

# Migration, knowledge diffusion and the comparative advantage of nations\*

Dany Bahar (Harvard University)<sup>†</sup>

Hillel Rapoport (Paris School of Economics and Bar Ilan University)

[Download Newest Version]

**Keywords:** migration, knowledge, comparative advantage, exports

**JEL Classification Numbers:** F14, F22, F62, O33, D83

---

\*This paper benefited from helpful comments at various stages from Martin Abel, Sam Asher, Sebastian Bustos, Michele Coscia, Ricardo Hausmann, Elhanan Helpman, Juan Ariel Jimenez, Bill Kerr, Michael Kremer, Frank Nefke, Paul Novosad, Nathan Nunn, Rodrigo Wagner, Muhammed Yildirim, Andres Zahler and the participants of numerous seminars at Harvard Economics Department, Harvard Kennedy School, the Hebrew University of Jerusalem and Bar Ilan University. All errors are my own.

<sup>†</sup>Corresponding Author: dbahar@fas.harvard.edu

## **Abstract**

To what extent are migrants a source of evolution of the comparative advantage of both their sending and receiving countries? We study the drivers of knowledge diffusion by looking at the dynamics of the export basket of countries. The main finding is that migration is a strong and robust driver of productive knowledge diffusion. In terms of their ability to induce exports, we find that an increase of only 65,000 people in the stock of migrants for the average country, is associated with about 15% increase in the likelihood of adding a new product to a country's export basket. We also find that, in terms of expanding the export basket of countries, a migrant is worth about US \$30,000 of foreign direct investment. For skilled migrants these same figures become 15,000 people and US \$160,000. In order to alleviate endogeneity concerns, we present results based on instrumenting for migration stocks using bilateral geographic and cultural variables.

# 1 Introduction

Franschhoek valley, a small town in the Western Cape province of South Africa, is known today for its beautiful scenery and for its high-quality wineries. The town was founded in the late 17th century by French Huguenot refugees, who settled there after being expelled from France following King Louis XIV elimination of the Edict of Nantes. As of today, the wineries in Franschhoek are among the main producers of South African wine exports. Likewise, Saxenian (2006) relates the story of Dov Frohman, an Israeli scientist who in 1974 returned home after years of having worked in Intel Corporation in the United States. Upon his return to Israel, Frohman founded in Haifa Intel's first design center outside the United States. As of today, Israel is an exporter of semiconductors related technologies. In this paper we explore the role of migrants in developing the comparative advantage of both their sending and receiving countries.

Ricardian models of trade assume as given the exogenous productivity parameters that define the export basket of countries which are generated in equilibrium. A large part of the literature has focused on understanding the characteristics of this equilibrium and the mechanisms through which it is conceived. However, a burgeoning literature has been dealing with understanding the evolution of what defines these productivity parameters. This paper contributes to this literature by documenting industry-specific productivity shifts as explained by the variation in international factors movement

with particular focus on migration. We study productivity by exploiting changes in the export baskets of countries. The key assumption is that, after controlling for product-specific shifts in demand, firms in a country will be able to export a good only after they have become productive enough to compete in global markets. Of all international factors flows, the results point to migration as the strongest of those drivers. We find that migrants, and even more so, skilled migrants, can explain variation in good-specific productivity as measured by the ability of countries to export those goods, for products that are intensively exported in the migrants' home/destination countries. In particular we find that, on average, a stock of migrants larger by 65,000 people is associated with a 15% increase in the likelihood of exporting a new product for a given country, whereas the same figure for skilled migrants is reduced to 15,000. Also, in terms of expanding the export basket of countries, a migrant is worth about US \$30,000 of foreign direct investment (FDI), while a skilled migrant is worth over \$160,000.

This differs from the previous approaches in the literature that look at the link between international factor flows and changes in aggregate productivity, as opposed to industry-specific productivity dynamics. That literature includes, for instance, the work of Coe and Helpman (1993) and Coe et. al. (2009) who study changes in aggregate productivity as a result of importing more from countries with higher R&D investment. Aitken & Harrison (1999) and Javorcik (2004) are among the long list of studies trying to establish whether FDI generates productivity spillovers on domestic firms, with

no definite answer emerging from all of them. Andersen and Dalgaard (2011) find a correlation between aggregate productivity and business travel flows.

We consider three alternative explanations on how migration could be associated to good-specific productivity increases. First, if a given country  $c$  receives migrants from countries exporters of a given product  $p$ , then there could be a local shift in demand for product  $p$ , given the plausible shift in aggregate preferences. This could result in a demand-driven productivity shift, which could become exports to either the migrants' sending/receiving country or global exports to the rest of the world, supplying the increase in global demand<sup>1</sup>.

Second, migrant networks could generate lower transaction costs for bilateral trade in specific goods, thus inducing bilateral exports between the sending and receiving country of the migrants (Kugler and Rapoport 2011 ).

Finally, migrants can serve as a transmission vehicle of product-specific knowledge, which could induce productivity shifts and in turn inducing global competition in certain goods. The acquisition of industry-specific knowledge is an important input for the productivity dynamics of a firm: more knowledge (either through learning or experience) allows economic agents to do more with the same resources. Then, the question remains: can migrants induce exports through knowledge transmission? Bahar et. al. (2014) present evidence suggesting that, after controlling for product-specific global

---

<sup>1</sup>Linder (1961) suggests, in this case, country  $c$  will become a trade partner of the home countries of the migrants.

demand, the evolution of the export basket of a country, both in its extensive and intensive margin, could be explained by the documented local geographic character of knowledge diffusion (e.g. Jaffe, Trajtenberg and Henderson 1993; Bottazzi and Peri 2003; Keller 2002; Keller 2004). They attribute their result to the fact that knowledge is often non-easily transferrable, mostly because a large component of it is "tacit" (Polanyi 1966). Polanyi explained tacit knowledge by saying that we know more than we can tell, and Kenneth Arrow (1969) suggested that the drivers for knowledge transmission are human minds rather than written words, also when it comes to economic processes. Thus, intuitively, migration would be the natural candidate among all international factor flows to serve as a driver of tacit knowledge and thus induce exports. This is, precisely, what this paper documents.

The paper uses a worldwide dataset that includes bilateral trade, FDI and migration stock figures for years 1990 and 2000. From it, we construct a sample that includes for each country, product and year the total exports to the rest of the world. The sample also includes the computed total stocks of trade, FDI and migration (disaggregated in immigrants and emigrants) to or from partner countries that export that each product in years 1990 and 2000.

The undertaken empirical exercise looks at how migration figures correlate with a country's extensive and intensive margin of trade. The extensive margin is measured by looking at the future addition of a new product to a country's export basket, while the intensive margin refers to the future annual

growth rate of a product that is already exported by a country. We control for global demand of each good by adding product-year fixed effects. We also add country-year fixed effects which would control for all country level variables characteristics that would make a given country more likely to export and receive migrants at the same time. We also calculate all of the specifications using an alteration of the dependent variable, which measures exports to the rest of the world excluding flows to countries where migrants are in or from. These controls, we argue, would rule out the first two explanations discussed above.

We are left with the third explanation, which is our preferred one. Yet, endogeneity concerns are present. They are reduced by adding proper controls and, also, by presenting a set of results that instrument for migration stocks using geographic and cultural bilateral variables between countries the sending and receiving countries of the migrants. The instruments provide an exogenous variation to the number of migrants in/from countries. They are based on the share of the migrants' sending and receiving countries exporting the product under consideration which have a common language, a common colonizer or a (former) colony-colonizer relationship with the country under analysis.

The body of the paper discusses in detail all the data collection, the empirical strategies and present the results. The paper is divided as follows. The next section describes the empirical strategy and the data. Section 3 presents the main results, and Section 4 discusses them. Section 5 concludes.

## 2 Empirical Strategy

### 2.1 Research Question and Empirical Challenges

The empirical strategy studies the relationship between international factor flows and the dynamics in the export basket of the receiving and sending countries, with emphasis on migration. In particular, the question is: can migrants induce product-specific productivity shifts in their sending (destination) countries, on products already intensively exported in their destination (sending) countries?

For the sake of better understanding, we use the following hypothetical example. Suppose there are two countries in the world: Italy (a pizza exporter), and the US (a hamburger exporter). The analogous question then becomes whether the presence of more Italians in the US is associated with the ability of the US to export pizza, and, whether this same presence is also associated with the ability of Italians to export hamburgers.

There are two main empirical challenges in studying the relationship between productivity and international factor flows (i.e. goods, capital and people). First, all flows are highly correlated among themselves. Moreover, several empirical studies have shown that migration networks are an important determinant of bilateral trade flows and bilateral FDI.<sup>2</sup>

Kugler and Rapoport (2011) even find evidence of complementarity be-

---

<sup>2</sup>e.g. Gould, 1994; Rauch and Trindade, 2002; Combes, Lafourcade and Mayer, 2005; Iranzo and Peri 2009; Felbermayr and Jung, 2009; Tong, 2005; Kugler and Rapoport, 2007, 2011; Javorcik et. al. 2011



tween the three types of flows. More specifically, the authors claim, migration generate links that lower transaction costs inducing bilateral FDI and trade.

Hence, the positive correlation between international flows of capital, goods and labor is a matter of consideration to any study of this kind. In fact, in the sample for year 2000, the correlation matrices between total migration, FDI and trade across countries are all positive, and above 0.4, with the exception of migration and FDI per capita (see Tables 1 and 2). That is, countries that receive/send more FDI tend to also receive/send more migrants and export/import in larger quantities. Hence, to deal with this challenge, the empirical specification controls for all three factors simultaneously.

[Table 1 about here.]

[Table 2 about here.]

The second challenge refers to the risk of having biased estimates generated by endogeneity, even after controlling for all factor flows. For instance, migrants could relocate themselves based on future potential of specific sectors of growing, or a third variable (i.e. “openness” shock) could induce migration and induce exports at the same time (though not necessarily of specific products). While the controls in the empirical specification intends to deal with some of these problems, the identification problem remains. In order to further reduce these concerns, we implement a number of instrumental variables, which exploit the variation in migration stocks explained

by bilateral cultural/historic characteristics between the sending and receiving countries of these migrants.

## 2.2 Empirical Specification

The aim of the paper is to study the dynamics of the extensive and intensive margin of trade (with exports to the rest of the world) given different levels of migration stocks, controlling for FDI and trade stocks. The specification will disentangle between immigration and emigration, and between all vs. skilled migrants.

Throughout the paper we will use the concept of Revealed Comparative Advantage (RCA) by Balassa (1965), which will be used to construct export-related variables both in the left-hand-side and right-hand-side of the specification. RCA is defined as follows:

$$RCA_{c,p} \equiv \frac{exp_{c,p} / \sum_p exp_{c,p}}{\sum_c exp_{c,p} / \sum_c \sum_p exp_{c,p}}$$

where  $exp_{c,p}$  is the exported value of product  $p$  by country  $c$ . This is a yearly measure.

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.

The empirical specification is defined as follows:

$$\begin{aligned}
Y_{c,p,t \rightarrow T} &= \beta_{im} \sum_{c'} immigrants_{c,c',t} \times R_{c',p,t} + \beta_{em} \sum_{c'} emigrants_{c,c',t} \times R_{c',p,t} \\
&+ \beta_{FDI} \sum_{c'} FDI_{c,c',t} \times R_{c',p,t} + \beta_{trade} \sum_{c'} trade_{c,c',t} \times R_{c',p,t} \\
&+ \gamma Controls_{c,p,t} + \alpha_{c,t} + \eta_{p,t} + \varepsilon_{c,p,t}
\end{aligned} \tag{1}$$

The definition of the dependent, or left hand side (LHS) variable,  $Y_{c,p,t \rightarrow T}$ , alternates according to whether the specification is studying the intensive or the extensive margin of trade for a specific product  $p$  and country  $c$ . When studying the *extensive* margin,  $Y_{c,p,t \rightarrow T}$  is 1 if country  $c$  achieved an *RCA* of 1 or more in product  $p$  in the period of time between  $t$  and  $T$  (conditional on having an  $RCA_{c,p,t} = 0$ ). That is:

$$Y_{c,p,t \rightarrow T} = 1 \text{ if } RCA_{c,p,t} = 0 \text{ and } RCA_{c,p,T} \geq 1$$

In addition, we also condition  $Y_{c,p,t \rightarrow T} = 1$  to whether whether  $RCA_{c,p,t-1} = 0$ , where  $t - 1$  refers to the beginning of the previous period (i.e. for the period 2000-2010, it is 1990)<sup>3</sup>. This eliminates cases in which the country was already exporting such product but for some reason they stopped doing so at the beginning of the period under consideration.

When studying the *intensive margin*,  $Y_{c,p,t \rightarrow T}$  is the annual (log) growth rate in the exports value of product  $p$ , between years  $t$  and  $T$ , conditional on having  $exports_{c,p,t} > 0$ . That is:

---

<sup>3</sup>For the period 1990-2000,  $t - 1$  is 1985 given data limitations

$$Y_{c,p,t \rightarrow T} = \frac{\ln(exports_{c,p,T} + 1) - \ln(exports_{c,p,t} + 1)}{T - t} \text{ if } exports_{c,p,t} > 0$$

The independent variables include the following:

- The sum of the stock of immigrants and of emigrants (in logs) from and to other countries (denoted by  $c'$ ) at time  $t$ , weighted by a dummy  $R_{c',p,t}$  which is 1 if  $RCA_{c',p,t} \geq 1$ <sup>4</sup>.
- The sum of stock of FDI and stock of trade (in logs) according to the previous logic.
- A vector of controls of baseline variables (when applicable): the baseline level of exports (in logs) for that same product; the average annual (logarithmic) growth rate of the export value in the previous ten year period (in order to control for previous trends in the export dynamics for that product); in order to correct for undefined growth rates caused by zeros in the denominator, we compute the growth rate using  $exports_{c,p,t} + 1$  for all observations; in addition, to control for our own correction, we also add as a control a dummy variable indicating whether  $exports_{c,p,t} = 0$ , which correspond to the observations most likely to be distorted.
- Country-by-year and product-by-year fixed effects.

---

<sup>4</sup>In fact, we define several thresholds for  $R_{c',p,t} = 1$ , which are  $RCA_{c',p,t} \geq x$  where  $x \in [1, 5]$

## 2.3 Data and Sample

Bilateral migration data comes from Docquier, Ozden and Perry (2010). The dataset consists of total bilateral working age (25 to 65 years old) foreign born individuals in 1990 and 2000. The data provide figures for skilled and non-skilled migrants at the bilateral level as well. Skilled migrants are considered to have completed some tertiary education at the time of the census. Figures 1 and 2 represent the migration data in year 2000.

[Figure 1 about here.]

[Figure 2 about here.]

Bilateral FDI positions are from the OECD International Direct Investment Statistics (2012). It tracks FDI from and to OECD members since 1985 until 2009. Using this data we compute 10-year stocks of capital flows for each country in 1990 and 2000<sup>5</sup>. Negative FDI stocks are treated as zero<sup>6</sup>.

Bilateral trade data comes from Hausmann et. al. 2011, based on the UN Comtrade data from 1984 to 2010. We use the trade dataset to construct two variables. First, total exports per product per country to the rest of the world, to be used to compute the dependent variable in the empirical specifications, to study the intensive and extensive margin of trade. The list of products is fairly disaggregated. An example of a product is

---

<sup>5</sup>For 1990 we use the stock from 1985 to 1990 due to limitations of the data.

<sup>6</sup>This follows the same methodology suggested by Kugler and Rapoport (2011)[Kugler and Rapoport(2011)]. Only 1.7% of the original dataset is affected by this.

"Knitted/Crocheted Fabrics Elastic Or Rubberized" (SITC code 6553), or "Electrical Measuring, Checking, Analyzing Instruments" (SITC code 8748). The words product, good and industry interchangeably referring to the same concept throughout the paper. Similarly to FDI, we also compute 10-year stocks for bilateral trade (imports plus exports) to be used in the RHS. Both the 10-year Trade and FDI stocks are deflated using the US GDP deflator (base year 2000) from the World Development Indicators by the World Bank.

Former Soviet Union countries are excluded from the sample given their poor trade data in the period 1990-2000. The final sample consists of 136 countries and 781 products. We define two 10-year periods for the analysis due to the limitations imposed by the bilateral migration data, which are 1990-2000 and 2000-2010.

Furthermore, we incorporate variables from the GeoDist dataset from CEPII on bilateral relationships such distance, common colonizer, colony-colonizer relationship, and common language, to construct the instrumental variables (Mayer and Zignago 2011).

The summary statistics for the variables to be used in the analysis are in Table 3. Panel A presents the summary statistics for the extensive margin sample, while Panel B does it for the intensive margin sample. The left and right sides of both panels present the same variables used in the RHS varying the threshold  $R_{c',p,t} = 1$ , based on  $RCA_{c',p,t} \geq 1$  and  $RCA_{c',p,t} \geq 5$ , respectively. The variables that do not depend on this threshold are reported in the left panel only.

[Table 3 about here.]

From Panel A we see that the unconditional probability of achieving an RCA above 1 (starting with an none exports) for the average country-product is 3.5%. Similarly, from Panel B , the average country-product exports value annual growth rate is close to zero in the data. The tables also include the sum of immigrants and emigrants for the average country and year from and in countries exporting a product with RCA above 1 (left) and above 5 (right). It presents the same statistics for aggregated FDI and Trade figures in million USD, after the delation process explained above. Note that FDI and Trade variables total inwards and outwards stock figures.

### 3 Results

Table 4 presents the estimation of specification (1) for all products in the dataset. The upper panel estimates the extensive margin (measured by the likelihood of adding a new product to a country's export basket) while the lower panel estimates the intensive margin (measured by the annual growth in exports of a product already in the country's export basket). It is important to notice that the dependent variables in both panels are computed using exports from country  $c$  to product  $p$  to the rest of the world. The columns titled "R1" indicate that the threshold used for constructing the RHS variables was  $RCA_{c',p,t} \geq 1$ , whereas the columns titled "R5" indicate the use of  $RCA_{c',p,t} \geq 5$ , instead. The table also presents results using all

migrants (columns 1 and 3) and skilled migrants (columns 2 and 4) in the RHS. The upper panel of Table 4 uses country-product pairs which had zero exports in the baseline years (1990 and 2000), which corresponds to almost 84,000 observations (thus, baseline variables are not included because lack of variation).

[Table 4 about here.]

The results in Panel A indicate that a country with 10% increase in its stock of immigrants from nations exporters of product  $p$ , is associated with an increase of 0.4% in the likelihood the receiving country will export product  $p$  with an RCA above 1 in the next ten years.<sup>7</sup> Similarly, a 10% increase in the stock of emigrants residing in countries exporters of product  $p$ , will tend to increase the same likelihood for country  $c$  by 1%. On average in the sample these numbers amount to about 1,400 immigrants and 5,000 emigrants. The correspondent correlation for an increase of 10% in the total stock of (incoming and outgoing) FDI is about 0.4%<sup>8</sup>. That is, a 3.5 times increase in the stock of FDI, which is about USD 210 million based on the sample average, would account for the same marginal effect of a stock increase of about 6500 migrants (both immigrants plus emigrants). In other words, each migrant is worth about \$30,000 of FDI in this context.

---

<sup>7</sup>A 10% increase is associated with an increase in such likelihood of 0.015 percentage points, which based on the unconditional probability of 3.5% (see Table 3), corresponds to a 0.4% increase.

<sup>8</sup>Or even zero, given that the estimator for the FDI coefficient is statistically insignificant.



Column 2 in Table 4 limits the migration figures to skilled immigrants and emigrants only. This significantly reduces the variation in the RHS, as can be seen in the summary statistics. The results suggest that, in the case of immigration, the coefficient estimator is almost doubled and statistically significant. This implies that the probability of a country adding a new product  $p$  to its export basket is larger by 0.8% for each 10% increase its stock of immigrants from countries exporters of  $p$  (about 200 people, on average). In terms of skilled emigration, a 10% increase in the stock of emigrants in nations exporters of product  $p$  (about 1300 people), is associated with an increase of 0.8% in the likelihood a given country will start exporting  $p$  in the next 10 years. In total, the figures suggest that, on average, an increase of 1500 in the stock of migrants in and from countries with comparative advantage in  $p$ , is associated with a probability of exporting  $p$  larger by 1.6%. To achieve the same result with FDI would require a fourfold increase in FDI, which amounts to about USD 240 million. Using the same logic as above, this means that each skilled migrant is worth about USD \$160,000 in terms of expanding a country's export basket in the ways studied by this paper.<sup>9</sup>

Panel B of Table 4 uses the annual export value (logarithmic) growth rate as the dependent variable, in order to study the intensive margin of trade. The number of observations is different than the sample used for

---

<sup>9</sup>The comparison of migrants with trade figures is not possible given that the estimated coefficients for trade are negative. This is a natural result given that countries tend to trade less with other countries that export the same goods.

Panel A, because we are using all products with export value above zero in the baseline year. The results present evidence that both the presence of immigrants from and of emigrants in countries exporters of product  $p$ , is associated with a larger future rate of growth in export value of product  $p$  in the country under consideration. In particular, for a given product  $p$ , a 10% increase in the stock of immigrants from countries exporting such product (about 17,000 people) is associated with an increase in the future annual growth rate in export value for the receiving country of about 0.03 points. Similarly, future annual growth in exports value of product  $p$  tends to be 0.07 percentage points higher with a 10% larger stock of emigrants in countries exporters of the same product (about 19,000 people). That is, about 40,000 more migrants, on average, is associated with a annual growth rate that is larger by 0.1 points. The coefficients in Column 2, which use skilled migration in the right hand side, are larger in magnitude in the case of immigrants, though the coefficient for emigration becomes statistically insignificant for the case of emigrants.

Columns 3 and 4 repeat the exercise of Columns 1 and 2 in both panels, but redefining  $R_{c',p,t} = 1$  if  $RCA_{c',p,t} \geq 5$ . This means that the right hand side variables are weighted by whether the partner countries have a revealed comparative advantage that is 5 times the world average in product  $p$  at time  $t$ . Intuitively, the average migrant from and/or in these countries will have a higher likelihood of being exposed to the productive knowledge that is required to efficiently produce (and thus, being able to export) product  $p$ .

At the same time, the variation for all variables is considerably less.

For this case, the results show no significant correlation with immigrants. However, the results suggest that a 10% increase in the stock of emigrants, which amounts to about 660 people, is associated with a larger probability of achieving comparative advantage by 0.5%. The same figure for skilled emigrants, which corresponds to about 100 people on average, is also 0.5%. That is, less than 1000 emigrants, on average, in countries that export product  $p$  with an RCA above 5, could increase the likelihood of their sending countries achieving comparative advantage in the same product by 1%. In this case, the comparison with FDI suggest that each emigrant is worth USD 800,000, or USD 900,000 of trade!

An interesting implication of the results is that FDI and Trade figures, in most cases, seem not to correlate with the ability of countries to expand their the export baskets under the studied context. That is, trading with countries which are exporters of a particular product is not associated with the likelihood of gaining comparative advantage in that same product. However, when it comes to the intensive margin of trade seem to positively correlate with the future annual growth of export value. Precedents of this result tracks to Coe and Helpman (1995), where they find evidence on how trade leads to increases in aggregate productivity.

Based on valid concerns on how much of these results are being driven by good pinned down by geology or climate conditions, we estimate specification (1) excluding those products from the sample. The results are robust to their

exclusion. See Section A.2.2 in the Appendix for more details.

All the specifications presented above include product-by-year fixed effects and country-by-year fixed effects. The former set of fixed effects would control for global demand for all products. Given that we are looking at exports to the rest of the world, the shifts we identify must be related to the supply side. The country-by-year fixed effects would control for time invariant countries' characteristics, as well as country-level aggregate demand and supply shocks, that would rule out a country-level third factor that positively correlates with both migrant figures and overall productivity, such as an openness shock.

### **3.1 Increased productivity or lower bilateral trade costs?**

A valid concern would be that the partial correlations we are observing are being driven by bilateral trade: the country is exporting more of the product to those countries where the migrants are from/in. This relates to the evidence presented by Kugler and Rapoport (2011), who find that migrants facilitate the creation of business networks which induces bilateral trade and capital flows. Under this possibility, it would be harder to attribute the results to a gain in productivity, but to a decrease in trade costs. In order to deal with this we estimate again the same specification, but we exclude from the dependent variable all exports to countries where migrants are from/in. That is, we reconstruct the dataset such that the export value to the rest of the world for each product and country combination, excludes exports to

nations that send or receive that same country's migrants.

Of course, a critical caveat is that the exclusion requires defining a threshold on the number of sending/receiving migrants. If one migrant is enough to activate this rule we will probably clean all world trade, given that there is always one alien citizen of every country in most developed nations, which generate the largest share of world trade. In this sense, we define a number of arbitrary thresholds which are 500, 1000, 2500 and 5000 migrants. For example, let's suppose we are looking at Canadian exports of television sets to the rest of the world in year 1990. We will exclude from that figure exports of TV sets from Canada to countries that (1) have a number  $X$  of Canadians migrants and (2) a number  $Y$  of their citizens are migrants in Canada, as long as  $X+Y$  is larger than 500, 1000, 2500 and 5000. The assumption is that in order to create an effective business network one would need more than 500, 1000, 2500 or 5000 migrants among the two countries.

In fact, Figure 3 shows the magnitude of the reduction of total trade figures after revising the exports figures as explained above. For instance, with the 500 threshold world trade figures are reduced by about 92.5%; while using the 5000 threshold reduces total trade figures by about 83%.

[Figure 3 about here.]

Nevertheless, despite the strong decline in the variation of the dependent variable, the results show consistent patterns with the previous results. For instance, Table 5 shows results using the 500 threshold (the most conserva-

tive one), while the tables using the other thresholds are presented in the Appendix Section A.2.2. The results are qualitatively the same as in Table 4. However, all in all, the observed correlation between migrant stocks and increased exports is not explained by migration-induced lower trade costs, but rather good-specific productivity increases. In fact, for Panel A, the estimates are similar in magnitude than in Table 4. In Panel B, the results are less robust, but the correlation with emigrants in countries with an RCA above 5 is consistent with the previous results.

[Table 5 about here.]

In spite of having shown that the results are not driven by bilateral migrant networks there is still room for endogeneity concerns, which keeps us from concluding anything causal on the relationships we have found so far. The next subsection deals with this issue and attempts to solve some of these concerns.

### **3.2 Instrumental variables approach**

The documented correlations can be partly driven by endogeneity: migrants relocate themselves following potential growth in products they are familiar with. In order to reduce endogeneity concerns, we generate a number of instrumental variables that will serve as exogenous variation to migration figures.

Given that there are two regressors we are interested in (immigrants and emigrants), an instrumental variable approach would require at least two instruments. In the spirit of Frankel & Romer (1999) we construct a number of instruments based on bilateral characteristics of the country under consideration and the countries of its immigrants/emigrants, using data from CEPII's GeoDist database.

For each combination  $c, p$  (with  $RCA > 1$ ) and  $t$  there is a defined set of countries  $c'$  where immigrants are *from* and emigrants are *in* (i.e.  $\Theta^I$  and  $\Theta^E$  respectively). Based on this, there are three pairs of instruments that are constructed. They are the *share* of countries  $\Theta^I$  and  $\Theta^E$  that:

- speak a common language as  $c$
- have the same current or historic colonizer as  $c$
- have/had a colony-colonizer relationship with  $c$

For the instruments to be valid, the exclusion restriction must be that, *product specific* exports to the the whole world are not correlated with common bilateral cultural/historic ties with its migrants' countries, once we control for country-year fixed effect. Furthermore, we assume that countries do not engage into *product specific* export-inducing agreements based on their cultural or historical ties, which *are not* captured via flows such as FDI or trade.

The assumptions above are not testable, and their validity -as any other exclusion restriction- are not fully guaranteed. However, following previous

literature, we believe the assumptions are reasonable, and present the results using these instruments.

The relevance of the instrument is fully testable. For intuition purposes, figure 4 presents the analogous of a first stage in a one endogenous variable 2SLS regression, using the United States and South Africa as examples. In both the left and right panels, each observation is a product labeled with its SITC 4-digit code. The left panel uses the United States in year 2000 as an example. The vertical axis measures the number of immigrants (in logs) it has received from countries exporters of each product, while the horizontal axis measures what share of those countries speak english. The right panel uses South Africa in year 2000, and measures the same relationships through emigrants, rather than immigrants. In both panels we see a strong positive correlation, which represents what would be a first stage.

[Figure 4 about here.]

However, since the specification includes  $n > 1$  endogenous regressors, testing for the relevance of the instruments is not straightforward. Stock and Yogo (2002) define critical values on the a number of cases involving  $n > 1$  endogenous regressors, to be used with the Kleibergen-Paap F statistic (when not assuming homoskedasticity). The critical value is 15.72 for the case of 2 endogenous and 6 instruments. Being above this value implies the bias caused by a weak instrument is lower than the OLS bias with 95% certainty. The Kleibergen-Paap F statistic will be reported in all regressions.



Results using the instrumental variables through the generalized methods of moments are presented in Table 6.

[Table 6 about here.]

Note that in all specifications in Panel B, the Kleibergen-Paap F statistics reveal the weakness of the instruments, so there is little we can say about the intensive margin of trade according to these results. In Panel A, however, the instrument seems to have a strong enough first stage (besides in Column 2), which allow us to draw conclusions.

First, immigration seems not to be correlated with the ability of countries to add new goods to their export baskets. However, when it comes to emigration, the magnitude of the coefficients are much larger, as compared to the ones in Table 4. For instance, based on Columns 1, a 10% increase in the stock of emigrants (about 5000 people) increases the likelihood of exporting a product exported in the receiving countries of these migrants by 3% (or 0.11 percentage points). Column 2 shows that the same increase in the stock of skilled emigrants (about 1300 people) could increase by 4% (or 0.15 percentage points) the likelihood of including a new good to the export basket. Again, we see how skilled emigrants have a much stronger correlation.

Columns 3 and 4 show a similar picture. While the magnitude of the estimators are smaller, the variance in the number of emigrants using the  $RCA \geq 5$  threshold is also lower. Thus, the effect is much larger when measuring according to the actual number of emigrants that could boost the

export basket of their sending countries. That is, only about 700 emigrants and 100 skilled emigrants residing in countries that export a particular good with an RCA that is five times the world average, would increase the likelihood of exporting that same good by 1.5% or 2.2%, respectively.

If the exclusion restrictions presented before are valid, and the results cannot be attributed to a third uncontrolled for variable, then these results are particularly strong and a solid contribution to the literature. The presence of migrants from or in nations that export a particular good induce a productivity shift in the sending and receiving country of the migrants, which results in the diversification of their export baskets.

## 4 Discussion and Interpretation

The results in the previous section show through a variety of ways that migration, in both directions, is a determinant of the evolution of the comparative advantage of nations. What is the logic behind that?

If knowledge is tacit, and thus it requires human interaction for its transmission and diffusion, then we could expect that migrants are a driver of such process, which results in increased productivity of the particular sectors that are especially productive in the sending and receiving countries of the migrants.<sup>10</sup> The results are consistent with such hypothesis.

The idea that immigrants could bring productive knowledge is obvious.

---

<sup>10</sup>Section A.1 of the Appendix outlines a simple model that formalizes this idea.

They are physically present in their receiving country, and thus they interact with the local population in ways that could lead to the diffusion of knowledge. But, why would we see a similar pattern with emigrants? There are two possible mechanisms in place: return migration or communication with their family and friends back home.

Return migration seems like a very plausible one. After all, estimates show that about 30% of emigrants return to their home countries after some period of time (e.g. Borjas and Bratsberg, 1996). These migrants spend enough time in the foreign country to be part of the labor force, which eventually could lead to generate industry-specific productivity shifts back home. More recently, Choudhury (2014) shows how Indian return migrants induce productivity improvements in their firm back home, after spending time in the multinational corporation headquarters abroad.

The second mechanism relates to links and open communication channels between the emigrants and their co-nationals back home. Thus, industry-specific knowledge diffusion are more prevalent whenever these links between individuals across nations are open: more communication, more short-term travel, etc. The identification of the exact mechanism, however, is part of our future research agenda.

## 5 Concluding Remarks

This paper presents evidence suggesting that migrants are a source of evolution for the comparative advantage of nations; a relationship that has not been documented in the literature thus far. The results contribute to the growing literature that aims to explain the evolution of industry-specific productivity of countries, and to the literature of international trade that aims to understand, in a Ricardian framework, dynamics of the comparative advantage of nations. It also contributes to the literature of international knowledge diffusion by studying the possible drivers of knowledge across borders, using the setting suggested by Bahar et. al. 2014, which uses product-level exports figures as a measure of knowledge acquisition, after controlling for global demand.

The main result in all these settings is that people matter: by serving as international drivers of productive-knowledge, they can shape the comparative advantage of nations. In all of the specifications we include controls for a set of variables that leave us with empirical evidence suggestive that this is the mechanism in place. The instrumental variables approach also reduces possible endogeneity concerns.

This finding is particularly important to understand some known characteristics of knowledge diffusion. First, the short-ranged character of knowledge diffusion can be explained by the fact that part of knowledge is embedded in people, that tend to move in a more localized manner than goods or

capital. Second, the fact that diffusion of knowledge and technology is more widespread today than decades ago (i.e. the diffusion process has accelerated over time) can be explained by the fact that people flows, such as migration or short term travel, have also increased rapidly.

The results suggest a process that is observable. In fact, the sample contains examples that are consistent with anecdotal evidence. For instance, the Tanzanian soap industry which benefited from Kenyan migration<sup>11</sup>; as well as paper products exports from Chile which occurred simultaneously to the presence of Chileans in Sweden during Chile's military dictatorship.

The importance of these results, however, go beyond the pure relationship between migration and exports. It serves as a proof of concept that mobility is a crucial element in the evolution of the comparative advantage of countries and productivity, which is known to be highly explanatory of income and growth.

## References

Aitken, B.J., and A.E. Harrison. "Do domestic firms benefit from direct foreign investment? Evidence from Venezuela." *American Economic Review* 89, 3: (1999) 605–618.

Andersen, Thomas Barnebeck, and Carl-Johan Dalgaard. "Flows of people, flows of ideas, and the inequality of nations." *Journal of Economic Growth*

---

<sup>11</sup>[http://www.unitedworld-usa.com/usatoday/tanzania/interviews/hemal\\_shah.htm](http://www.unitedworld-usa.com/usatoday/tanzania/interviews/hemal_shah.htm)

16, 1: (2011) 1–32.

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

Bahar, Dany, Ricardo Hausmann, and Cesar A. Hidalgo. “Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion?” *Journal of International Economics* 92, 1: (2014) 111–123.

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

Borjas, George J., and Bernt Bratsberg. “Who Leaves? The Outmigration of the Foreign-Born.” *The Review of Economics and Statistics* 78, 1: (1996) 165.

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

Choudhury, Prithwiraj. “Return Migration and Geography of Innovation in Mnes: A Natural Experiment of On-the-Job Learning of Knowledge Production by Local Workers Reporting to.” *Harvard Business School Working Paper* , 14-078.

Coe, David T, and Elhanan Helpman. “International R&D spillovers.” *European Economic Review* 2921, 94: (1995) 859–887.

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

Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer. “The trade-creating effects of business and social networks: evidence from France.” *Journal of International Economics* 66, 1: (2005) 1–29.

Docquier, Frédéric, Abdeslam Marfouk, C. Ozden, and C.R. Parsons. “Geographic, gender and skill structure of international migration.” In *Report written for the Economic Research Forum*. 2010, 1–27.

Felbermayr, Gabriel J., and Benjamin Jung. “The pro-trade effect of the brain drain: Sorting out confounding factors.” *Economics Letters* 104, 2: (2009) 72–75.

Frankel, Jeffrey, and David Romer. “Does Trade Cause Growth?” *American Economic Review* 89, 3: (1999) 379–399.

Gould, DM. “Immigrant links to the home country: empirical implications for US bilateral trade flows.” *The Review of Economics and Statistics* 76, 2: (1994) 302–316.

Hausmann, Ricardo, César A Hidalgo, Sebastián Bustos, Michele Coscia, Sarah Chung, Juan Jiménez, Alexander Simoes, and Muhammed A.

Yildirim. *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. Cambridge, MA, 2011.

Iranzo, Susana, and Giovanni Peri. “Migration and Trade: Theory with an Application to the Eastern-Western European Integration.” *Journal of International Economics* 79: (2009) 1–19.

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.

Javorcik, Beata S. “Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages.” *The American Economic Review* 94, 3: (2004) 605–627.

Javorcik, Beata S., Caglar Ozden, Mariana Spatareanu, and Cristina Neagu. “Migrant networks and foreign direct investment.” *Journal of Development Economics* 94, 2: (2011) 231–241.

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.

Kugler, Maurice, and Hillel Rapoport. “International labor and capital flows: Complements or substitutes?” *Economics Letters* 94, 2: (2007) 155–162.



———. “Migration, FDI, and the Margins of Trade.” *CID Working Paper Series* , 222.

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

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

OECD. “International Direct Investment Statistics.”, 2012. <http://www.oecd/investment/statistics>.

Polanyi, M. *The Tacit Dimension*. Chicago; London: University of Chicago Press, 1966, 2009 edition.

Rauch, James E, and Vitor Trindade. “Ethnic chinese networks in international trade.” 84, February: (2002) 116–130.

Saxenian, AnnaLee. *The New Argonauts: Regional Advantage in a Global Economy*. Cambridge, MA: Harvard University Press, 2006.

Stock, JH, and M Yogo. “Testing for Weak Instruments in Linear IV Regression.” *NBER Technical Working Papers Series* , 284.

The World Bank. “World Development Indicators Online.”, 2013.

Tong, SY. “Ethnic Networks in FDI and the Impact of Institutional Development.” *Review of Development Economics* 9, 4: (2005) 563–580.

Figure 1: Cartogram Share of Migrants, Year 2000

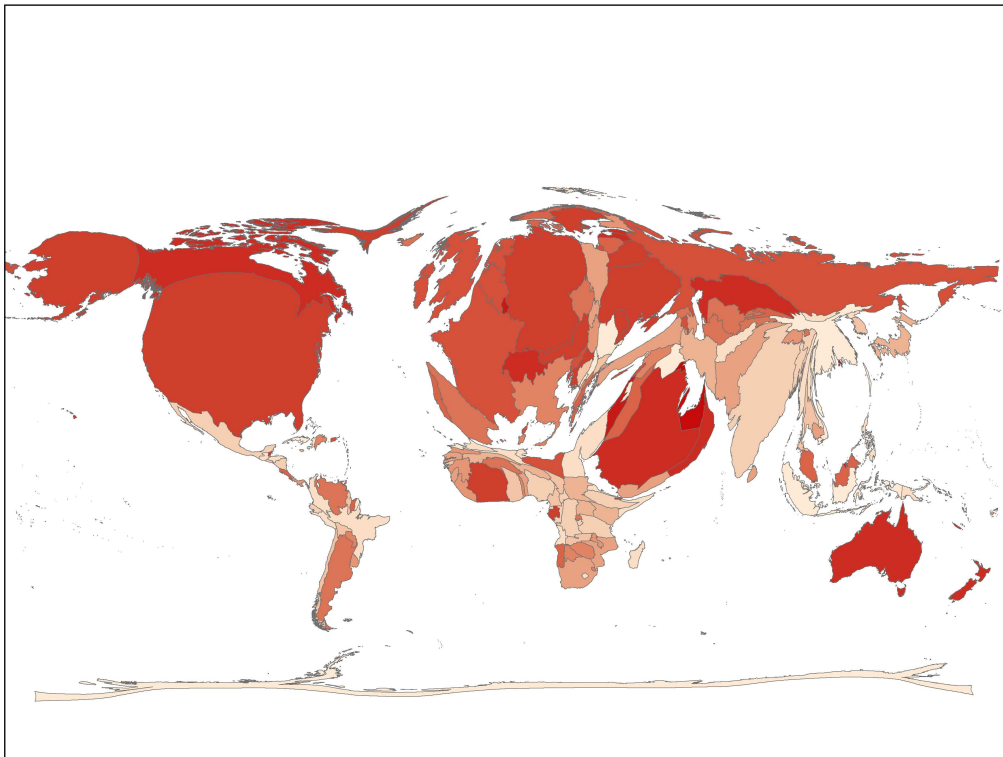


Figure 2: Cartogram of Migrants Per Capita, Year 2000



Figure 3: Proportion of World Trade Left

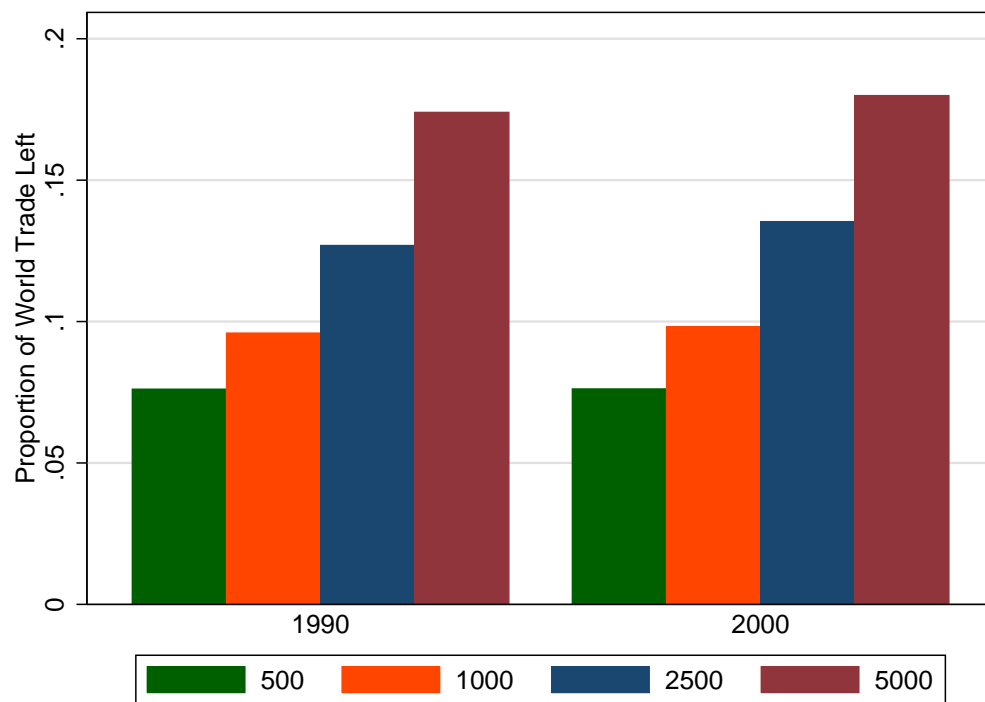


Figure 4: First stage, common language

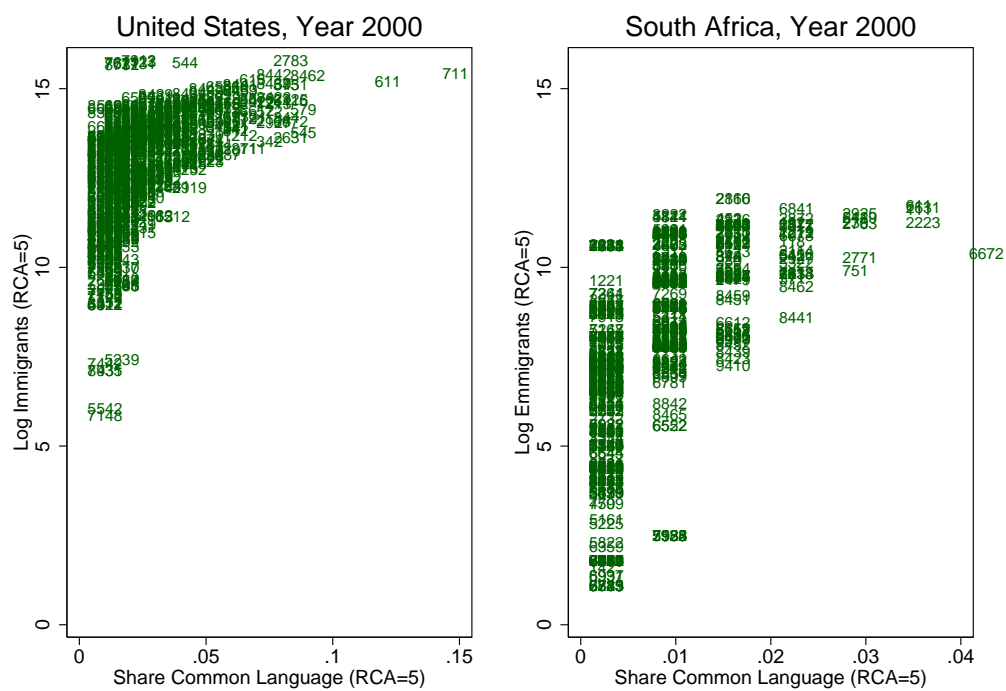


Table 1: Correlation Matrix International Flows (log)

Variables	Migrants (log)	FDI (log)	Trade (log)
Migrants (log)	1.000		
FDI (log)	0.455	1.000	
Trade (log)	0.457	0.738	1.000

Table 2: Correlation Matrix International Flows (per capita)

Variables	Migrants (p.c.)	FDI (p.c.)	Trade (p.c.)
Migrants (p.c.)	1.000		
FDI (p.c.)	0.159	1.000	
Trade (p.c.)	0.423	0.538	1.000

Table 3: Summary Statistics

Variable	RCA=1			RCA=5		
	N	Mean	sd	N	Mean	sd
<i>Panel A: Extensive Margin Sample (<math>exports_{c,p,t} = 0</math>)</i>						
New Product (RCA>1)	83,397	0.035	0.185	-	-	-
Immigrants	83,397	14233.76	69694.14	83,397	3869.30	32469.86
Emigrants	83,397	49819.20	139403.55	83,397	6657.21	57416.82
Immigrants (HS)	83,397	1862.31	9711.80	83,397	450.21	4396.31
Emigrants (HS)	83,397	12818.14	36970.25	83,397	987.11	5668.17
FDI (total, mn USD)	83,397	59.26	1767.27	83,397	8.12	482.81
Trade (total, mn USD)	83,397	1256.47	4014.39	83,397	129.22	733.90
<i>Panel B: Intensive Margin Sample (<math>exports_{c,p,t} &gt; 0</math>)</i>						
Growth Exports	129,035	-0.003	0.349	-	-	-
Baseline Log Exports	129,035	13.259	3.723	-	-	-
Immigrants	129,035	169327.61	582922.35	129,035	32345.20	165795.17
Emigrants	129,035	189433.86	445525.94	129,035	16373.44	90889.80
Immigrants (HS)	129,035	48736.68	220836.55	129,035	7964.71	53017.41
Emigrants (HS)	129,035	55520.94	120329.89	129,035	3422.09	18897.87
FDI (total, mn USD)	129,035	14387.69	60995.89	129,035	644.94	6656.45
Trade (total, mn USD)	129,035	29353.97	67416.95	129,035	2152.64	9568.08



Table 4: Fixed Effects

<b>Panel A: Extensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0015 (0.000)***	0.0028 (0.001)***	0.0005 (0.000)	0.0011 (0.001)
Ln Emigrants	0.0032 (0.001)***	0.0026 (0.001)**	0.0019 (0.001)***	0.0018 (0.001)***
Ln FDI, total	0.0015 (0.001)	0.0014 (0.001)	0.0001 (0.000)	0.0001 (0.001)
Ln Trade, total	-0.0080 (0.004)*	-0.0084 (0.004)*	0.0014 (0.001)	0.0014 (0.001)
N	83397	83397	83397	83397
r2	0.15	0.15	0.15	0.15
<b>Panel B: Intensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0030 (0.001)***	0.0043 (0.001)***	0.0029 (0.001)***	0.0021 (0.001)**
Ln Emigrants	0.0069 (0.002)***	-0.0006 (0.002)	0.0042 (0.001)***	0.0043 (0.001)***
Ln FDI, total	-0.0015 (0.000)***	-0.0013 (0.000)***	-0.0002 (0.000)	-0.0002 (0.000)
Ln Trade, total	0.0094 (0.005)*	0.0165 (0.005)***	0.0012 (0.001)*	0.0021 (0.001)***
Baseline Log Exports	-0.0298 (0.001)***	-0.0295 (0.001)***	-0.0290 (0.001)***	-0.0285 (0.001)***
Previous Exports Growth	-0.1373 (0.013)***	-0.1379 (0.013)***	-0.1397 (0.013)***	-0.1413 (0.013)***
Previous Exports Zero	-0.0044 (0.013)	-0.0042 (0.013)	-0.0007 (0.013)	0.0001 (0.013)
N	73193	73193	73193	73193
r2	0.35	0.35	0.35	0.35

All specifications include country-by-year and product-by-year fixed effects. SE clustered at the country level presented in parenthesis

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Table 5: Fixed Effects, excluding bilateral exports (500 migrants threshold)

<b>Panel A: Extensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0019 (0.000)***	0.0030 (0.001)***	0.0008 (0.000)**	0.0008 (0.001)
Ln Emigrants	0.0032 (0.001)***	0.0023 (0.001)**	0.0012 (0.000)***	0.0008 (0.000)*
Ln FDI, total	0.0008 (0.001)	0.0008 (0.001)	-0.0001 (0.000)	-0.0001 (0.000)
Ln Trade, total	-0.0063 (0.004)	-0.0063 (0.004)	0.0028 (0.001)**	0.0029 (0.001)**
N	119783	119783	119783	119783
r2	0.11	0.11	0.11	0.11
<b>Panel B: Intensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	-0.0007 (0.001)	0.0019 (0.002)	0.0016 (0.001)*	0.0012 (0.001)
Ln Emigrants	0.0084 (0.002)***	0.0012 (0.003)	0.0028 (0.001)***	0.0024 (0.001)*
Ln FDI, total	-0.0009 (0.001)	-0.0008 (0.001)	0.0010 (0.000)**	0.0010 (0.000)***
Ln Trade, total	0.0097 (0.006)	0.0147 (0.006)**	0.0021 (0.001)**	0.0027 (0.001)***
Baseline Log Exports	-0.0038 (0.002)*	-0.0037 (0.002)*	-0.0029 (0.002)	-0.0026 (0.002)
Previous Exports Growth	-0.1077 (0.019)***	-0.1073 (0.019)***	-0.1103 (0.019)***	-0.1114 (0.019)***
Previous Exports Zero	-0.0645 (0.018)***	-0.0650 (0.018)***	-0.0599 (0.019)***	-0.0595 (0.018)***
N	55580	55580	55580	55580
r2	0.32	0.32	0.32	0.32

All specifications include country-by-year and product-by-year fixed effects. The dependent variable in all specifications is constructed using exports of country  $c$  to the whole world excluding to countries  $c'$  where total migration between  $c$  and  $c'$  exceeds 500 people.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Table 6: Instrumental Variables Estimation (GMM)

<b>Panel A: Extensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	-0.0004 (0.002)	-0.0014 (0.004)	0.0002 (0.001)	-0.0005 (0.002)
Ln Emigrants	0.0110 (0.003)***	0.0154 (0.006)***	0.0050 (0.001)***	0.0082 (0.002)***
Ln FDI, total	-0.0001 (0.001)	-0.0003 (0.001)	-0.0004 (0.000)	-0.0004 (0.000)
Ln Trade, total	-0.0072 (0.005)	-0.0084 (0.005)*	0.0003 (0.001)	0.0004 (0.001)
N	83396	83396	83396	83396
r2	0.14	0.14	0.15	0.14
KP F Stat	20.38	15.34	84.01	24.19
<b>Panel B: Intensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0165 (0.008)**	0.0224 (0.007)***	-0.0022 (0.004)	-0.0033 (0.004)
Ln Emigrants	-0.0074 (0.013)	-0.0193 (0.014)	0.0146 (0.004)***	0.0162 (0.004)***
Ln FDI, total	-0.0013 (0.001)**	-0.0009 (0.001)	-0.0004 (0.000)	-0.0005 (0.000)
Ln Trade, total	0.0056 (0.009)	0.0098 (0.009)	-0.0002 (0.001)	0.0005 (0.001)
Baseline Log Exports	-0.0306 (0.002)***	-0.0307 (0.002)***	-0.0295 (0.002)***	-0.0287 (0.001)***
Previous Exports Growth	-0.1270 (0.013)***	-0.1284 (0.013)***	-0.1328 (0.013)***	-0.1330 (0.013)***
Previous Exports Zero	-0.0105 (0.013)	-0.0098 (0.013)	-0.0033 (0.012)	-0.0050 (0.012)
N	73193	73193	73193	73193
r2	0.34	0.34	0.34	0.34
KP F Stat	3.95	7.60	15.33	14.50

All specifications include country-by-year and product-by-year fixed effects. SE clustered at the country level presented in parenthesis.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

## A Appendix

### A.1 A Conceptual Framework

Consider a small open economy in a world with a fixed set  $\omega$  of goods, each one, indexed by  $i$ , with the same production function given by:

$$q_i = \varphi_i L_I^\alpha$$

Where  $0 < \alpha < 1$ ,  $L$  is units of labor, the only factor of production (which is inelastically supplied) and  $\varphi_i$  is product specific productivity. Prices are exogenously determined in world markets, and defined by  $p_i$ . Product specific productivity,  $\varphi_i$ , is an increasing function of product-specific tacit knowledge which is equals the amount of non-workers people in the economy,  $\eta_i$ , that have such product specific knowledge. Each economy has an initial endowed vector of  $\eta_i$  (for each product in the set  $\omega$ ) determined exogenously.

Assume the following functional form:

$$\varphi_i = \eta_i^\rho \text{ where } 0 < \rho < \frac{1-\alpha}{1+\alpha} < 1 \quad (2)$$

Each firm has to pay a fixed cost  $C(\varphi) = \frac{1}{\varphi_i^{\frac{1}{1-\alpha}}}$  to enter the market. Hence, the total cost of production is:

$$TC_i = w \left[ \frac{q_i}{\varphi_i} \right]^{1/\alpha} + C(\varphi_i)$$

Where  $w$  is wage, which is set to 1 for simplicity. Hence, a firm will produce the product only if it can achieve positive profits:

$$\Pi_i = p_i q_i - \left[ \frac{q_i}{\varphi_i} \right]^{1/\alpha} - C(\varphi_i) > 0$$

Hence, to have  $\Pi_i > 0$  we must have that:

$$\varphi_i > \left[ \frac{q_i^{1/\alpha} + 1}{p_i q_i} \right]^\alpha \quad (3)$$

Assuming equation (3) holds, we can compute the optimal amount  $q_i$  that a firm will produce given the exogenous price  $p_i$  (by equating price to MC):

$$p_i - \frac{1}{\alpha \varphi_i} \left[ \frac{q_i}{\varphi_i} \right]^{\frac{1-\alpha}{\alpha}} = 0$$

Solving this results in:

$$q_i^* = (\alpha p_i)^{\frac{\alpha}{1-\alpha}} \varphi_i^{\frac{1+\alpha}{1-\alpha}} = (\alpha p_i)^{\frac{\alpha}{1-\alpha}} \eta_i^{\frac{\rho(1+\alpha)}{1-\alpha}} \quad (4)$$

It is easy to see that  $\frac{\partial q_i}{\partial \eta_i} > 0$ , and  $\frac{\partial^2 q_i}{\partial \eta_i^2} < 0$  given the assumptions in (2).

This brings us to Proposition 1:

**Proposition 1.** *A firm's decision on the quantity of the product it produces, once such enters the market, is increasing in the amount of knowledge specific to that product available in the country.*

Thus, one testable implication based on Proposition 1, at the country level, is that a country will increase the level of production/exports of certain

product whenever it has acquired more knowledge specific to that product.

With the equilibrium amount  $q_i^*$  from (4), we can find the required amount of knowledge needed for a firm in order to enter the market. Hence, plugging  $q_i^*$  into (3) we have:

$$\Gamma(\eta_i) = \eta_i - \left[ \frac{\eta_i^{\frac{\rho}{1-\alpha}} (\alpha p_i)^{\frac{\alpha}{1-\alpha}} \left[ 1 + \left( \eta_i^{\frac{\rho}{1-\alpha}} (\alpha p_i)^{\frac{\alpha}{1-\alpha}} \right)^{\frac{1}{\alpha}} \right]}{p_i} \right]^{\frac{\alpha}{\rho}} > 0$$

Therefore, it can be inferred:

$$q_i > 0 \text{ if } \Gamma(\eta_i) > 0, \text{ with } \frac{\partial \Gamma(\eta_i)}{\partial \eta_i} > 0$$

**Proposition 2.** *A product will be produced by a firm in the country if the country has the minimum required amount of specific product knowledge to make it profitable to the firm. The probability of a product being produced by a firm in a country is an increasing function of the number of people in the economy that have the product specific knowledge.*

A second testable implication based on Proposition 2, at the country level, implies that an increase in the amount of specific knowledge on certain product should be positively correlated with the ability of that country to start producing/exporting such product.

$$\frac{\partial q_i}{\partial \eta_i} = \frac{\partial q_i}{\partial \varphi_i} \frac{\partial \varphi_i}{\partial \eta_i} \quad (5)$$

The empirical section estimates changes in  $q_i$  as a result of variation in  $\eta_i$ . Changes in  $\eta_i$  specific to each product are proxied by exploiting the variation on bilateral relationships (i.e. migration, trade and FDI) from and to countries that export intensively product  $i$ .

## A.2 Robustness Tests

### A.2.1 Excluding products pinned down from geology or climate

In order to limit the sample to products that are not pinned down by geology or climate, we follow the classification provided by Lall (2000), shown in Table A1. Lall's classification is used to create two categories of products: Primary and Resource Based (PRB) products and Non-Primary or 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 A1), whereas NPRB products are the ones contained in all other categories.

[Table A1 about here.]

The results, presented in Tables A2, are consistent with the ones shown in the main body of the paper. In the Table, the upper panel shows the results for the extensive margin while the lower panel studies the intensive margin of trade. The results are qualitatively similar to those in Table 4. While not robust to all specifications, in most of them immigrants and emigrants correlate to the ability of countries to export new NPRB products. In the

intensive margin, immigrants and emigrants appear to have positive and statistically significant coefficients when using the  $RCA \geq 5$  threshold, while the evidence is mixed for the  $RCA \geq 1$  threshold.

[Table A2 about here.]

### **A.2.2 Excluding Bilateral Exports**

Kugler and Rapoport (2011) find evidence that migrants reduce transaction costs inducing bilateral trade and capital flows. This section expand on the test presented in the main body of the paper which discard the results being driven by this pattern, instead of a gain in productivity.

As explained in the main body of the paper, we reconstruct the dataset such that the export value for each product and country to the rest of the world excludes exports to nations where migrants are from/in. The number of minimum number of migrants (immigrants + emigrants) used to exclude those bilateral flows are defined in thresholds. Tables A3 to A5 present results using as thresholds 1000, 2500 and 5000 migrants, to complement the result in the main body of the paper that presented the 500 threshold (i.e., if there are more than 500 migrants between any two given countries, we exclude all bilateral trade between those two countries, to construct the dependent variable).

[Table A3 about here.]

[Table A4 about here.]



[Table A5 about here.]

Table A1: 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

Table A2: Fixed Effects, NPRB products

<b>Panel A: Extensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0010 (0.000)**	0.0019 (0.001)***	-0.0004 (0.001)	0.0013 (0.001)
Ln Emigrants	0.0029 (0.001)**	0.0028 (0.001)*	0.0022 (0.001)***	0.0025 (0.001)***
Ln FDI, total	0.0017 (0.001)	0.0016 (0.001)	0.0011 (0.001)*	0.0009 (0.001)
Ln Trade, total	-0.0040 (0.004)	-0.0045 (0.004)	0.0007 (0.001)	0.0007 (0.001)
N	41215	41215	41215	41215
r2	0.21	0.21	0.21	0.21
<b>Panel B: Intensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0014 (0.001)	0.0041 (0.001)***	0.0018 (0.001)**	0.0017 (0.001)*
Ln Emigrants	0.0048 (0.002)**	-0.0031 (0.002)	0.0037 (0.001)***	0.0036 (0.001)***
Ln FDI, total	-0.0016 (0.001)***	-0.0013 (0.001)**	-0.0001 (0.000)	-0.0001 (0.000)
Ln Trade, total	0.0175 (0.006)***	0.0227 (0.006)***	0.0009 (0.001)	0.0014 (0.001)*
Baseline Log Exports	-0.0446 (0.002)***	-0.0446 (0.002)***	-0.0436 (0.002)***	-0.0433 (0.002)***
Previous Exports Growth	-0.0853 (0.016)***	-0.0850 (0.016)***	-0.0885 (0.017)***	-0.0896 (0.017)***
Previous Exports Zero	-0.0310 (0.016)**	-0.0310 (0.015)**	-0.0266 (0.016)*	-0.0258 (0.016)*
N	42737	42737	42737	42737
r2	0.40	0.40	0.40	0.39

All specifications include country-by-year and product-by-year fixed effects. SE clustered at the country level presented in parenthesis

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Table A3: FE, excluding bilateral exports (1000 migrants threshold)

<b>Panel A: Extensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0017 (0.000)***	0.0028 (0.001)***	0.0011 (0.000)***	0.0012 (0.001)*
Ln Emigrants	0.0032 (0.001)***	0.0024 (0.001)**	0.0013 (0.000)***	0.0008 (0.000)*
Ln FDI, total	0.0008 (0.001)	0.0007 (0.001)	-0.0001 (0.000)	-0.0001 (0.000)
Ln Trade, total	-0.0061 (0.004)	-0.0063 (0.004)	0.0027 (0.001)**	0.0028 (0.001)**
N	114227	114227	114227	114227
r2	0.11	0.11	0.11	0.11
<b>Panel B: Intensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0005 (0.001)	0.0034 (0.002)**	0.0020 (0.001)**	0.0015 (0.001)
Ln Emigrants	0.0072 (0.002)***	-0.0007 (0.003)	0.0031 (0.001)***	0.0030 (0.001)**
Ln FDI, total	-0.0009 (0.001)	-0.0007 (0.001)	0.0010 (0.000)***	0.0010 (0.000)***
Ln Trade, total	0.0124 (0.006)**	0.0175 (0.006)***	0.0016 (0.001)*	0.0022 (0.001)**
Baseline Log Exports	-0.0072 (0.002)***	-0.0072 (0.002)***	-0.0062 (0.002)***	-0.0058 (0.002)***
Previous Exports Growth	-0.1045 (0.017)***	-0.1041 (0.017)***	-0.1077 (0.017)***	-0.1088 (0.017)***
Previous Exports Zero	-0.0691 (0.016)***	-0.0695 (0.016)***	-0.0639 (0.016)***	-0.0635 (0.016)***
N	57956	57956	57956	57956
r2	0.32	0.32	0.32	0.32

All specifications include country-by-year and product-by-year fixed effects. The dependent variable in all specifications is constructed using exports of country  $c$  to the whole world excluding to countries  $c'$  where total migration between  $c$  and  $c'$  exceeds 1000 people.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Table A4: FE, excluding bilateral exports (2500 migrants threshold)

<b>Panel A: Extensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0018 (0.000)***	0.0028 (0.001)***	0.0009 (0.000)**	0.0009 (0.001)
Ln Emigrants	0.0029 (0.001)***	0.0024 (0.001)**	0.0013 (0.000)***	0.0012 (0.001)**
Ln FDI, total	0.0009 (0.001)	0.0008 (0.001)	-0.0002 (0.000)	-0.0002 (0.000)
Ln Trade, total	-0.0063 (0.004)	-0.0065 (0.004)	0.0023 (0.001)*	0.0023 (0.001)*
N	106881	106881	106881	106881
r2	0.11	0.11	0.11	0.11
<b>Panel B: Intensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0012 (0.001)	0.0029 (0.001)**	0.0020 (0.001)**	0.0014 (0.001)
Ln Emigrants	0.0067 (0.002)***	-0.0009 (0.003)	0.0034 (0.001)***	0.0030 (0.001)**
Ln FDI, total	-0.0006 (0.001)	-0.0004 (0.001)	0.0007 (0.000)*	0.0008 (0.000)**
Ln Trade, total	0.0094 (0.006)*	0.0157 (0.006)***	0.0017 (0.001)**	0.0024 (0.001)***
Baseline Log Exports	-0.0126 (0.002)***	-0.0124 (0.002)***	-0.0118 (0.002)***	-0.0114 (0.002)***
Previous Exports Growth	-0.1070 (0.016)***	-0.1072 (0.016)***	-0.1096 (0.016)***	-0.1109 (0.016)***
Previous Exports Zero	-0.0610 (0.015)***	-0.0611 (0.015)***	-0.0567 (0.015)***	-0.0561 (0.015)***
N	61476	61476	61476	61476
r2	0.31	0.31	0.31	0.31

All specifications include country-by-year and product-by-year fixed effects. The dependent variable in all specifications is constructed using exports of country  $c$  to the whole world excluding to countries  $c'$  where total migration between  $c$  and  $c'$  exceeds 2500 people.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Table A5: FE, excluding bilateral exports (5000 migrants threshold)

<b>Panel A: Extensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0020 (0.000)***	0.0026 (0.001)***	0.0009 (0.000)**	0.0009 (0.001)
Ln Emigrants	0.0026 (0.001)***	0.0023 (0.001)**	0.0016 (0.000)***	0.0015 (0.001)***
Ln FDI, total	0.0009 (0.001)	0.0009 (0.001)	-0.0000 (0.000)	-0.0000 (0.000)
Ln Trade, total	-0.0068 (0.004)	-0.0069 (0.004)	0.0020 (0.001)	0.0021 (0.001)
N	100988	100988	100988	100988
r2	0.12	0.12	0.12	0.12
<b>Panel B: Intensive Margin</b>				
	R1		R5	
	All	Skilled	All	Skilled
Ln Immigrants	0.0017 (0.001)	0.0040 (0.001)***	0.0020 (0.001)***	0.0015 (0.001)
Ln Emigrants	0.0062 (0.002)***	-0.0025 (0.003)	0.0039 (0.001)***	0.0035 (0.001)***
Ln FDI, total	-0.0009 (0.000)*	-0.0007 (0.000)	0.0005 (0.000)	0.0005 (0.000)
Ln Trade, total	0.0087 (0.005)	0.0152 (0.005)***	0.0017 (0.001)**	0.0024 (0.001)***
Baseline Log Exports	-0.0165 (0.002)***	-0.0163 (0.002)***	-0.0158 (0.002)***	-0.0153 (0.002)***
Previous Exports Growth	-0.1060 (0.016)***	-0.1060 (0.016)***	-0.1083 (0.016)***	-0.1097 (0.016)***
Previous Exports Zero	-0.0549 (0.015)***	-0.0551 (0.015)***	-0.0508 (0.015)***	-0.0503 (0.015)***
N	64343	64343	64343	64343
r2	0.32	0.32	0.32	0.32

All specifications include country-by-year and product-by-year fixed effects. The dependent variable in all specifications is constructed using exports of country  $c$  to the whole world excluding to countries  $c'$  where total migration between  $c$  and  $c'$  exceeds 5000 people.

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$