables that economic theory suggests should shape the export basket of countries. These stylized facts are consistent with a world in which productive knowledge diffuses locally as measured by the mix of products that countries export.

While we believe that the previous results control for an important set of alternative explanations, we consider now whether the similarity in the export basket of neighboring countries is driven by similarity of preferences. Following the Linder Trade Hypothesis (Linder, 1961), countries with similar preferences and hence demand structure, are likely to trade more, which in a Helpman-Krugman interpretation is due to the fact that they enjoy different varieties of similar products (Helpman & Krugman, 1985). Since neighbors trade more intensively, the similarity in bilateral trade may be driving our results. We check for this by comparing the similarity index $S_{c,c'}$ of the bilateral exports of neighbors with the similarity index of their exports to the rest of the world. The results for year 2000 can be seen in Figure 4. As is clear from the graph, neighbors are more similar in terms of what they export to the rest of the world than what they trade among themselves. This implies that export similarity is not a consequence of the composition of bilateral trade between neighbors (see section A.2 for more details).

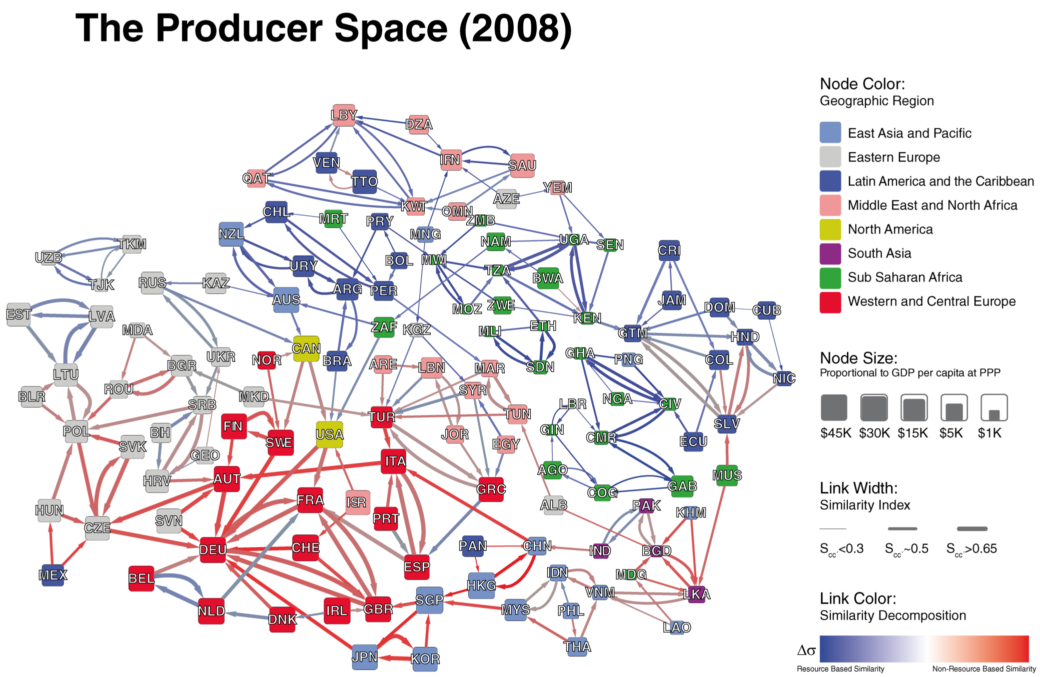
In sum, there is a puzzling similarity in the export basket of countries that

Figure 4: Neighbors Similarity (on bilateral exports vis-à-vis ROW exports)



is strongly affected by variables that proxy for distance and that is robust to the inclusion of institutional, income and factor endowment variables. In fact, one way to illustrate the strength of the similarity between neighboring countries is to represent the matrix of export similarity as a network where each country is connected to the two other countries most similar to it. Figure 5 presents the network of export similarity for year 2008 as a graphical network where each node represent a country, and each country is connected to the two other countries with the most similar export baskets, as measured by the Export Similarity Index $S_{c,c'}$. Countries are colored according to geographic regions, showing that the clusters defined by export similarity correlate strongly with physical distance. The width of the links is propor-

Figure 5: The Network of Exports Similarity (Year 2008)



tional to the similarity index and the color of the link indicates whether the similarity is driven by PRB products (blue) or by NPRB products (red) (see section A.3 for more details). We note that, in a large number of cases, the country with the most similar export structure is an immediate neighbor, such as in the case of France, Germany, Austria, the Czech Republic, Hungary and Slovakia or in the case of India, Pakistan, and Bangladesh. This visualization illustrates the strong association between proximity and export structure that characterizes the world economy.

3 Dynamics of Productive Knowledge Diffusion

The similarity of the export basket of neighboring countries, even after controlling for other potential determinants of similarity, is suggestive evidence of an accumulated history of intra-industry knowledge spillovers. Export similarity is remarkably similar among neighbors in spite of the fact that proximity-induced trade would have strengthened the forces of specialization and differentiation. Nevertheless, other omitted factors may be behind our observations. To address this concern, we look at the dynamics of the extensive margin of trade.

We study the evolution of the mix of products that each country exports to shed light on whether the evolution of this product mix is affected by the export basket of a country's neighbors, defined as those sharing a border. We study the probability that a country will add a product to its export basket in period T (i.e. "jump" to the product) if it has neighbors that are already exporting that product in period t (with $T > t$). We define a "jump" as a tenfold or more increase in the RCA of country c in product p, from $RCA_{c,p} \leq 0.1$ to $RCA_{c,p} \geq 1$ in a ten year period[7]. Moreover, we restrict jumps to two conditions. First, a jump needs to keep an RCA above 1 for four years after t+10 (the forward condition). Second, we restrict jumps to products that had an RCA below 0.1 for two years before the beginning of the period (the backward condition). These two conditions intend to rule

[7]With the exception of our last period which is a seven year period (2001-2008)

out the possibility of "temporary jumps" in the data driven by noise, errors, shocks in commodity prices or other exogenous reasons.[8]

We test our hypothesis using the following empirical specification:

$$\begin{aligned} J_{c,p,t\to T} &= \alpha + \beta_1 log(RCA_{c_N,p,t}) + \beta_2 log(RCA_{c_N,p,t})^2 \\ &\quad +\beta_3 RCA_{c,p,t} + \beta_4 g^{RCA}_{c,p,t-10\to t} + \beta_5 zeroexp_{c,p,t-10} \\ &\quad +\beta_6 density_{c,p,t} + \varphi_{p,t} + \mu_{c,c_N,t} + \varepsilon_{c,p,t} \end{aligned} \tag{4}$$

where $J_{(c,p,t\to T)}$ is a binary variable that takes the value of 1 when there was a "jump" between year t and T in product p and country c. The variable of interest, $log(RCA_{c_N,p,t})$, is the log value of the RCA of the neighboring country of c that has the largest RCA for that product, among all neighbors (named c_N). This variable is included both in its linear and squared form in order to allow for a quadratic relationship.[9] We also include a set of variables at the country-product level. This includes the initial RCA of country c in product p, the average annual growth rate of the RCA in the previous ten year period[10], a dummy variable indicating whether there were zero exports of this good at the beginning of the previous period, and the "density" of the country in the product at the beginning of the period. The variable "density",

[8] For the last period (2001-2008) we eliminate the forward condition due to data limitations.

[9] All the results are robust to the exclusion of this quadratic term.

[10] For the first period 1970-1980 we used the previous five year average annual growth rate (1965-1970) due to data limitations.

which distributes between 0 and 1, was developed by Hausmann and Klinger (2007) and used in Hidalgo et. al. (2007). It measures the proximity between a country's export basket and the product under consideration. Proximity is based on the probability that a pair of products is co-exported by the same country. In other words, the density of a product proxies for the existence of other exports that share similar technologies or inputs (as measured by their co-occurrence across countries and time). Density strongly affects the likelihood of a country adding the product to its export basket (Hausmann & Klinger, 2007; C. A. Hidalgo et al. 2007). We use it to control for the likelihood that a country would jump to a product given the initial composition of its export basket.[11] $\varphi_{p,t}$ are product-by-year fixed effects which control for any time-varying product characteristic such as global demand, price or productivity shocks. $\mu_{c,c_N,t}$ are country-neighbort-by-year fixed effects, using the neighbor that has the largest RCA in that product. By adding $\mu_{c,c_N,t}$ we control for constant and time varying country-neighbor characteristics such as commonalities in institutions, geography, climate, culture, history, productivity, economic development, population, initial factor endowments, etc.

Table 5 shows the summary statistics of the data used for this exercise. For this analysis we use four periods: 1970-1980, 1980-1990, 1990-2000 and 2001-2008.[12]

[11] All results are robust to the exclusion of this variable. In fact, the inclusion of this variable reduces the size of our estimator of interest.

[12] Since the original Feenstra data runs up to year 2000, and since 2001 and on was extended by the authors, we prefer to start the last period in 2001 to avoid discrepancies

Table 5: Summary Statistics Dynamics of Knowledge Diffusion (1970-2008)

Panel A: All Sample					
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Jump	267,170	0.009	0.09	0	1
Growth Rate RCA (10-years period)	267,170	1.20	11.47	-72.3	211.4
RCA	267,170	1.31	18.11	0	3110.2
Log RCA	267,170	-0.63	0.56	-1	3.49
Density	267,170	0.14	0.13	0	0.96
Log Maximum RCA of Neighbors	267,170	-0.33	0.70	-1	3.49
Neighbor Exports Product RCA>1	267,170	0.28	0.45	0	1
Log Maximum RCA of Random Neighbors	267,170	-0.26	0.70	-1	3.49
Neighbor (Random) Exports Product RCA>1	267,170	0.33	0.47	0	1
Panel B: Restricted Sample to observations with RCA<0.1 ("Eligible to Jump")					
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Jump	176,495	0.015	0.12	0	1
Growth Rate RCA (10-years period)	176,495	3.14	10.06	-9.42	211.4
RCA	176,495	0.0096	0.02	0	0.1
Log RCA	176,495	-0.96	0.069	-1	-0.69
Density	176,495	0.087	0.086	0	0.66
Log Maximum RCA of Neighbors	176,495	-0.52	0.63	-1	3.49
Neighbor Exports Product RCA>1	176,495	0.18	0.38	0	1
Log Maximum RCA of Random Neighbors	176,495	-0.31	0.69	-1	3.49
Neighbor (Random) Exports Product RCA>1	176,495	0.30	0.46	0	1

The dataset contains 114 countries, since we exclude countries without neighbors such as islands. We also exclude countries of the former Soviet Union for the periods before 2000. We also exclude all products that were not exported by any country at the beginning of each period. The total number of products in the sample is 777.

Panel A of Table 5 shows the summary statistics for the whole sample. Panel B is for the sample restricted to those observations with $RCA_{c,p,t} \leq 0.1$. These observations include only products "eligible to jump",[13] which we use in this exercise. In the summary statistics (Panel B) it is shown that the unconditional probability of developing a new export in a ten-year period is

in the data. Also, in the first period, all the "previous period" growth variables are computed for 1965-1970. '

[13]Each observation is at the country-product-year level.

1.5%. The table reports statistics for the compound average annual growth rate of export value and RCA for all country-product appearances in the sample for all four periods. In order not to have undefined growth rates we added 0.1 to all RCA values in the sample, thus pairing down the rate of growth for RCA for products below that threshold. However, we do not modify the levels of this variable when used as regressor. We also control for 'distorted' growth rates by adding a dummy when the RCA was zero at the initial year of the computed growth rate used in the right hand side of the specifications. The total number of observations in the sample is above 175,000.

Finally, we develop a counterfactual or surrogate data to have a benchmark with which to test for the significance of our results. In the surrogate dataset we replace a country's real neighbors with an equal number of randomly chosen neighbors. For instance, if South Africa has four neighbors: Botswana, Mozambique, Namibia and Zimbabwe, in our randomization, South Africa will still have four neighbors, which happen to be China, Italy, Malawi and Poland. If the effects we see are not present using our counterfactual sample of random neighbors, then this hints that the creation of new industries is indeed driven by the actual neighbors. The results we see are not present in samples with the same statistical mean and variance of the right-hand-side variables. The summary statistics in Table 5 present the correspondent indicators for the random neighbors as well.

We also present results using a slightly modified version of the empirical specification shown above. In it, we substitute the variable of interest by a dummy that takes the value of one if the RCA of c_N is above 1, and zero otherwise. We do this in order to ease the interpretation of our results. In addition, we also reproduce the results restricting the dataset to NPRB products only.

Table 6 shows the results using the whole universe of products. Panel A uses a continuous independent variable of interest in both its linear and quadratic form, while Panel B uses a binary version. Our variable of interest, the highest RCA (in logs) for the product in the neighborhood, is strongly significant in all specifications. Column 4, which includes all the controls, implies that an increase of one standard deviation in this variable is associated with an average increase in the likelihood of a country "jumping" to that product of 0.55 percentage points, or an increase of 36.7% (based on the 1.5% unconditional probability of jumping). The positive and statistical significant estimator for the quadratic form shows that this is convex.

Panel B of Table 6 has an easier interpretation: if the most successful of your neighbors exporting product p is doing so with an RCA above 1, then your chances of "jumping" to that product increases by 1 percentage point, which represents for the average product an increase of roughly 65% in the probability of "jumping" (from 1.5% to 2.5%). All these estimates are statistically significant.

Both panels replicate the same specification using the counterfactual of an equal number of random neighbors. In all cases, the estimators are sharply

Table 6: Dynamics of Productive Knowledge Diffusion

Dependent Binary Variable: Jump (New Product in 10 years)

Panel A: Continous Independent Variable

	Real Neighbors (1-4)				Random Neighbors (5-8)			
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Max RCA Neighbor	0.011***	0.007***	0.007***	0.008***	-0.000	-0.001*	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log Max RCA Neighbor^2	0.002*	0.003***	0.002**	0.001	0.003***	0.001**	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log Initial RCA		-0.002	-0.022***	-0.022***		0.003	-0.020***	-0.020***
		(0.009)	(0.004)	(0.004)		(0.009)	(0.004)	(0.004)
Initial Density		0.107***	0.100***	0.096***		0.113***	0.108***	0.100***
		(0.011)	(0.009)	(0.018)		(0.011)	(0.009)	(0.018)
Growth Rate RCA		-0.001***	-0.000***	-0.000***		-0.001***	-0.000***	-0.000***
Previous Period		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Zero RCA		0.013***	0.005***	0.003***		0.013***	0.005***	0.003***
Previous Period		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Constant	0.020***	-0.005	0.001	0.002	0.013***	-0.004	0.001	0.001
	(0.001)	(0.009)	(0.001)	(0.001)	(0.001)	(0.009)	(0.001)	(0.001)
Product-by-Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Country-Neighbor-by-Year FE	No	No	No	Yes	No	No	No	Yes
Observations	176,495	165,376	165,376	165,376	176,495	165,376	165,376	165,376
Adjusted R-squared	0.003	0.008	0.007	0.003	0.000	0.007	0.006	0.002

Panel B: Binary Independent Variable

	Real Neighbors (1-4)				Random Neighbors (5-8)			
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neighbor Exports (RCA>1)	0.017***	0.010***	0.010***	0.010***	0.001	-0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Log Initial RCA		-0.002	-0.022***	-0.021***		0.001	-0.020***	-0.020***
		(0.009)	(0.004)	(0.004)		(0.009)	(0.004)	(0.004)
Initial Density		0.107***	0.101***	0.095***		0.115***	0.108***	0.100***
		(0.011)	(0.009)	(0.018)		(0.011)	(0.009)	(0.018)
Growth Rate RCA		-0.001***	-0.000***	-0.000***		-0.001***	-0.000***	-0.000***
Previous Period		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Zero RCA		0.013***	0.005***	0.003***		0.013***	0.005***	0.003***
Previous Period		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Constant	0.012***	-0.009	0.001	0.002	0.015***	-0.004	0.001	0.001
	(0.001)	(0.009)	(0.001)	(0.001)	(0.001)	(0.009)	(0.001)	(0.001)
Product-by-Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Country-Neighbor-by-Year FE	No	No	No	Yes	No	No	No	Yes
Observations	176,495	165,376	165,376	165,376	176,495	165,376	165,376	165,376
Adjusted R-squared	0.003	0.008	0.006	0.003	0.000	0.007	0.006	0.002

Robust standard errors in parentheses. Standard Erros clustered at the Country-Neighbor-Year level.
*** p<0.01, ** p<0.05, * p<0.1

reduced in value and lose their statistical significance. We take this as evidence that the capacity of countries to move into new exports is affected by the presence of those products in the export basket of its neighbors. This is suggestive of intra-industry knowledge diffusion.

Table 7 replicates the same exercise but restricting the dataset to NPRB products only. The results are consistent with the previous table and even stronger, hinting once again that this effect is not driven by geology or climate. Using only NPRB products, according to Panel A of Table 7, an increase of one standard deviation in the maximum log RCA of the neighborhood is associated with an increase in the probability of jumping of 49.25%.[14]

Similarly, Panel B shows the likelihood of "jumping" for the average NPRB product in our sample is enhanced by roughly 65% (from 1.53% to 2.53%) in the presence of a exporter of that product in the neighborhood with an RCA above 1.

We believe our controls rule out interpretations other than product-specific or intra-industry knowledge spillover from a neighbor to a country. Our specification includes country-neighbor-by-year fixed effects, which takes care of all time varying country-neighbor characteristics. We also control for product-by-year fixed effects, which would take account of all prices, demand and technology shocks affecting each product globally. In addition, we control for the own country's export growth (in RCA) of the product

[14]Based on the 1.53% unconditional probability of jumping and a standard deviation of 0.5453 for the variable of interest when the sample is restricted to NPRB products only.

Table 7: Dynamics of Productive Knowledge Diffusion (NPRB products only)

Dependent Binary Variable: Jump (New Product in 10 years) - NPRB Products Only

Panel A: Continous Independent Variable

		Real Neighbors (1-4)			Random Neighbors (5-8)			
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log Max RCA Neighbor	0.017 ***	0.010 ***	0.009 ***	0.010 ***	0.001	0.000	0.001	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Log Max RCA Neighbor^2	0.011 ***	0.012 ***	0.007 ***	0.007 ***	0.012 ***	0.008 ***	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log Initial RCA		-0.052 ***	-0.041 ***	-0.040 ***		-0.050 ***	-0.039 ***	-0.038 ***
		(0.013)	(0.006)	(0.006)		(0.014)	(0.006)	(0.006)
Initial Density		0.209 ***	0.189 ***	0.176 ***		0.209 ***	0.197 ***	0.179 ***
		(0.022)	(0.018)	(0.040)		(0.020)	(0.018)	(0.040)
Growth Rate RCA		-0.000 **	-0.000 **	-0.000 ***		-0.000 **	-0.000 **	-0.000 ***
Previous Period		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Zero RCA		0.013 ***	0.008 ***	0.006 ***		0.013 ***	0.008 ***	0.006 ***
Previous Period		(0.002)	(0.002)	(0.002)		(0.002)	(0.002)	(0.002)
Constant	0.018 ***	-0.063 ***	0.003	0.002	0.009 ***	-0.063 ***	0.003	0.002
	(0.001)	(0.015)	(0.002)	(0.002)	(0.001)	(0.014)	(0.002)	(0.002)
Product-by-Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Country-Neighbor-by-Year FE	No	No	No	Yes	No	No	No	Yes
Observations	92,735	86,938	86,938	86,938	92,735	86,938	86,938	86,938
Adjusted R-squared	0.004	0.018	0.014	0.005	0.002	0.017	0.013	0.003

Panel B: Binary Independent Variable

		Real Neighbors (1-4)			Random Neighbors (5-8)			
Independent Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neighbor Exports (RCA>1)	0.019 ***	0.009 ***	0.009 ***	0.010 ***	-0.001	-0.001	0.001	0.002
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)
Log Initial RCA		-0.058 ***	-0.041 ***	-0.040 ***		-0.055 ***	-0.039 ***	-0.038 ***
		(0.014)	(0.006)	(0.006)		(0.014)	(0.006)	(0.006)
Initial Density		0.207 ***	0.189 ***	0.176 ***		0.215 ***	0.197 ***	0.179 ***
		(0.021)	(0.017)	(0.040)		(0.021)	(0.017)	(0.040)
Growth Rate RCA		-0.000 **	-0.000 **	-0.000 ***		-0.000 **	-0.000 **	-0.000 ***
Previous Period		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Zero RCA		0.014 ***	0.008 ***	0.006 ***		0.014 ***	0.008 ***	0.006 ***
Previous Period		(0.002)	(0.002)	(0.002)		(0.002)	(0.002)	(0.002)
Constant	0.013 ***	-0.068 ***	0.003	0.002	0.016 ***	-0.065 ***	0.003	0.002
	(0.001)	(0.015)	(0.002)	(0.002)	(0.001)	(0.014)	(0.002)	(0.002)
Product-by-Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Country-Neighbor-by-Year FE	No	No	No	Yes	No	No	No	Yes
Observations	92,735	86,938	86,938	86,938	92,735	86,938	86,938	86,938
Adjusted R-squared	0.003	0.017	0.014	0.004	0.000	0.016	0.013	0.003

Robust standard errors in parentheses. Standard Erros clustered at the Country-Neighbor-Year level.
*** p<0.01, ** p<0.05, * p<0.1

in the previous decade in order to consider the possibility that neighboring countries have parallel productivity trends in the same products but with a time delay. We also control for the country's own likelihood of developing the product by using the density variable mentioned above. After all these controls, we are left with a direct effect of the presence of the industry in one country on its neighbors.

4 Patterns of Diffusion: Products' Complexity and Time Trends

A question that remains open is how easily know-how diffuses as a function of the amount of tacit knowledge a product requires. In other words, if a product is more complex or knowledge-intensive, does it diffuse with more difficulty?

Keller and Yeaple (2010) perform a similar exercise by studying the extent to which sales by foreign affiliates of multinational companies are affected by the intensity in R&D of the industry they belong to. They find that the sales of subsidiaries decline with distance from headquarters and that this effect is stronger in more R&D-intensive industries.

We explore instead the impact of the complexity of the product on the likelihood that it will diffuse to neighbors. To do so we interact the largest RCA in the neighborhood of the country (our variable of interest) with the Product Complexity Index (PCI). The PCI is a continuous product-year variable that measures the complexity of a good based on an iterative formula that takes into account the ubiquity of the product (i.e. the number of countries that make it) and the diversification of the countries exporting that product (i.e. the number of other products that they make) (Hidalgo & Hausmann, 2009; Hausmann et. al. 2011). The first iteration will measure a product as more complex if few highly diversified countries make it. The second iteration will look at the other products that are made by the exporters of a given product and look at whether those are also made by few highly diversified countries and so on. A product that requires a lot of tacit knowledge will be hard to make and hence will be made by few countries. However, if these countries have a lot of tacit knowledge, they should be able to use it to make more products and hence should be more diversified. A product with a high PCI requires more complex know-how in order to be produced/exported. In fact, using the Lall categories as a benchmark (see Table 2), the Primary and Resource Based (PRB) products have an average PCI of -0.07 (s.d. 1.68), while NPRB products have an average PCI of 1.72 (s.d. 1.52) in the year 2001. For all products, PCI distributes (in the year 2001) continuously from -4.2127 to 5.2405, with mean 0.91 (s.d. 1.83).

Table 8 shows results with the same sample used in the previous section, while the columns 2 and 3 restrict the sample to the NPRB and PRB products only (respectively). While we still get a positive and significant coefficient for the highest RCA in the neighborhood, this effect seems to be dampened the more complex a product is. These negative and significant coefficients in columns 1 and 2 are precisely capturing that diffusion is less common for knowledge-intensive products. The results in column 3, which uses a sample of only PRB products, show a non-statistically significant estimator for the interaction parameter. This provides evidence that the knowledge intensity for this subset of already less-complex products is less relevant for diffusion. The included quadratic terms to allow for non-linearity are statistically insignificant in all the different samples.

These results are illustrative of the role of tacit knowledge in explaining the slow pace of international technological diffusion. The more complex a product, the weaker will its diffusion be. We interpret this as the consequence of the increased tacit knowledge that more complex products require.

Our next analysis relates to diffusion across time. With global communication and international travel become more common, one would expect that tacit knowledge diffusion, which by definition requires person-to-person interaction, would become stronger. This idea is not new. Keller (2002), for example, shows that international technological diffusion has become less localized in more recent periods. Comin and Hobijn (2010) also show how adoption lags of technologies have been shorter for newer technologies.

We proceed to test this within our framework by interacting our variable of interest with a time trend. We introduce the time trend in both its linear and quadratic forms using the four periods in our sample described above. The results are shown in Table 9. The first two columns are the results using all products in the sample, while the last two columns restrict the sample

Table 8: Diffusion and Product Complexity

Dependent Binary Variable: Jump (New Product in 10 years)

Independent Variables	All Products (1)		NPRB (2)		PRB (3)	
Log Max RCA Neighbor X Product Complexity Index	-0.001	**	-0.005	***	-0.001	
	(0.001)		(0.002)		(0.001)	
(Log Max RCA Neighbor X Product Complexity Index)^2	-0.000		-0.000		-0.000	
	(0.000)		(0.000)		(0.000)	
Log Max RCA Neighbor	0.008	***	0.014	***	0.006	***
	(0.001)		(0.003)		(0.001)	
Log Initial RCA	-0.023	***	-0.043	***	-0.015	***
	(0.004)		(0.006)		(0.004)	
Initial Density	0.098	***	0.182	***	0.063	***
	(0.018)		(0.040)		(0.015)	
Growth Rate RCA	-0.000	***	-0.000	***	-0.000	***
Previous Period	(0.000)		(0.000)		(0.000)	
Zero RCA	0.004	***	0.006	***	-0.000	
Previous Period	(0.001)		(0.002)		(0.001)	
Constant	0.001		0.002		0.001	
	(0.001)		(0.002)		(0.001)	
Observations	165,376		86,938		76,448	
Adjusted R-squared	0.003		0.005		0.003	

All specifications include Product-by-Year fixed effects and Country-Neighbor-by-Year fixed effects.
Robust standard errors in parentheses. Standard Erros clustered at the Country-Neighbor-Year level.
*** p<0.01, ** p<0.05, * p<0.1

only to NPRB products. In the first two columns, when using all the products, we find no evidence of faster diffusion across time. However, we do fin evidence when looking only at NPRB products. Despite the non-significance of the interaction term in column 3, the results in column 4 show that the interaction in both its linear and quadratic forms are jointly significant (p-value=0.0146). The positive value of the quadratic term suggests that the speed of diffusion is increasing over time. The fact that NPRB products are more complex than PRB products suggests that faster diffusion may be affecting predominantly more complex products.

Figure 6 graphs the marginal return to one standard deviation increase in the highest RCA among the neighbors of a country (in logs) on the probability of "jumping" for the average product, using the results of column 4.[15] In it we clearly see the convex shape of this relationship. In fact, in the period 1970-1980 the average return to the probability of "jumping" due to an increase in one standard deviation in the variable of interest is 0.586 percentage points, which corresponds to an increase of 38% (using the 1.53% unconditional probability of jumping, suing our NPRB estimates). Doing the same exercise for our latest period, 2001-2008, we find that the increase is of 0.866 percentage points corresponding to an increase of 56.56%. This evidence suggests that indeed diffusion has become more widespread in the more recent period, which could be the direct result of more human interaction across

[15]The 95% confidence interval of the estimation was computed by bootstrapping using 10000 repetitions.

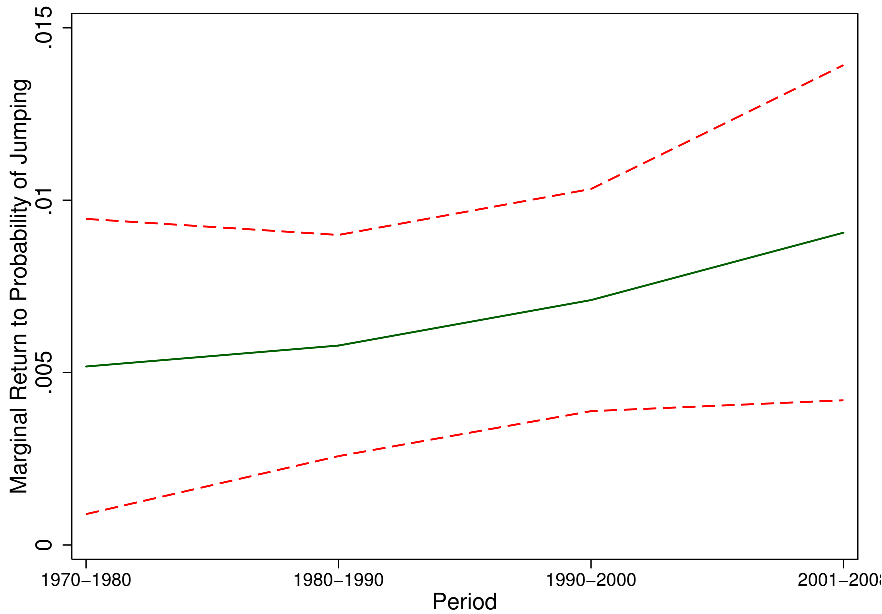
Table 9: Diffusion Across Time

Dependent Binary Variable: Jump (New Product in 10 years)

Independent Variables	All Products (1)		All Products (2)		NPRB (3)		NPRB (4)	
Log Max RCA Neighbor X Period	-0.0002		-0.0003		-0.0008		-0.0006	
	(0.001)		(0.001)		(0.002)		(0.002)	
(Log Max RCA Neighbor X Period)^2			0.0001				0.0005	***
			(0.000)				(0.000)	
Log Max RCA Neighbor	0.009	***	0.009	***	0.011	*	0.011	*
	(0.002)		(0.002)		(0.006)		(0.006)	
Log Initial RCA	-0.022	***	-0.022	***	-0.040	***	-0.040	***
	(0.004)		(0.004)		(0.006)		(0.006)	
Initial Density	0.098	***	0.097	***	0.182	***	0.178	***
	(0.018)		(0.018)		(0.040)		(0.040)	
Growth Rate RCA	-0.000	***	-0.000	***	-0.000	***	-0.000	***
Previous Period	(0.000)		(0.000)		(0.000)		(0.000)	
Zero RCA	0.003	***	0.003	***	0.006	***	0.006	***
Previous Period	(0.001)		(0.001)		(0.002)		(0.002)	
Constant	0.002		0.002		0.003		0.002	
	(0.001)		(0.001)		(0.002)		(0.002)	
Observations	165,376		165,376		86,938		86,938	
Adjusted R-squared	0.003		0.003		0.004		0.005	

All specifications include Product-by-Year fixed effects and Country-Neighbor-by-Year fixed effects.
Robust standard errors in parentheses. Standard Erros clustered at the Country-Neighbor-Year level.
*** p<0.01, ** p<0.05, * p<0.1

Figure 6: Marginal Return to the Probability of "Jumping" over time



countries, proxied here by a country and its neighbors.

5 Concluding Remarks

The economic literature has moved from an emphasis on traditional factors of production to a stress on the role of intermediate inputs in determining productivity. While technology can be partially diffused by embedding it in tradable products, accounting for the large differences in productivity across countries requires an assumption regarding the limited tradability of key intermediate inputs (Rodriguez-Clare, 1996; Rodrik, 1996). A key candidate for this non-tradability is the tacit knowledge that is required to make products.

If tacit knowledge has difficulty in moving over space, then we should expect neighbors to be similar in terms of the products they export. In this respect, we show that there is a very strong similarity in the export basket of neighboring countries. We find this surprising because, as predicted by gravity models, neighbors trade more intensely, and this should specialize countries in different rather than similar products, in order to exploit the gains from trade. The fact that neighbors tend to export the same products, even when primary and resource-based products are excluded, is suggestive of localized intra-industry technological spillovers.

The phenomenon is more convincingly shown by looking at the dynamic introduction of new products into the export basket of countries. To the best of our knowledge, we are the first ones to use the appearance of products in the export basket of countries as a measure of intra-industry international technological diffusion. In our view, it signals that a country has acquired the required tacit knowledge related to that particular product. Our results using this outcome variable are in line and consistent with previous economic literature regarding the short radius of diffusion patterns. These results had been previously shown using patent citations and productivity trends.

We believe that we convincingly show intra-industry international diffusion by estimating the impact of having a neighbor that successfully exports a product on the likelihood that a country will be able to introduce that product into its export basket. Our findings control for country-neighbor-by-year and product-by-year fixed effects, indicating that the result is not driven by time varying country-pair or product characteristics. We also control for the country's own predisposition to move into that product and for its previous history in that product. We further check that the effect comes not from the number of neighbors by creating a surrogate dataset in which we substitute a country's neighbors with an equal number of random neighbors.

In addition, we find that (i) more complex products, i.e. those that require more knowledge, diffuse less; and (ii) diffusion has accelerated over time, hinting that bridges between people in different countries have increased, as global communication develop, allowing for faster diffusion. All these results, are consistent with previous literature and provide evidence of the limited tradability of key ingredients of production, such as tacit knowledge.

The limited tradability of tacit knowledge can help explain the well-known fact that rich and poor countries tend to be geographically clustered. In the context of this paper, this may be related to the fact that countries are affected by the tacit knowledge that exists in their neighborhood. Countries in a neighborhood more richly endowed with tacit knowledge will be able to develop the capacity to export more and more complex goods. Previous research has shown the robust link between the diversification and complexity of a country's export basket and its future growth (Hausmann, Hwang and Rodrik, 2007; Hidalgo and Hausmann, 2009, Hausmann et. al. 2011).

Future research should explore the precise channel through which these spillovers happen. In particular, it would be interesting to know the relative importance of trade, foreign direct investment, migration or other more tacit channels in the intensity of international intra-industry spillovers. These are important questions, because as this paper has shown, the geography of knowledge diffusion is strong enough to have shaped the comparative advantage of nations.

References

Acharya, R.C., and W. Keller. "Technology transfer through imports." *Canadian Journal of Economics* 42, 4: (2009) 1411–1448.

Aghion, Philippe, and Peter Howitt. *The Economics of Growth.* Cambridge, MA: MIT Press, 2009.

Arrow, Kenneth J. "Classificatory Notes on the Production and Transmission of Technologcal Knowledge." *The American Economic Review* 59, 2: (1969) 29–35.

Balassa, B. "Trade Liberalisation and Revealed Comparative Advantage." *The Manchester School* 33, 2: (1965) 99–123.

Bottazzi, Laura, and Giovanni Peri. "Innovation and spillovers in regions : Evidence from European patent data." *European Economic Review* 47: (2003) 687–710.

Branstetter, Lee G. “Are knowledge spillovers international or intranational in scope ? Microeconometric evidence from the U.S. and Japan.” *Journal of International Economics* 53: (2001) 53–79.

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

Coe, David T, and Elhanan Helpman. “International R&D spillovers.” *European Economic Review* 2921, 94.

Coe, David T., Elhanan Helpman, and Alexander W. Hoffmaister. “International R&D spillovers and institutions.” *European Economic Review* 53, 7: (2009) 723–741.

Comin, Diego, and Bart Hobijn. “An Exploration of Technology Diffusion.” *American Economic Review* 100, 5: (2010) 2031–2059.

Dornbusch, R, and S Fischer. “Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods.” *The American Economic Review* 67, 5: (1977) 823–839.

Eaton, Jonathan, and Samuel Kortum. “Technology, trade, and growth : A unified framework.” *European Economic Review* 45, 4-6: (2001) 742–755.

Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo. “World Trade Flows: 1962-2000.” *NBER Working Paper No. 11040* 1962–2000.

Finger, JM, and ME Kreinin. "A Measure of 'Export Similarity' and Its Possible Uses." *The Economic Journal* 89, 356: (1979) 905–912.

Grossman, Gene M., and Elhanan Helpman. *Innovation and Growth in the Global Economy.* Cambridge, MA: MIT Press, 1991.

Hausmann, Ricardo, César A Hidalgo, Sebastián Bustos, Michele Coscia, Sarah Chung, Juan Jímenez, Alexander Simoes, and Muhammed A. Yildirim. *The Atlas of Economic Complexity: Mapping Paths to Prosperity.* Cambridge, MA, 2011.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. "What you export matters." *Journal of Economic Growth* 12, 1: (2006) 1–25.

Hausmann, Ricardo, and Bailey Klinger. "The Structure of the Product Space and the Evolution of Comparative Advantage." *CID Working Paper Series* , 146.

Heckscher, Eli, and Bertil Ohlin. *Heckscher-Ohlin Trade Theory.* Cambridge: MIT Press, 1991.

Helpman, Elhanan, and Paul Krugman. *Market Structure and International Trade.* Cambridge: MIT Press, 1985.

Hidalgo, César A, and Ricardo Hausmann. "The building blocks of economic complexity." *Proceedings of the National Academy of Sciences of the United States of America* 106, 26: (2009) 10,570–5.

Hidalgo, César A, Bailey Klinger, AL Barabási, and Ricardo Hausmann. "The product space conditions the development of nations." *Science (New York, N.Y.)* 317, 5837: (2007) 482–7.

Jaffe, A.B., M. Trajtenberg, and R. Henderson. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *The Quarterly Journal of Economics* 108, 3: (1993) 577.

Keller, Wolfgang. "Geographic localization of international technology diffusion." *American Economic Review* 92, 1: (2002) 120–142.

———. "International Technology Diffusion." *Journal of Economic Literature* XLII, September: (2004) 752–782.

Keller, Wolfgang, and Stephen R Yeaple. "Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States." *The Review of Economics and Statistics* 91, November: (2009) 821–831.

———. "Gravity in the Weightless Economy." .

Krugman, Paul. *Geography and Trade*, volume 1 of *MIT Press Books*. Cambridge, MA: The MIT Press, 1992.

La Porta, R, F Lopez-de Silanes, A Shleifer, and R Vishny. "The quality of government." *The Journal of Law, Economics & Organization* 15, 1: (1999) 222–279.

Lall, Sanjaya. “The Technological Structure and Performance of Developing Country Manufactured Exports, 1985-98.” *Oxford Development Studies* 28, 3: (2000) 337–369.

Larrain, F., L.F. Lopez-Calva, and A. Rodriguez-Clare. “Intel : A Case Study of Foreign Direct Investment in Central America.” *Center for International Development Working Paper* 58.

Leamer, EE, and James Levinsohn. “International Trade Theory: The Evidence.” *Handbook of international economics* III.

Linder, Staffan Burenstam. *An essay on trade and transformation.* Stockholm: Almqvist & Wicksell, 1961.

Maddison, A. *Monitoring the world economy, 1820-1992.* Etudes du Centre de d{é}veloppement. Development Centre of the Organisation for Economic Co-operation and Development, 1995.

Markusen, James R, and Anthony J Venables. “Foreign direct investment as a catalyst for industrial development.” *European Economic Review* 43: (1999) 335–356.

Mayer, Thierry, and Soledad Zignago. “Notes on CEPII distances measures : The GeoDist database.” *CEPII Working Paper* , 25.

Polanyi, M. *Personal knowledge: Towards a post-critical philosophy.* London, UK: Routledge, 1962.

Pritchett, Lant. "Divergence, big time." *The Journal of Economic Perspectives* 11, 3: (1997) 3–17.

Ricardo, David. *On the Principles of Political Economy and Taxation*. London, 1817, 1951 edition.

Rivera-Batiz, L.A., and P.M. Romer. "Economic integration and endogenous growth." *The Quarterly Journal of Economics* 106, 2: (1990) 531–555.

Rodríguez-Clare, Andrés. "The division of labor and economic development." *Journal of Development Economics* 49, 1: (1996) 3–32.

Rodrik, Dani. "Coordination failures and government policy: A model with applications to East Asia and Eastern Europe." *Journal of International Economics* 40, 1-2: (1996) 1–22.

Romer, P. "Endogenous Technological Change." *The Journal of Political Economy* 98, 5: (1991) S71–S102.

Shirotori, M, B Tumurchudur, and O Cadot. "Revealed Factor Intensity Indices at the Product Level." *Policy Issues in International Trade and Commodities* , 44.

Tinbergen, J. "Shaping the world economy." *The International Executive* 5, 1: (1963) 27–30.

United Nations. "COMTRADE database.", 2010. `http://comtrade.un.org/`.

World Bank. "World Development Indicators Online.", 2010. http://data.worldbank.org/.

Zipf, George Kingsley. "The P1 P2/D Hypothesis: On the Intercity Movement of Persons." *American Sociological Review* 11, 6: (1946) pp. 677–686.

A Appendix

A.1 Robustness of the Stylized Facts

We replicate the results shown in the main body of the paper using the Finger & Kreinin (F&K) Export Similarity Index (Finger & Kreinin, 1979). F&K Similarity Index is constructed using the formula:

$$S_{c,c'}^{F\&K} = \sum_{p} min(X_p^c, X_p^{c'})$$

where p represents products, c and c' represent any two countries and x_p^c is the share of product p exported by country c out of the total export baskets for country c. Hence, two countries c and c' that export the exact same products in the exact same proportion would have $S_{c,c'}^{F\&K} = 1$.

Figure A1 shows the scatter of both export similarity indices – our own named BBH Export Similarity Index and F&K's one – showing a strong positive correlation between them (ρ = 0.69), implying that both indexes capture much of the same information.

Figure A2 replicates Figure 1 in the main body of the paper using the F&K Similarity Index. Once again, on average, export similarity is higher among neighbors than for all the pairs of countries together, and it is also decreasing with distance.

Table A1 replicates Table 3 in the main body of the paper. It estimates the adapted gravity model using the F&K Export Similarity Index as the

Figure A1: Correlation Between Similarity Indices Measures (Year 2000)

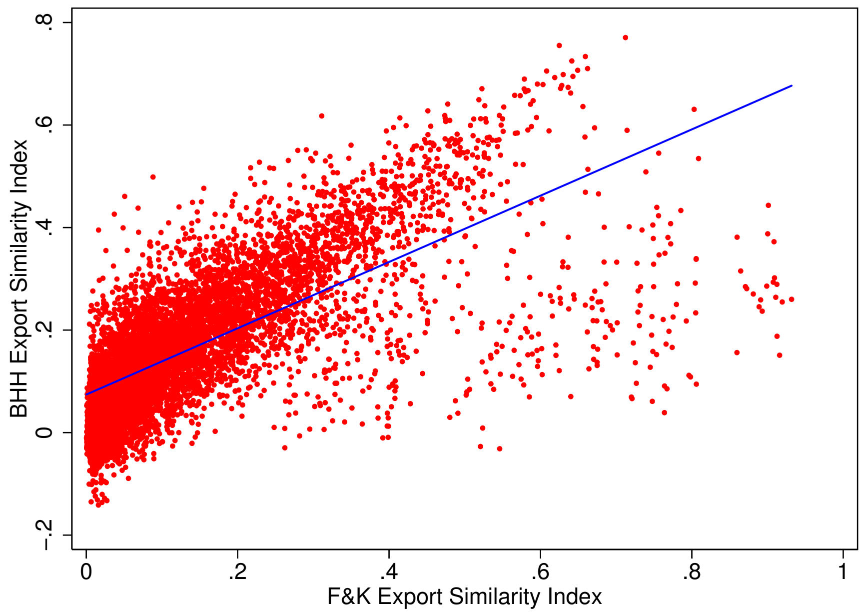
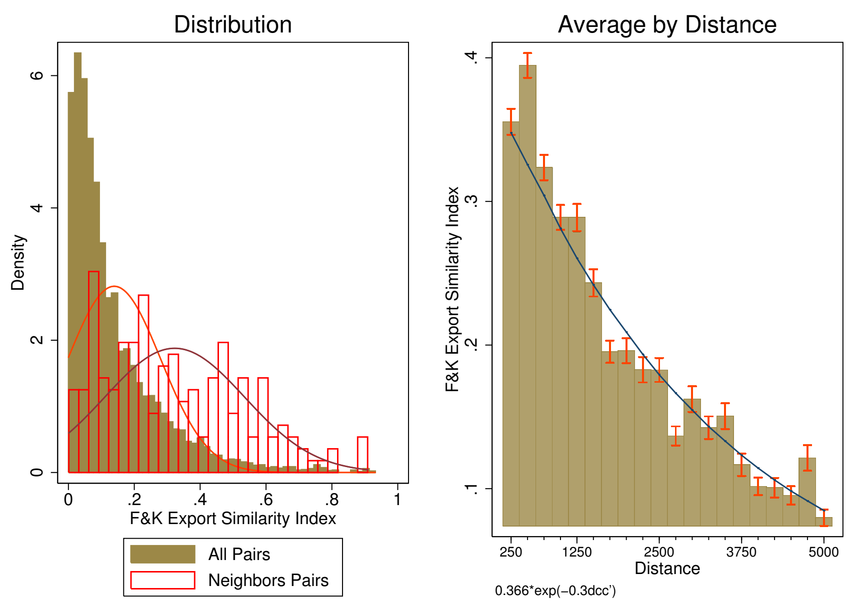


Figure A2: F&K Export Similarity Index (Year 2000)



dependent variable. The results are consistent with the ones using our own index.

Despite the strong correlation between our own *Export Similarity Index* and the one suggested by F&K, we pursue all of our analysis with ours because, as explained in the body of the paper, the F&K measure does not distinguish between products that are exported by one country and not the other from those that are exported by neither. This is corrected by our index, which takes values lower than zero when countries are more different than similar. We believe this to be an important feature of an export similarity index, since it allows studying specialization between countries.

A.2 Bilateral Trade does not Drive Similarity in Exports

A regression supporting the finding presented in Figure 4 is in Table A2. We run a linear regression based on specification (3) measuring the power of explanation of both the decompositions of the Similarity Index on $S_{c,c'}$ for year 2000. The results are presented in Table A2. Here we show that the Export Similarity Index measured with bilateral exports between country pairs does not explain, under any specification, the overall Similarity Index: its estimator is small, statistically insignificant and has the wrong sign. However, the similarity index as measured by exports to the rest of the world has the holds a strong explanatory power in the regression.

Table A1: Correlates of the Export Similarity Index (Year 2008)

Dependent Variable: F&K Export Similarity Index	**(1)**		**(2)**		**(3)**		**(4)**		**(5)**		**(6)**	
Log Distance	-0.027	***	-0.027	***	-0.026	***	-0.021	***	-0.017	***	-0.022	***
	(0.003)		(0.003)		(0.003)		(0.004)		(0.004)		(0.004)	
Share Border	0.089	***	0.090	***	0.077	***	0.073	***	0.047	**	0.049	***
	(0.015)		(0.015)		(0.015)		(0.016)		(0.018)		(0.018)	
Same Region	0.044	***	0.042	***	0.025	***	0.028	***	0.033	***	0.035	***
	(0.006)		(0.006)		(0.006)		(0.008)		(0.010)		(0.010)	
Common Language			0.007		0.017	**	0.026	***	0.023	***	0.025	***
			(0.007)		(0.007)		(0.008)		(0.009)		(0.009)	
Colonial Relationship			-0.022	*	-0.020		-0.026	**	0.001		0.005	
			(0.013)		(0.012)		(0.013)		(0.013)		(0.012)	
Common Colonizer			0.004		-0.001		0.003		0.010		0.014	
			(0.008)		(0.008)		(0.009)		(0.013)		(0.013)	
Absolute Difference					-0.032	***	-0.025	***	-0.020	**	-0.022	***
Ln GDPPC					(0.002)		(0.005)		(0.008)		(0.008)	
Absolute Difference					-0.006	***	-0.008	***	-0.003		-0.003	
Ln Pop					(0.002)		(0.002)		(0.002)		(0.002)	
Absolute Difference							-0.010	**	-0.015	*	-0.014	*
Ln Physical Capital Per Worker							(0.005)		(0.008)		(0.008)	
Absolute Difference							0.001		-0.001		-0.001	
Ln Human Capital Per Worker							(0.001)		(0.001)		(0.001)	
Absolute Difference							-0.006		-0.010	*	-0.010	*
Ln Land Per Worker							(0.005)		(0.005)		(0.005)	
Log Bilateral Trade											-0.011	***
											(0.004)	
Constant	0.444	***	0.443	***	0.542	***	0.505	***	0.167	***	0.241	***
	(0.035)		(0.036)		(0.041)		(0.057)		(0.039)		(0.042)	
Observations	7,875		7,875		7,626		5,671		3,175		3,175	
Adjusted R-squared	0.293		0.293		0.326		0.378		0.465		0.467	

All specifications include country dummies.
Robust standard errors in parentheses, clustered at the country-pair level.
*** p<0.01, ** p<0.05, * p<0.1

Table A2: Neighbors Similarity (on bilateral exports vis-à-vis ROW exports)

Dependent Variable: Export Similarity Index	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Export Similarity Index	-0.008	-0.008	-0.011	-0.007	-0.009	-0.009
(measured with Bilateral Exports)	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)
Export Similarity Index	0.870 ***	0.870 ***	0.854 ***	0.846 ***	0.848 ***	0.848 ***
(measured with Exports to ROW)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Log Distance	-0.000	0.000	-0.000	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Share Border	-0.007 *	-0.007	-0.005	-0.007 *	-0.008 **	-0.008 *
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Same Region	0.011 ***	0.010 ***	0.008 ***	0.007 ***	0.006 ***	0.006 ***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Common Language		0.006 **	0.007 ***	0.009 ***	0.009 ***	0.009 ***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Colonial Relationship		-0.011 **	-0.010 **	-0.009	-0.006	-0.006
		(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Common Colonizer		0.015 ***	0.014 ***	0.023 ***	0.022 ***	0.022 ***
		(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Absolute Difference			-0.007 ***	-0.003	-0.004 *	-0.004 *
Ln GDPPC			(0.001)	(0.002)	(0.002)	(0.002)
Absolute Difference			-0.002 ***	-0.002 ***	-0.001 **	-0.001 **
Ln Pop			(0.001)	(0.001)	(0.001)	(0.001)
Absolute Difference				-0.001	-0.001	-0.001
Ln Physical Capital Per Worker				(0.002)	(0.002)	(0.002)
Absolute Difference				0.001 ***	0.001 ***	0.001 ***
Ln Human Capital Per Worker				(0.000)	(0.000)	(0.000)
Absolute Difference				0.004 ***	0.004 ***	0.004 ***
Ln Land Per Worker				(0.001)	(0.001)	(0.001)
Log Bilateral Trade						-0.000
						(0.001)
Constant	0.152 ***	0.151 ***	-0.034 *	-0.051 **	-0.049 **	-0.049 **
	(0.014)	(0.016)	(0.018)	(0.020)	(0.020)	(0.021)
Observations	3,149	3,149	3,083	2,551	2435	2435
Adjusted R-squared	0.973	0.973	0.974	0.977	0.977	0.977

All specifications include country dummies.
Robust standard errors in parentheses, clustered at the country-pair level.
*** p<0.01, ** p<0.05, * p<0.1

A.3 Decomposing Similarity

The observed similarity through the network in Figure 5 is based on the correlated export of resource-based products for some country-pairs (blue links) and by non-resource-based products for others (red links).

We created a measure to determine whether for a pair of countries' similarity is a reflection of the export of primary and resource based (PRB) products or, on the contrary, non primary nor resource based (NPRB) products. The measure is based on decomposing the relative contribution of PRB and NPRB products to export similarity by separating products into these two categories and counting the fraction of PRB and NPRB products that both countries export with an RCA above their respective means. We take the difference between these two fractions as an estimate of the contribution of PRB and NPRB products to export similarity. Formally, we define:

$$\Delta\sigma_{c,c'} = \sigma_{c,c'}^{NPRB} - \sigma_{c,c'}^{PRB} \quad (5)$$

where

$$\sigma_{c,c'}^{NPRB} = \frac{1}{N_{NPRB}} \sum_{p \epsilon NPRB} \delta_{c,c',p} \quad (6)$$

and N_{NPRB} is the total number of NPRB products and

$$\delta_{c,c',p} = \begin{cases} 1 & \text{if } RCA_{c,p} \geq \overline{RCA_c} \text{ and } RCA_{c',p} \geq \overline{RCA_{c'}} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $\overline{RCA_c}$ is the average RCA of country c over all products.

The definition for $\sigma^{PRB}_{c,c'}$ can be obtained by changing NPRB for PRB in (6).

From equation (5), $\Delta\sigma_{c,c'} > 0$ if the major contributors to the export similarity between c and c' are NPRB products, such as manufactures and chemicals, and negative in the opposite case. As an example, Figure A3 plots Japan's and Korea's RCA in all products in 2008 and shows NRBP products in red and PRB products in blue. The horizontal flat lines represents the average RCA over all products for Korea, while the vertical flat line does so for Japan. In this case $\sigma^{NPRB}_{c,c'c} = 0.6517$, $\sigma^{PRB}_{c,c'} = 0.3471$ and $\Delta\sigma_{c,c'} = 0.3046$, indicating that Japan and Korea export 61.75% of all of their NPRB products with an RCA above their respective means (in the upper right part of the graph), compared to only 34.71% for PRB products. This shows that the similarity between Japan and Korea that we are measuring comes mainly from their correlated export of NPRB products.

By using these measures we are able to document for any pair of countries whether their exports similarity is driven by NPRB or by PRB products. Not all countries similarity is driven by the same kind of products. Figure (A4) summarizes this information by showing, within each region of the World,

Figure A3: Decomposition of Similarity Index for Korea and Japan in 2008

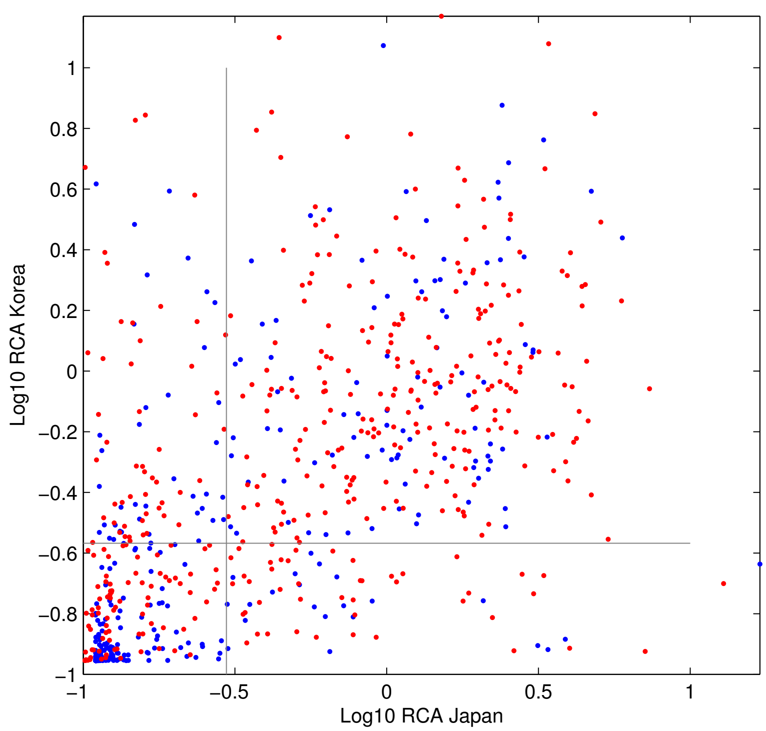
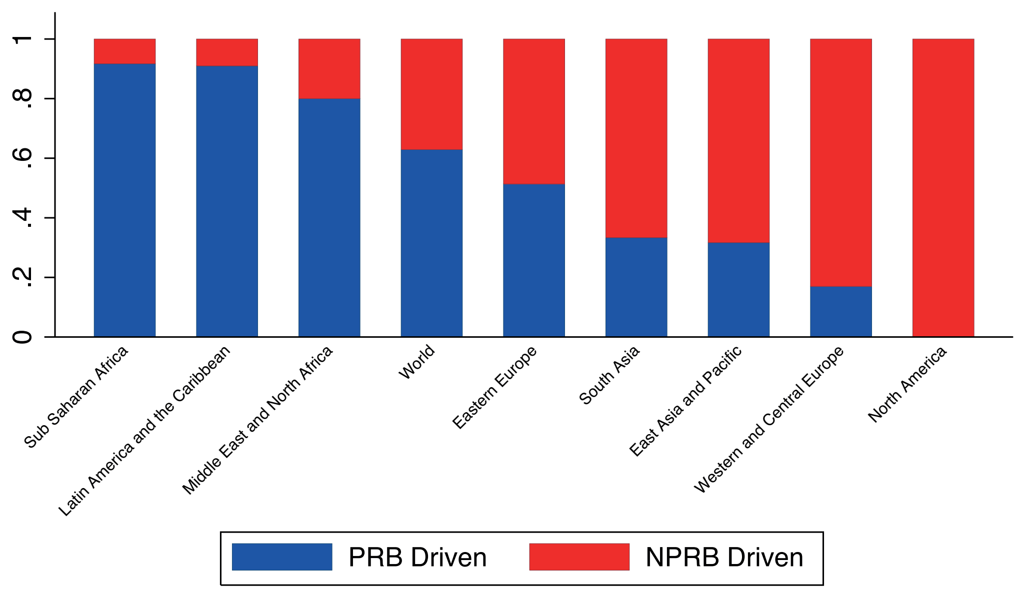


Figure A4: Category of Products Driving Similarities per Region (Year 2008)



what proportion of country-pairs are similar mostly due to NPRB products, or PRB products.