International Competitiveness and Migration: Diversity, Networks or Knowledge Diffusion?

Preliminary and incomplete, please do not circulate[*]

Gianluca Orefice (CEPII)[†] Hillel Rapoport (PSE)[‡] Gianluca Santoni (CEPII) [§]

January 9, 2019

Abstract

Immigrants have a positive effect on the export performance of host countries. Diversity, knowledge diffusion and networks effects have been shown (separately) to be valid explanations for such an evidence. This paper tests these three mechanisms in a unified empirical framework and provides evidence of the heterogeneous effects of birthplace diversity on the international competitiveness of host countries across sectors with different characteristics. We show that networks, knowledge diffusion and birthplace diversity stimulate the international competitiveness of the host economy through both the *intensive* and the *extensive* margins of trade. Also, we find that sectors whose production process is complex and intensive in cognitive tasks benefit more from birthplace diversity. This suggests that a more diverse labor supply is beneficial when problem solving capabilities are needed in production. We address the potential endogeneity problem by adopting an instrumental variable (IV) approach based on three extensions of the traditional shift share IV.

Key Words: International Trade, Birthplace Diversity, Migration, Productivity.

JEL Codes: F14, F16, F22, O47.

[*]We are grateful to Elisabetta Lodigiani and Thierry Mayer and participant to the 7th OECD Conference on Migration for insightful comments. The views expressed in this article are those of the authors and also do not reflect those of the CEPII.

[†]CEPII, 20 Avenue de Segur 75007 Paris. Email: gianluca.orefice@cepii.fr.

[‡]PSE, 48 Boulevard Jourdan 75014 Paris. Email: hillel.rapoport@univ-paris1.fr.

[§]CEPII, 20 avenue de Segur 75007 Paris. Email: gianluca.santoni@cepii.fr.

1 Introduction

The positive effect of migration on international trade has been widely documented in the previous literature. By providing information on their country of origin, immigrants reduce the transaction cost of trade and favor the export flows towards their origin countries, i.e. networks or transaction cost channel (Gould (1994); Head & Ries (1998); Rauch (2001); Felbermayr & Toubal (2012) and Parsons & Vezina (2018)). Also, immigrants have been shown bringing along knowledge when they move for origin to destination country. The migration-driven transfer of knowledge increases the efficiency of production in destination countries, and therefore improves their export performances (comparative advantage channel). Namely, migrants arriving from countries having comparative advantage in a given sector, may contribute to the development of such a sector in the country of destination, i.e. *knowledge* diffusion channel (Bahar & Rapoport (2018).

Recent papers observe an additional channel through which migrants may affect the export performance of destination country: migrants may affect the average productivity of destination countries through the *diversity* channel (Ortega & Peri (2014); Alesina, Harnoss & Rapoport (2016); Docquier, Turati, Valette & Vasilakis (2018)). People originating form a heterogeneous set of countries bring at destination a more diverse set of skills, experiences, ideas, expertise and problem-solving capabilities that may be useful to improve the efficiency of the production process and the overall performance of the firm (Lazear (1999); Hong & Page (2001); Horwitz & Horwitz (2007)). Although diversity in production units also implies coordination costs (linguistic barriers), and in some extreme cases the segregation of minorities in production, Lazear (1999) shows that diversity in team-working may offset the coordination cost is presence of a certain degree of complementarity across abilities in production.

In this paper we test the effect of Birthplace Diversity on the export performances of the receiving country by explicitly controlling for both the network and the knowledge diffusion effect. To the best of our knowledge this is the first paper that test these three channels in a unified empirical framework. This is crucial contribution of our paper. Indeed, testing one of the three channels not explicitly controlling for the others (as done in previous studies) implies an unclear identification of the channel under the lens. After testing the effect of the three channels on the international competitiveness of destination countries, we focus on the effect of birthplace diversity. In particular we study the effect of birthplace diversity on several measures of country's export performances and test the mechanism at play. By testing the heterogeneous effect of Birthplace Diversity across sectors with different characteristics, we argue the mechanism through which diversity may impact the international competitiveness of a country. These research objectives are shown empirically in two steps. First, using bilateral specific regressions we test the effect of the three channels (i.e. networks, knowledge diffusion and diversity) on the *intensive* and *extensive* margin of trade. Then, in a second part, we focus our attention to the Birthplace Diversity channel and test, using country-sector aggregated regressions, the mechanism through which it may affect the international competitiveness of countries.[1]

In the second part of the paper we adopt the ex-ante Revealed Comparative Advantage (RCA), in the vein of Costinot, Donaldson & Komunjer (2012), as an alternative proxy for the international competitiveness of a country. Based on a first stage bilateral trade regression, as in Costinot et al. (2012), we use predicted country-year fixed effects to build synthetic country specific measures of RCA that can be used in regression as a proxy for the international competitiveness of a country in a given sector. Therefore, as a side product of the present paper, we collect these synthetic measures of revealed comparative advantages in a new CEPII database freely available for scholars and practitioners in this field. Adopting an instrumental variable (IV) approach based on three extensions of the traditional shift share IV à la Card (2001), the contribution of the present paper is also methodological. Our first IV bases on the pure supply-driven component of migration inflows, by removing any demand-driven factor from a predicted bilateral stock of migrants. The second IV bases on the (supply-driven) inflows of immigrants occurred after natural disasters at origin countries. Finally, we also provide an instrumental variable approach based on the recent contribution by Jaeger, Ruist & Stuhler (2018). [Literature on network and knowledge diffusion to be included] .....

The effect of diversity and productivity of production teams as been shown in many papers. Hong & Page (2001) theoretically show that a group of more diverse problem solvers may perform better than a group of homogeneous but more ables problem solvers. Hoogendoorn & van Praag (2012) use a randomized field experiment to show that more ethnically diverse teams have better performance than ethnically homogeneous teams: in diverse teams the coordination costs from ethnic and linguistic diversity are offset by the wider availability of relevant skills.[2] Abstracting from team of workers in production, Kahane, Longley & Simmons (2013) analyze the ethnic composition of National Hockey League teams in the US and find that more diverse teams have better performance. Interestingly, Kahane et al. (2013) conclude that the "productivity" premium provided by diverse teams is driven by complementarity between native and foreign-born players' skills.[3] Trax, Brunow & Suedekum (2015) use German establishment level data to show that diversity of foreign born workers increases the productivity of plants. Parrotta, Pozzoli & Pytlikova (2014) using matched employer-employee data from Denmark, show that ethnic diversity has a negative effect on total factor productivity at the level of the firm, while education diversity has a positive effect on productivity. At the aggregate local labor market level, Ottaviano & Peri (2006) find that multicultural urban environment increases the productivity of US-born citizens. Similarly, a recent study by Rodríguez-Pose & von Berlepsch (2017) on US counties identifies the presence of a strong positive impact of population diversity on county-level economic development: counties

[1] In the second part we move to aggregate country-sector regressions to adopt the same dimension of the Birthplace Diversity measure and test its heterogeneous impact on competitiveness of different types of sectors.

[2] A recent research published by McKinsey&Company find a significant positive relationship between ethnic diverse teams and financial performances of firms (EBIT). Companies at the top quartile of ethnic diversity are 35 percent more likely to have outperform their national industry median (Vivian, Dennis & Sara 2015).

[3] Peri & Sparber (2009) provide empirical evidence of a productivity effect from the complementarity among immigrant and native workers.

that received migrants from more diverse set of origins over the late 19^{th} century are nowadays significantly richer than counties with a more homogeneous population at the time.

[Literature on diversity and productivity - macro to be included] To be cited: Ortega & Peri (2014); Alesina et al. (2016); Bahar & Rapoport (2018); Docquier et al. (2018).

To our knowledge, only few papers directly link population diversity and international trade. Maggi & Grossman (2000) develop a theoretical model in which the distribution of worker types (reflecting the diversification of abilities in the country) contributes to the country's comparative advantage and export performance. In Maggi & Grossman (2000), countries with a more diverse population have a comparative advantage in the production/export of goods characterized by high substitutability among employees in production (i.e. when the presence of highly-talented workers is relatively more important). From an empirical point of view, Parrotta, Pozzoli & Sala (2016) test the effect of ethnic diversity of the export performance of Danish firms. They find a strong positive effect of firm's workforce diversity on the extensive margin of exports (participation and number of export markets), and a non-significant effect on the intensive margin channel.

The rest of the paper is organized as follows. The next section describes the empirical strategy we adopt to test the three channels in a unified empirical model, the approach adopted to address the endogeneity problem and discuss the results. In section 3 we move at the country-sector level and discuss the empirical framework used to test the heterogeneous effect of birthplace diversity on different trade margins, and the results. The last section concludes.

2 Migration and international competitiveness: three channels and a unified empirical strategy

This section describes the empirical strategy used to test the effect of networks, knowledge diffusion and birthplace diversity channel in a unified empirical framework. In particular, we test the effect of these three channels on the international competitiveness of countries and its components (intensive and extensive margins). To this end we adapt a standard gravity model for trade to include the three channels mentioned above:[4]

$$y_{ikjt} = \beta_1 Mig_{ijt} + \beta_2 Mig_{ijt} * (TS)_{ijk} + \beta_3 (TS)_{ijk} + \beta_4 BD_{it} + X_{ijt} + \theta_{jkt} + \theta_{rlt} + \varepsilon_{ijkt} \tag{1}$$

where the dependent variable y_{ijkt} is in turn: (i) a dummy variable equal to one if the country i exports to j in a given sector k at time t (*extensive margin*); and (ii) the (log of) total exports of country i to j for sector k at time t, conditioned on being already serving the market jk at time $(t-1)$ - *intensive margin*. Notice that other trade margins, like the number of destinations and the number of exported products need different data

[4]See Head & Mayer (2014) for a discussion on the gravity model for trade.

aggregation (exporter-sector-year aggregated data) and will be covered in the second part of the paper.

Three main explanatory variables characterize the empirical exercise. First, the stock of immigrants (in ln) in destination i from origin j and time t (Mig_{ijt}) aims at capturing the transaction cost channel. The availability of immigrants coming from j origin is expected to boost the exports of i to j ($\beta_1 > 0$). Second, in line with Bahar & Rapoport (2018) we test the knowledge diffusion channel by interacting the stock of immigrants Mig_{ijt} with a dummy (TS_{ijk}) equal to one if country j has a Revealed Comparative Advantage in sector k (i.e. $RCA_{jk,1995} > 1$) and country i has not ($RCA_{ik,1995} < 1$). We use the Balassa Index in 1995 to approximate the ex-ante comparative advantage of country i and j in a given sector k. A Balassa Index greater (smaller) than one suggests a comparative advantage (*dis*advantage) of a country in sector k. As a sensitivity analysis, we use an alternative definition of TS_{ijk}, being equal to one if $RCA_{jk,1995} > 1.25$ and $RCA_{jk,1995} < 0.75$ or if $RCA_{jk,1995} > 1.5$ and $RCA_{jk,1995} < 0.5$. The interaction term - $Mig_{i,j,t}$*$TS_{i,j,k}$ - captures the effect of migrants originating from countries having a comparative advantage in k, on countries with comparative *dis*advantage in k. In presence of positive knowledge diffusion effect, we expect ($\beta_2 > 0$). Finally, in line with the previous literature we define Birthplace Diversity (BD) as one minus the Herfindahl-Hirschman (HH) concentration index applied to the population of immigrant population $BD_{i,t} = 1 - \sum_{j=1}^{J} s_{ijt}^2$, where s_{ijt} is the share of immigrants originating from country j in the total population of immigrants residing in country i at time t. The index of birthplace diversity $BD_{i,t}$ increases with the diversity in migrants' birthplaces in the country (it is equal to 0 if country i hosts immigrants coming from only one origin country). The birthplace diversity index $BD_{i,t}$ can be interpreted as the probability that two randomly selected *foreign born residents* are from different countries of origins; so it increases when the size of a minority group increases, and decreases with the size of a majority group. If the diversity of birthplace in the set of migrants' origins has a positive effect on exports, we should obtain $\beta_4 > 0$.

The set of control variables - X_{ijt} - includes geographic and trade policy variables often used in the empirics of gravity equation: (i) distance (in ln); (ii) time difference; (iii) common colonial ties; (iv) common language; (v) common border; (vi) a dummy for bilateral trade agreement RTA_{ijt} (capturing the effect of a preferential market access); (vi) the applied tariff, included as $log(1 + tariff)$ and controlling for the tariff level faced by country i in exporting to j in sector k.[5] Moreover, we include the stock of emigrants from i living in j to control for the effect of preferences and taste difference on bilateral trade.[6] The bigger the diaspora from i to j, the higher is the import demand of country j for goods produced in country i.

Two set of fixed effects are always included in estimations. Importer-sector-year fixed effects (θ_{jkt}) control for any unobserved factor at the level of the importer country-sector-year that may affect the export performances

[5]Notice that the effect of MFN tariff imposed in country j in sector k is captured by the importer-sector-year fixed effects. Data on applied tariffs are from the WITS-TRAINS database.

[6]The so-called *Transplanted home bias effect*, White (2007).

towards the market jk. In particular, this set of fixed effects controls for the multilateral resistance term on the importer side (Head & Mayer (2014)). Since one of the variable of interest ($BD_{i,t}$) is exporter country-year specific, we cannot include fixed effects on this dimension. Namely, exporter country-sector-year that would capture exactly the multilateral resistance term on the exporter side cannot be included. To (partially) address the potential omitted variable problem, we always include fixed effects specific to the macro region-income level (and year) of the exporter country, θ_{rlt}.[7] A similar strategy is used in Alesina et al. (2016) where authors include macro region fixed effects because country dummies would be perfectly collinear with birthplace diversity. By doing so, any unobserved shock specific to a macro region within a given income level is captured by fixed effects.[8] As further controls for the (exporter) multilateral resistance term we include: i) a country remoteness index; [9] ii) a series of dummy variables for the quartile of total exports of country i in a given sector k.[10] Total exports bins aim at capturing the export capacity of country i in sector k as suggested by standard gravity equation.[11] Therefore, within each region-income cell, total export quintile and conditional to the exporter market access (i.e. remoteness), countries are assumed to be plausibly homogeneous in terms of factor endowments (capital and skilled/unskilled labor force), technological level, quality of institutions and infrastructure.

2.1 Data and Descriptive evidence

All the migration related variables (i.e. bilateral migration stocks and birthplace diversity) are based on ij specific bilateral stocks of migrants from United Nations (2015). This dataset provides information on bilateral migration stocks for a 195*195 matrix of origin-destination combinations, for the years 1990, 1995, 2000, 2005, 2010, 2015.[12] In table 1, for a sub-sample of the countries covered in our empirical analysis, we report the stock of immigrants from all origins, the stock of immigrants expected to diffuse knowledge (i.e. from countries where $TS_{ijk} = 1$ for at least one k), and the value of Birthplace Diversity.

Export based measures of international competitiveness (i.e. total exports, intensive and extensive margins) are based on BACI (CEPII) data. We have information on bilateral export flows from/to 195 countries, over the

[7] The macro-region and the income levels of countries are obtained from World Bank classification. For example we have a dummy for South American countries belonging to the same income level (as defined by the World Bank, for the income level we consider year 1995).

[8] In Table A2 is reported a detailed description of each region-income level cell and the number of countries belonging to each cell.

[9] Following Yotov, Piermartini, Monteiro & Larch (2017) we construct the remoteness index for the exporting country as: $ln(Remote)_{it} = ln(\sum_j^J dist_{ij}/E_{jt}/Y_t)$.

[10] In order to reduce endogeneity concerns we exclude the direct exports of country i towards country j and other destinations of the same region of country j.

[11] In a standard gravity equation the export flow from country i to j depends on the overall international competitiveness of country i - i.e. marginal cost in Armington under perfect competition. This may be approximated by bins in overall export performance of country i purged by j specific factors.

[12] The dataset *Trends in International Migrant Stock: The 2015 Revision* (United Nations database, POP/DB/MIG/Stock/Rev.2015) is available at: `http://www.un.org/en/development/desa/population/migration/data/estimates2/estimates15.shtml`. However, given that trade data start in 1995, in the econometric estimations we cover the period 1995-2015. The main advantage of this dataset, with respect to other sources (e.g. IMD-OECD), is the balanced nature of the data which include all other non-OECD destination countries. For periods prior to 1990 (used to build our instrumental variable) we use data from the World Bank *Global Bilateral Migration Database*, see Ozden, Parsons, Schiff & Walmsley (2011).

period 1995-2015 covered in this paper, at product HS 6-digit level. However, to work with manageable dataset, we aggregate the trade data at the country pair-sector-year level where the sector is defined as HS 2-digit product classification. In table 2 we show in-sample descriptive statistics for the main variables included in our estimations. Figure 1 shows the simple correlations between the total exports of countries and three channels at the core of the empirical exercise: (i) total stock of immigrants (bilateral and country specific, respectively in top-left and bottom-left scatter), (ii) immigrants contributing to the diffusion of knowledge (i.e. coming from origins having comparative advantage in sectors in which the exporting countries has not) in the top-right scatter, and (iii) birthplace diversity in the bottom-right scatter. Figure 1 suggests the positive correlation between the three channels discussed above and the total export of country.

Data on the presence of Preferential Trade Agreements, distance, common language, border and colony come from CEPII databases. Data on GDP per capita (used to calculate the remoteness measures), income and regional classifications are from World Bank Development Indicators data.

2.2 Endogeneity

While the omitted variable concern is reduced by the inclusion of destination region-by-income level fixed effects, and country specific controls capturing the trade/migration costs as main determinants of bilateral migration, it may still be the case that unobserved country i specific shocks affect contemporaneously its export performances and the stock of immigrants coming from different origins (i.e. positive productivity shocks boosting export and attracting immigrants). Moreover, reverse causality may produce biased Ordinary Least Squared (OLS) estimations if the international competitiveness of a country (and thus its production for the foreign market) has an impact on the labor demand for immigrants workers. We address these endogeneity concerns by adopting an Instrumental Variable approach that uses (in turn) two *original* IVs and a third instrumental variable approach in the vein of Jaeger et al. (2018). The first instrumental variable bases on the predicted supply-driven migration stocks purged from any demand-driven effect - see section 2.2.1. In the same vein, a second IV also bases on the predicted supply-driven migration but uses the time variation of immigrants coming from origins that experienced natural disaster (i.e. exogenous shock in the push factors) - see section 2.2.3. Finally, following the main idea in Jaeger et al. (2018), in a third IV strategy we remove the feedback effect in the predicted supply-driven migration stocks.

2.2.1 IV 1: a modified shift-share based instrumental variable

Previous papers on the labor market effect of immigration have often adopted the shift share instrument à la Card (2001) to solve the endogeneity problem.[13] This instrument is based on the idea that *contemporaneous*

[13] See for example Ottaviano & Peri (2006); Peri & Requena-Silvente (2010); Card (2009)

outflows of migrants from a given origin country are allocated across different destinations based on the *historical* geographical distribution of migrants from the same origin country. While this is a feasible solution in many empirical settings, it cannot be directly applied in our framework: labor demand shocks related to the international competitiveness of an exporting country may directly attract migrants from a given origin country (reverse causality). We therefore propose an alternative instrumental variable that can be used when the shift share instrument à la Card (2001) cannot be directly adopted. The idea is obtaining the *predicted supply-driven* stock of immigrants in each country i based on the estimation of the following structural gravity model for migration:

$$Mig_{ijt} = \delta_{it} + \delta_{jt} + \delta_{ij}^{ij=ji} + \beta_1 Immi\ Sh_{ij,60} * ln(Mig)_{jt} + \beta_2 Immi\ Sh_{ij,60} + \mu_{ijt} \tag{2}$$

where the destination-year and the origin-year fixed effects (δ_{it}, δ_{jt}) respectively capture the pull and the push factors of bilateral migration; while the symmetric country-pair fixed effects ($\delta_{ij}^{ij=ji}$) account for the time invariant and pair specific symmetric migration costs. The inclusion of the interaction between the share of emigrants of j resident in the host country i in 1960 ($Immi\ Sh_{ij,60}$) and the total stock of outflow immigrants from j at time t (variable $ln(IMig)_{jt}$) mimics the idea of Card (2001), i.e. that immigrant tend to settle in destinations where previous immigrants from the same origin already reside. From equation (2) we take the predicted value $\widehat{Immi}_{ijt}$ (fit of the regression) and subtract the destination-year fixed effect:

$$\widehat{AdjImmi}_{ijt} = \widehat{Immi}_{ijt} - \widehat{\delta_{it}}. \tag{3}$$

By doing so, we purge the predicted bilateral stock of migrants from the "problematic" demand-driven component that makes the standard shift share instrument à la Card (2001) endogenous in our empirical setting. Gravity-based 2SLS strategy has been widely used in the trade literature to predict trade openness out of bilateral trade flows (see the seminal paper by Frankel & Romer (1999)). This approach has been successively adopted in the migration literature to instrument migration share at destination - see Ortega & Peri (2014) and Docquier, Lodigiani, Rapoport & Schiff (2016) among others. In particular, Ortega & Peri (2014) and Docquier et al. (2016) estimate a bilateral migration gravity equation including several proxies for geographic and cultural distance as explanatory variables. The fit of this equation is then used as an instrumental variable for the total stock of migrants at destination. While geographic and cultural distance can fairly be assumed exogenous with respect to the economic performances of the destination country (exclusion restriction), both papers include other gravity-related variables (i.e. population and immigration policy at destination) that may be affected by the economic outcomes of the destination country.[14] Under this circumstance, the exclusion restriction is

[14] Negative economic shock may reflect into a change in the destination country's population and immigration policy setting.

not satisfied. It is therefore important to remove from the gravity-based predictor, the estimated destination country-time specific component.

$\widehat{AdjImmi}_{ijt}$ is therefore used to instrument the bilateral stock of migrants Mig_{ijt} and its interaction with TS_{ijk}. Finally, we build the instrumental variable for the Birthplace Diversity index as described above, but using $\widehat{AdjImmi}_{ijt}$:

$$\widehat{BD_{it}^{PPML}} = 1 - \sum_{j=1}^{J} \left(\frac{\widehat{AdjImmi}_{ijt}}{\sum_{j=1}^{J} \widehat{AdjImmi}_{ijt}} \right)^2 \tag{4}$$

Notice that $\left(\frac{\widehat{AdjImmi}_{ijt}}{\sum_{j=1}^{J} \widehat{AdjImmi}_{ijt})} \right)$ is the share in the total population of *supply driven* predicted number of migrants in country i originating from country j. Therefore, our instrumental variable $\widehat{BD_{it}^{PPML}}$ is built using the pure supply-driven component of the bilateral migration stocks, and can be used safely as an instrumental variable. In this case, the exclusion restriction assumption is that the diversity index based on the *predicted supply-driven* migration stocks ($\widehat{BD_{it}^{PPML}}$) affects the competitiveness of a country only through the BD_{it} index based on the *observed* migration stocks. This is plausible because the variability of the IV bases on j specific outflows of immigrants and not on the i specific component of bilateral migration (i.e. unrelated by construction to any country i specific shock). Another usual criticism of the standard shift-share instrument is the non-orthogonality of the initial distribution of immigrants used to allocate subsequent migration inflows. By using the distribution of immigrants in 1960 (35 years before the initial year of our estimations), this concern is reduced here.[15]

2.2.2 IV 2: a modified shift-share based instrumental variable controlling for feedback effects

As discussed in Jaeger et al. (2018), the country of origin mix for a given destination can be quite invariant over time (i.e. high persistence of immigrants' settlement across destination countries). This implies a high degree of autocorrelation in the shift-share instrument, that therefore captures both the short- and the long-term effect of immigration at destination country.[16] If the short- and the long-term effect have the opposite expected sign on the outcome variable (international competitiveness here), then the resulting estimate using the standard shift-share approach have an unclear interpretation. To address this potential bias, in the spirit of Jaeger et al. (2018), we remove from the predicted bilateral migration flows the long run component, i.e. the lag of the shift-share instrument.[17] Namely, we obtain the *predicted supply-driven* stock of immigrants in each country i

[15] Notice that being in a bilateral setting we cannot apply the procedure suggested by Goldsmith-Pinkham, Sorkin & Swift (2018) aiming at identifying the relevant "shares" driving the estimates. We reckon that the validity of our instruments relies on the fact that, conditional on local "demand pull factors" (i.e. $\widehat{\delta_{it}}$), the distribution of immigrant shares in 1960 is plausibly orthogonal to trade flows in 1995-2015.

[16] Jaeger et al. (2018) show a positive correlation between shift-share instrument and its lag equal to 0.96.

[17] We do not follow exactly the multiple instrumentation proposed by Jaeger et al. (2018) as a higher number of instruments in presence of a large set of fixed effects would produce low-efficient estimator. But, we definitely follow Jaeger et al. (2018) in the spirit by removing from the predicted immigration flow the long-term component $Immi\ Sh_{ij,60} * ln(Immi)_{jt-5}$.

based on the estimation of the following structural gravity model for migration:

$$Mig_{ijt} = \delta_{it}+\delta_{jt}+\delta_{ij}^{ij=ji}+\beta_1 Immi\ Sh_{ij,60}*ln(Mig)_{jt}+\beta_2 Immi\ Sh_{ij,60}*ln(Mig)_{jt-5}+\beta_3 Immi\ Sh_{ij,60}+\mu_{ijt} \quad (5)$$

where the variables have the same meaning as in equation 2 and we split the shift-share instrument into short- and long-term component, i.e. $Immi\ Sh_{ij,60} * ln(Immi)_{jt}$ and $Immi\ Sh_{ij,60} * ln(Immi)_{jt-5}$ respectively. From equation (5) we take the predicted value $\widehat{Immi_{ijt}}$ (fit of the regression) and subtract the destination-year fixed effect, the initial settlement of immigrants in 1960 and the long-term component ($Immi\ Sh_{ij,60} * ln(Immi)_{jt-5}$):

$$\widehat{AdjImmi}_{ijt}^{short} = \widehat{Immi_{ijt}} - \widehat{\delta_{it}} - \widehat{\beta_2} Immi\ Sh_{ij,60} * ln(Immi)_{jt-5} - \widehat{\beta_3} Immi\ Sh_{ij,60}. \quad (6)$$

By doing so, we purge the predicted *supply driven* bilateral stock of immigrants from any demand driven effect and from the long-term component highlighted by Jaeger et al. (2018). Finally, we use $\widehat{AdjImmi}_{ijt}^{short}$ to build an IV for the Birthplace Diversity index and to instrument Mig_{ijt} and its interaction with TS_{ijk}.

2.2.3 IV 3: a natural disaster based instrumental variable

An alternative IV to solve the endogeneity concern has been inspired by the natural experiment approach literature. Natural disasters (tsunami, earthquakes, floods, etc) have been proven to be one of the main causes of human mobility in many developing countries (Gray & Mueller (2012) and Beine & Parsons (2017)).[18] We therefore compute the Birthplace Diversity index based on the *predicted supply-driven* stocks of immigrants induced by countries that experienced (at least one) natural disaster in the pre-treatment period, i.e 1985-1990.[19] To do this, we use the predicted number of immigrant ($\widehat{Immi_{ijt}}$) from equation (2) net of destination-year fixed effects, but only from the sub-sample of origins j with at least one natural disaster over the period 1985-1990. This variable is used also to instrument Mig_{ijt} and its interaction with TS_{ijk}.

Notice that for a precise calculation of the diversity measure, we cannot omit migrants communities from the rest of other origins (i.e. countries that did not experience natural disasters in the period 1985-1990). Indeed, if we compute the BD index using the $\widehat{AdjImmi}_{ijt}$ from the subsample of countries that have experienced natural disasters, we would miss a consistent number of origins and the resulting BD would be strongly biased.[20] We therefore use the bilateral stock of immigrants in 1960 to include the origin countries that did not experience natural disasters in the period 1985-1990. This formula describes this alternative instrumental variable:

[18] Beine & Parsons (2017) show that while natural disaster *per se* have a null (slightly negative) effect on emigration, they considerably boost emigration towards destinations with low migration cost.

[19] See Appendix A for details on natural disaster data used in the paper.

[20] The total number of countries j affected by a catastrophic natural event in our sample are 41.

$$BD_{it}^{\widehat{PPML},ND} = 1 - \sum_{j=1}^{J} \left[\left(I_{j,85-90} \frac{\widehat{AdjImmi}_{ijt}}{\sum_{j=1}^{J} \widehat{AdjImmi}_{ijt}} \right)^2 + \left((1 - I_{j,85-90}) \frac{Immi_{ij,60}}{\sum_{j=1}^{J} Immi_{ij,60}} \right)^2 \right] \tag{7}$$

where $Immi_{ij,60}$ is the stock of immigrants in i from origin j in 1960, and $I_{j,85-90}$ is a dummy variable equal to one if country j experienced a natural disaster in the period 1985-1990. The first term of the squared bracket in equation (7) activates for origins with natural disaster and uses the predicted supply-driven component of bilateral migration ($\widehat{AdjImmi}_{ijt}$) to compute the BD index. The second term of the squared bracket activates for countries without natural disaster and uses bilateral stock of immigrants in 1960 to compute the squared share of immigrant from j (taken in 1960 to avoid any endogeneity concern in the time variation of the stocks of migrants coming form disasters-free origins). Figure 2 qualitatively demonstrate the identification strategy used in this case. Figure 2 show a clear positive relationship between natural disaster events occurred in the period 1985-1990 in origin countries j and subsequent outward migration (univariate R-square 0.79).

2.3 Bilateral Results

In this section we present results on country pair-sectors specific trade margins based on equation (1). We first show results using simple OLS estimator (benchmark), then we move to a 2SLS approach aiming at addressing the potential endogeneity issue with the settlement of immigrants across destinations.

2.3.1 Baseline Results

Results using OLS estimator on equation 1 are reported in table 3 and 4 respectively for intensive and extensive margins of trade. The structure of the two tables is similar. In columns (1)-(4) we show results by introducing progressively additional control variables in the baseline specification 1; in columns (5)-(7) we adopt the full baseline specification and standardize the three variables of interest (Mig_{ijt}, $Mig_{ijt} * TS_{ijk}$ and BD_{it}) to shed light on the relative magnitude of the three channels. By doing so, the magnitude of coefficients on the three effects may be directly compared (as all expressed in standard deviation units from the mean). In columns (6)-(7) we use alternative definitions for TS_{ijk} as a sensitive analysis on the threshold used to consider a country having comparative advantage in a given sector k.

In line with the previous literature, we find that migrants from j and residing i boost the exports from i to j - see columns (1)-(7) in table 3. This is the standard transaction cost channel highlighted in several previous papers: the presence of migrants from a specific origin provides additional information to firms at destination on how to export in j (i.e. consumers' taste, regulation, distribution channel). In particular (using our preferred specification in column 4 of table 3), a 10 percent increase in the stock of immigrant from j boosts the export of i to j by 0.62 percent. This effect is magnified if the country of origin of immigrants (j) has a comparative advantage in sector k (i.e. Balassa index above 1) and country i has not (i.e. Balassa index below 1). This is the knowledge diffusion channel: when migrants in i originate from a country having a comparative advantage in sector k, then the transaction cost effect increases substantially. See column 4 in table 3. Results hold by adopting standardized variables - see column (5). The knowledge diffusion effect is robust to alternative definition of TS_{ijt} dummy, equal to one if the country of origin of immigrants has a Balassa index on sector k bigger that 1.25 (1.5) and destination i has Balassa index below 0.75 (0.5). See columns 6 and 7.

On top of the transaction cost and knowledge diffusion channel, the diversity in the origins of migrants - BD_{it} - has a positive and statistically significant effect on the intensive margin of exports in all the specifications (see columns 1 to 7). Since in the computation of BD_{it} we consider the bilateral stock of immigrants from j to i, it might be the case that the diversity channel and the transaction cost channel blur together. For this reason, in table A3 we include a birthplace diversity index computed by getting rid of the bilateral specific migration stocks - $BD_{\mathbf{ijt}}$.[21] Also for this (slightly) different version of the diversity variable, the effect on the intensive margin of trade is positive and significant.

Notice that the control variables included in equation 1 all have the expected sign. While distance between i and j reduces bilateral exports, common language, border and colony all have positive effect on exports. The presence of a Regional Trade Agreement in force boosts bilateral exports. Applied tariffs reduce bilateral trade, in particular a 1% increase in bilateral (sector specific) tariff reduces exports by 2.2%.[22] Also, the fact that country j (origin of immigrants) has a comparative advantage in sector k, while country i has not (dummy TS_{ijk}), negatively affect the exports of country i to j. Coherently with the gravity model for trade, a higher export capacity of country i (bins of total i's exports toward extra j's region destinations) has a positive effect on bilateral specific exports. Interestingly, the positive coefficient on $Ln(Emig)$ suggests that the presence of emigrants from i to j stimulates the import demand of j from i - preference channel in import demand.

< Table 3 about here >

In table 4 we report the estimation results for the extensive margin (i.e. export participation). The structure of the table is the same as table 3. The transaction cost and knowledge diffusion channel have both a positive and significant effect on the extensive margin. The higher the stock of immigrant in i from origin j, the higher the probability that country i exports towards j - see table 4. Using our preferred specification in column (4), a 10 percent increase in the bilateral stock of immigrants increase the probability for i to export in j by 0.11. This effect is magnified if the migrants' origin country j has a comparative advantage in sector k, i.e. when the dummy TS_{ijk} is equal to one. Birthplace diversity has a significant positive effect also on the extensive

[21] Where $BD_{\mathbf{ijt}} = 1 - \sum_{k \neq j}^{J} s_{ikt}^2$

[22] The coefficient on tariff elasticity is coherent with many previous studies. See Berthou & Fontagné (2016), Fitzgerald & Haller (2018) and Buono & Lalanne (2012).

margin of trade.[23]. In this case, one standard deviation increase in the Birthplace Diversity index implies a 0.64 percent increase in the probability of observing an export flows between i and j in sector k at time t. In columns (5)-(7) of table 4 we use standardized variables for both network, knowledge diffusion and diversity channels (i.e. Mig_{ijt}; Mig_{ijt} and BD_{it}) to check the relative size of the three channels. On the extensive margin the effect of diversity is only one-fifth of the transaction cost channel.

Notice that also in this case all control variables have the expected sign. The role of tariff here is less relevant than in the case of intensive margin because a change in the variable cost of trade is expected to have a null/reduced effect on the extensive margin of trade.

$<$ Table 4 about here $>$

2.3.2 2SLS Results

Based on the three instrumental variables presented respectively in sections 2.2.1, 2.2.2 and 2.2.3, in table 5 we present the results using 2SLS approach to solve the endogeneity problem concerning the variable Mig_{ijt}, BD_{it} and $Mig_{ijt} * TS_{ijk}$. The idea is instrumenting the bilateral stock of immigrants (and its interaction with the TS_{ijk} dummy) with the predicted supply driven bilateral migration flows. Then we instrument the Birthplace Diversity index BD_{ijt} with a diversity measure based on the *predicted* (rather than observed) bilateral migration stocks. All the three instruments are based on the exogenous variation of the supply of immigrants in the origin country j. Any i specific labor demand effect has been removed from the IVs. See detailed description in section 2.2.

In column (1)-(3) of table 5 we report 2SLS estimations on the intensive margin of exports by using each of the three instruments described above. Irrespectively of the IV used, the effect of the bilateral stock of migrants is positive and significant (network effect) and magnified if immigrant originate from countries having a comparative advantage in sector k (knowledge diffusion). Also in the case of 2SLS, the positive effect of birthplace diversity remains positive and statistically significant across the three instruments. In columns (4)-(6) of table 5 we report 2SLS estimations on the extensive margin of exports. Results confirm what we obtained with OLS: the three channels, i.e. network, knowledge diffusion and diversity, positively affect the export participation probability of country i in market kj.

$<$ Table 5 about here $>$

The relevance of the three instrumental variables is reported at the bottom of table 5. The three instruments are highly relevant and do not suffer any problem of potential weak instrument (F-stat above 10). The validity of the three instruments cannot be tested with a Sargan test (exact identified model), but being based exclusively

[23] The results are robust to the way we compute the diversity variable, in table A4 we report the results using BD_{ijt}, i.e. excluding j specific immigrants in the calculation of the Herfindhal Index

on the supply of immigrants from country j are plausibly valid (i.e. unrelated by construction to any country i specific shock). Since the labor demand component of country i has been explicitly removed in all the three IVs, the exclusion restriction here is that immigrants residing in i because "pushed away" from country j affect the export performances of country i only through their effect on bilateral migration stocks. In other words, bilateral export performances are expected to be orthogonal with respect to the push component of emigration from j. Notice that the allocation of exogenous "push" migration is made on the distribution of immigrants from j across destinations i in 1960 - see equation 2. With a lag of thirty years we are confident about the validity of our IVs.

3 The heterogeneous effect of Birthplace Diversity

The previous section showed the statistic significant effect of birthplace diversity channel on the export performances of countries by explicitly controlling for the two other channels through which migrants may affect exports, i.e network and knowledge diffusion channels. In this section we focus on the role of birthplace diversity in affecting the export performance of the destination country, and test the channel through which the diversity in the country of origin of immigrants has an impact of trade.

3.1 From bilateral to country specific estimation: how and why

The three migration related channels discussed above have different identifying variation. While the diversity channel is country-year specific (it does not varies across origins j), the transaction cost and the knowledge diffusion channels are respectively country pair and country pair-sector specific. This may drive to a potential aggregation bias in estimations. A more compelling way of analyzing the trade effect of birthplace diversity is estimating it at the same level of aggregation as the variable BD_{it}. Therefore, in the present section we focus on the role of birthplace diversity by aggregating data at the country-(sector)-year level. While the diversity variable is actually country-year specific, we keep the sector dimension k to test the mechanism at play. By testing the heterogeneous effect of diversity based on several sector characteristics, we shed light on the mechanism through which diversity affects the international competitiveness of a country.

We aggregate trade related variables at country-sector-year level (i.e. total exports, number of destination served by country i in sector k, number of products exported by country i in sector k), and test the effect of birthplace diversity $BD_{i,t}$, controlling for the total stock of immigrants residing in country i with the aim of capturing the affect of other migration-related shocks (other than diversity). The advantage of this specification is to use the same level of aggregation of the variable of interest interacted with a bunch of sector characteristics.

Based on the theoretical model by Maggi & Grossman (2000), we do expect birthplace diversity positively affecting the aggregate export performances of country i, with a magnified effect for sectors whose production process is characterized by sub-modular production function, where the presence of a high-talented workers is more likely to be paired with workers at the opposite end of the ability distribution, i.e. sectors where a more diverse workforce is extremely important. Maggi & Grossman (2000) propose a theoretical model in which two countries have different workforce ability distributions (one more disperse than the other), and produce two types of goods: (i) one characterized by a *super-modular* production function; and the other (ii) by a *sub-modular* production function. In presence of super-modular production function, performing better one task raises the marginal value of a better performance in the other task. This is the case of industries in which the precision in a long sequence of production steps contributes to the success of the overall production (example the automotive industry). In presence of sub-modular production function, performing better one task mitigates the need for better performances in the other tasks. This is the case of industries requiring creativity and problem solving abilities (such as fashion, design and cultural goods), where the overall success of the production process strongly depends on the presence of an extreme brilliant worker in production, i.e. when the marginal value of having a more able worker increases when the ability of the other co-workers in the same task are less. Under these assumptions, Maggi & Grossman (2000) show that the country with a more diverse distribution of worker abilities (i.e. birthplace diverse workforce for our purposes), by having higher possibilities for matching extreme brilliant with more modest workers in the same task, will have a comparative advantage in the sector characterized by sub-modular technology (where creativity and problem-solving are more needed). Conversely, the country having less disperse worker ability distribution will specialize in the production of products characterized by super-modular technology.

Assuming that immigrants coming from different origins are imperfectly substitute in production (they have different characteristics), host countries with a higher birthplace diversity index have a more disperse distribution of worker abilities. We can therefore conclude that:

Conjecture 1 *Host countries with higher index of Birthplace Diversity will export more in sectors characterized by sub-modular technology (i.e. cognitive and problem solving intensive sectors).*

3.2 Empirical Strategy

With country-sector-year aggregated data at hand we run the following *baseline* econometric specification:

$$y_{ikt} = \beta_1 BD_{i,t} + \beta_2 Mig_{i,t} + \beta_3 X_{it} + \theta_i + \theta_{k,t} + \epsilon_{i,k,t} \quad (8)$$

where the dependent variable y_{ikt} is in turn: (i) total exports (in ln) of country i in sector k and time t; (ii) the number of destination countries reached by i on sector k and time t (in ln); and (iii) the number of exported products (HS 6-digit products) within a sector k across all destinations (in ln). $BD_{i,t}$ is the birthplace diversity measure for country i at time t as described above, $Mig_{i,t}$ is the total stock of immigrants (in ln) in country j at time t and aims at controlling for the pure effect of migrants on the total export of the country (network and knowledge diffusion). The set of control variables X_{ij} includes the number of preferential trade agreements in force for country i (proxy for average market access), and the GDP of the country (in ln). Since the main variable of interest (BD_{it}) if country-year specific, we cannot include country-year fixed effect, thus in equation (8) we simply control for country (θ_i) and sector-by-year (θ_{kt}) fixed effects.

Given the aggregate nature of equation (8), the skeptical reader may not be convinced that the inclusion of $Mig_{i,t}$ would control for the transaction cost channel in a compelling way, so that the coefficient on BD_{it} would also capture some of the effect of bilateral migration (i.e. transaction cost channel). To control for this potential effect we estimate equation (8) on a further dependent variable: the *estimated* Revealed Comparative Advantage *net* of the bilateral migration effect. In the vein of Costinot et al. (2012), we compute a synthetic measure of the export performance of country i in sector k and time t conditioned on the effect of bilateral migration. This is obtained by keeping the country-sector-year fixed effect from the following auxiliary regression:

$$Export_{ijkt} = \gamma_{ikt} + \gamma_{jkt} + \gamma_{ijk} + \beta Mig_{ijt} + \beta Mig_{ijt} * (TS)_{ijk} + PTA_{ijt} + \mu_{ijkt} \tag{9}$$

where γ_{ikt}, γ_{jkt} and γ_{ijk} are respectively exporter-sector-year, importer-sector-year and country pair-sector fixed effect; whereas Mig_{ijt} and $Mig_{ijt} * (TS)_{ijk}$ control for the effect of bilateral migration on exports (both transaction cost and technological diffusion channels).[24] From eq.(9) we recover the estimated exporter-sector-year fixed effects ($\widehat{\gamma_{ikt}}$) which represents a synthetic measure of the export performance (RCA) of country i in sector k at time t. In a Ricardian type model of trade, Costinot et al. (2012) show that exporter-sector-year fixed effects from a reduced form model as in equation (9) exactly mirror the ex-ante Ricardian comparative advantage of a country. Being a synthetic measure of *revealed* comparative advantage, we use this variable as an alternative way of estimating the international competitiveness effect of birthplace diversity. This measure of revealed comparative advantage is made freely available for the interested scholars and practitioners in the CEPII web page (see appendix section C for a description of the dataset). Notice that equation (9) is estimated using a PPML model to control for the incidence of zeros and heteroscedasticity of trade flows as suggested in Silva & Tenreyro (2006). In equation (9) we also include a dummy for the presence of a common Preferential Trade Agreement controlling for any preferential market access boosting bilateral trade. Being conditioned on bilateral migration, the RCA measure built as described above, is purged from any migration driven effect. In this case the coefficient of BD_{it} will precisely capture the effect of diversity and not any other migration related

[24]Notice that the direct effect of TS_{ijk} is captured by the country-sector-pair fixed effects γ_{ijk}.

effect.

In this section we want to go deeper in understanding the mechanism through which birthplace diversity may affect the international competitiveness of a country. To this end we test the heterogeneous effect of birthplace diversity by interacting the variable BD_{it} with several sectors specific characteristics - SH_k - such as: (i) skill intensity of the sector (dummy variable based on UNCTAD classification); (ii) technology intensity of the sector (dummy variable based on UNCTAD classification); and (iii) differentiated sector dummy (based on Rauch classification). Since all these sector characteristics are provided at SITC 3-digit level, our aggregated estimations adopt the SITC classification. The idea is that the birthplace diversity channel should work in particular for sectors intensive in complex tasks where problem solving capabilities are relative more important (where a more diverse distribution of worker types is more valuable). Hong & Page (2001) theoretically show that a group of more diverse problem solvers may perform better than a group of homogeneous but more ables problem solvers.[25] To test the heterogeneous impact of Birthplace Diversity we adopt the following empirical specification:

$$ln(y_{ikt}) = \beta_1 BD_{i,t} * (SH)_k + \theta_{it} + \theta_{kt} + \epsilon_{i,k,t} \tag{10}$$

this "augmented" version of eq (8) allows us include country-year and sector-year fixed effects to address any concern of omitted variable bias. Our interest is now on the interaction term $BD_{i,t} * (SH)_k$ and the inclusion of θ_{it} and θ_{kt} fixed effects would control for all country-year and sector-year specific factors. Notice that any concern that high-exporting countries attract immigrants from a varieties of origin is perfectly controlled in this specification by the inclusion of country-year fixed effects. The drawback of the specification (10) is the impossibility to obtain the average effect of BD_{it} because perfectly collinear with fixed effects. On this respect the reader can find informative the results on BD_{it} estimated with a less conservative set of fixed effects as showed in equation (8).

3.3 Aggregate Results

This section presents the results concerning the heterogeneous effect of birthplace diversity on the country-sector-year aggregated export performances of country i. The structure of tables presented in what follows is the same across the four export performances analyzed here, i.e. intensive margin, number of products, number of destinations, and synthetic measure of comparative advantage á la Costinot et al. (2012). The first column aims at presenting the effect of Birthplace Diversity on the average sector (country-year fixed effects not

[25] People migrating from different origin countries bring at their destination country a diverse set of skills, experiences, ideas, expertise and problem-solving capabilities that may be useful to improve the efficiency of the production process and the overall performance of the firm. See Lazear (1999); Hong & Page (2001); Horwitz & Horwitz (2007).

included because perfect collinear with BD_{it}). In column 2, with the same set of fixed effects, we introduce the interaction between Birthplace Diversity and a dummy on whether the sector k covers differentiate (rather than homogeneous) products. In columns (3)-(6) we estimate the augmented equation (10) with both country-year and sector-year fixed effects and we focus on the heterogeneous effect of birthplace diversity, i.e. its interaction with the three sectors characteristics detailed above.

In table 6 we report results on total export of country i on sector k time t. On average, a more diverse set of migrants' origins positively affect the export of the destination country i on the average sector (see first line in columns 1 and 2). This effect is magnified if the sector covers differentiated products. When we include country-year and sector-year fixed effects (and therefore solve any omitted variable concern), we find that the effect of diversity is particularly relevant for skill and high technology intensive sectors (columns 4 and 5), and for differentiated sectors.

< Table 6 about here >

In table 7 we report results on the number of destinations served by country i on sector k. The diversity in the country of birth of immigrants has a positive and significant effect on the destination extensive margin of export of country i. In particular, one standard deviation increase in the BD_{it} index implies a 1.52 percent increase in the number of destinations served by country i in sector k. Interestingly, this effect is magnified for skill intensive and high technology sectors.

< Table 7 about here >

In table 8 we report results on the synthetic measure of export performance described in section 3. We find that (on average) birthplace diversity has a positive effect on the revealed comparative advantage measure (purged from the effect of bilateral migration). The effect of diversity on the aggregate international competitiveness of country i is magnified for differentiated sectors (columns 2 and 3), for high skill (column 4) and high-technology intensive sectors (column 5).

< Table 8 about here >

The evidence discussed so far points to the fact that birthplace diversity is particularly beneficial for sectors in which the varieties of ideas and the problem solving capabilities are relevant (differentiated, skill and technology intensive sectors). As a further test on the channel through which diversity may affect the international competitiveness of countries, we use the job complexity of the sector. In this case we adopt the SIC sector classification to be coherent with the job complexity measure as provided by Costinot (2009). This measure approximates the complexity of tasks needed in each SIC (3 digit) sector. We find strong evidence that birthplace diversity is beneficial in particular for sectors with high job complexity - see columns (1)-(5) in table 9.

< Table 9 about here >

All in all we may conclude that the diversity in the country of origin of immigrants translates into better export performances through improved problem solving capability. This is testified by the fact that BD_{it} is more beneficial in sectors where ideas and problem solving capabilities are relatively more relevant (skill, technology intensive and job-complex sectors). This is coherent with the theoretical predictions in Hong & Page (2001) and Maggi & Grossman (2000). Hong & Page (2001) show that a group of more diverse problem solvers may perform better than a group of homogeneous but more ables problem solvers. Maggi & Grossman (2000) demonstrate that a more disperse distribution of worker ability is more beneficial in sectors intensive in problem solving capabilities.

Therefore in the next sections we dig more into this dimension and test the effect of birthplace diversity on *manual* vs *abstract* intensive sectors with the aim of using the effect of Birthplace Diversity in manual intensive sectors as a placebo test.

3.4 Placebo test: Birthplace Diversity and manual intensity of sectors

To further strengthen the above mentioned channel, this section provides a placebo test by analyzing the effect of Birthplace Diversity on abstract vs manual intensive sectors. Ideas and problem solving capabilities are not relevant in manual intensive sectors, so if the mechanism highlighted above is true, birthplace diversity is expected to have a null/reduced effect on export performances for manual intensive sectors. In table 10 we compare the effect of diversity on manual vs abstract intensive sectors. In line with the intuition developed so far, diversity has positive and significant effect on the abstract intensive sectors only and not on manual sectors. Compare columns (1)-(4) with (5)-(8).

< Table 10 about here >

4 Conclusion

Immigration is often considered a simple positive labor supply shock in the host countries. This paper shows that this is far from being the case. For the first time in the literature (to the best of our knowledge) we test in a unified empirical framework the effect of migration on the three channels through which it may affect the international competitiveness in the host country, i.e. transaction, knowledge diffusion and diversity channel. We show that these three channels are contemporaneously at play. An increase in the bilateral stock of immigrant boosts the host country's exports toward the country of origin of immigrants (*transaction cost* channel). This effect is magnified if immigrants arrive from an origin country having a comparative advantage in a specific sectors (and the host country has not) - *knowledge diffusion* channel.

In the second part of the paper we focused on the heterogeneous effect of Birthplace Diversity on several trade-related outcomes. In line with Maggi & Grossman (2000), we find that diversity in the country of origin of immigrants is particularly beneficial for sectors intensive in problem solving capabilities (abstract intensive, skill and job-complex intensive sectors.)

Tables and Figures

Table 1: Migration Stocks, Knowledge Diffusion and Diversity.

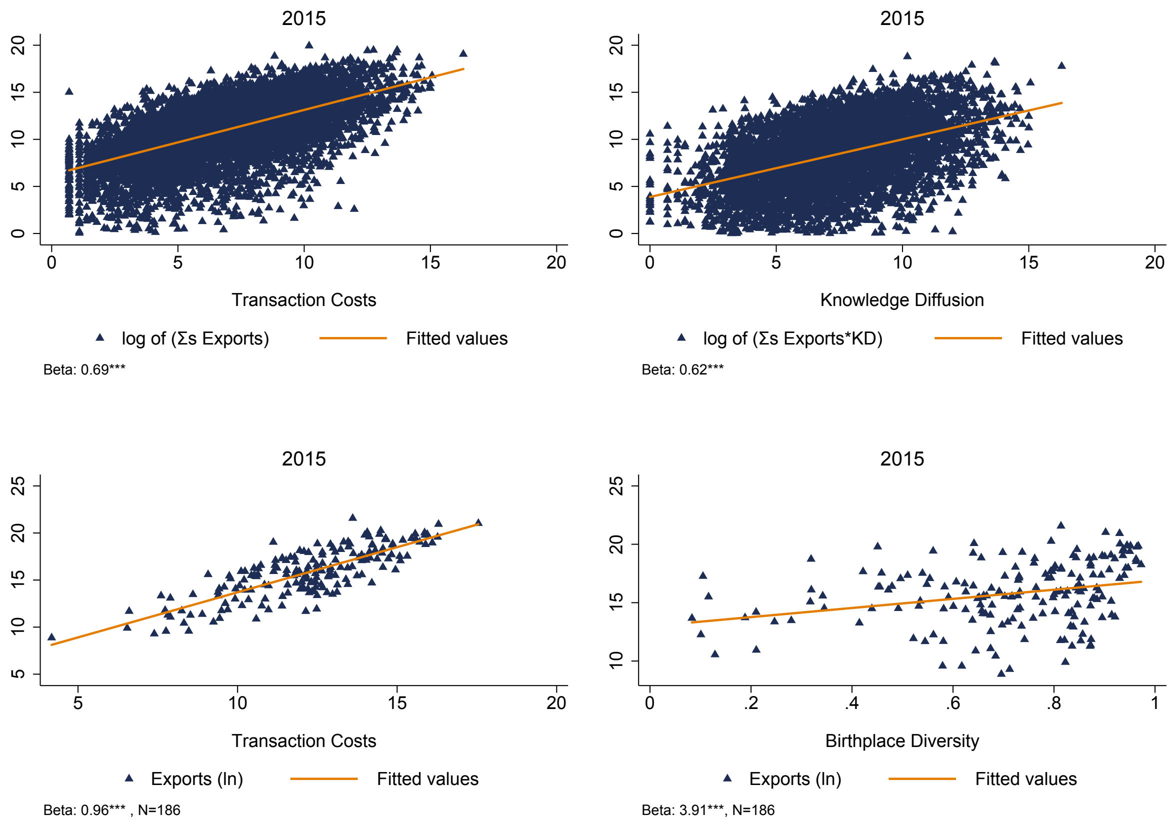
	Year 1995			**Year 2005**			**Year 2015**		
	Mig Stock Tot	Mig Stock $\overline{TS_{ij}} * Tot$	BD	Mig Stock Tot	Mig Stock $\overline{TS_{ij}} * Tot$	BD	Mig Stock $\overline{TS_{ij}} * Tot$	Mig Stock	BD
United States	24.53	5.40	0.91	34.53	7.60	0.90	41.77	9.23	0.90
Germany	7.12	1.58	0.92	10.11	2.22	0.93	11.97	2.63	0.93
Russia	11.92	2.45	0.84	11.66	2.35	0.85	11.64	2.33	0.85
Saudi Arabia	4.95	1.34	0.90	6.28	1.70	0.90	9.84	2.66	0.90
UK	4.09	0.88	0.95	5.83	1.33	0.97	8.42	1.98	0.97
UAE	1.77	0.52	0.81	3.17	0.94	0.78	7.98	2.39	0.76
Canada	4.85	1.23	0.96	6.06	1.51	0.96	7.81	1.93	0.96
France	6.08	1.13	0.93	6.71	1.23	0.92	7.75	1.44	0.93
Australia	4.11	0.89	0.91	4.80	1.02	0.92	6.65	1.39	0.94
Spain	1.01	0.18	0.93	4.10	0.66	0.95	5.85	0.96	0.95

Table 2: In-sample descriptive statistics

Variable	N	Mean	Sd	p25	p50	p75
Bilateral Sample:						
log(Exports)	1,737,471	5.9	2.9	3.8	5.8	7.9
# Destinations	3,000,682	0.8	0.4	1.0	1.0	1.0
log(Immigrants)	3,000,682	4.7	4.1	0.0	5.3	8.1
TS	3,000,682	0.2	0.4	0.0	0.0	0.0
Diversity	3,000,682	0.8	0.2	0.7	0.8	0.9
Aggregate Sample:						
log(Exports)	171,028	7.7	4.0	4.8	7.8	10.7
Diversity	171,028	0.7	0.2	0.6	0.8	0.9
log(Immigrants)	171,028	12.5	2.0	11.2	12.5	13.9

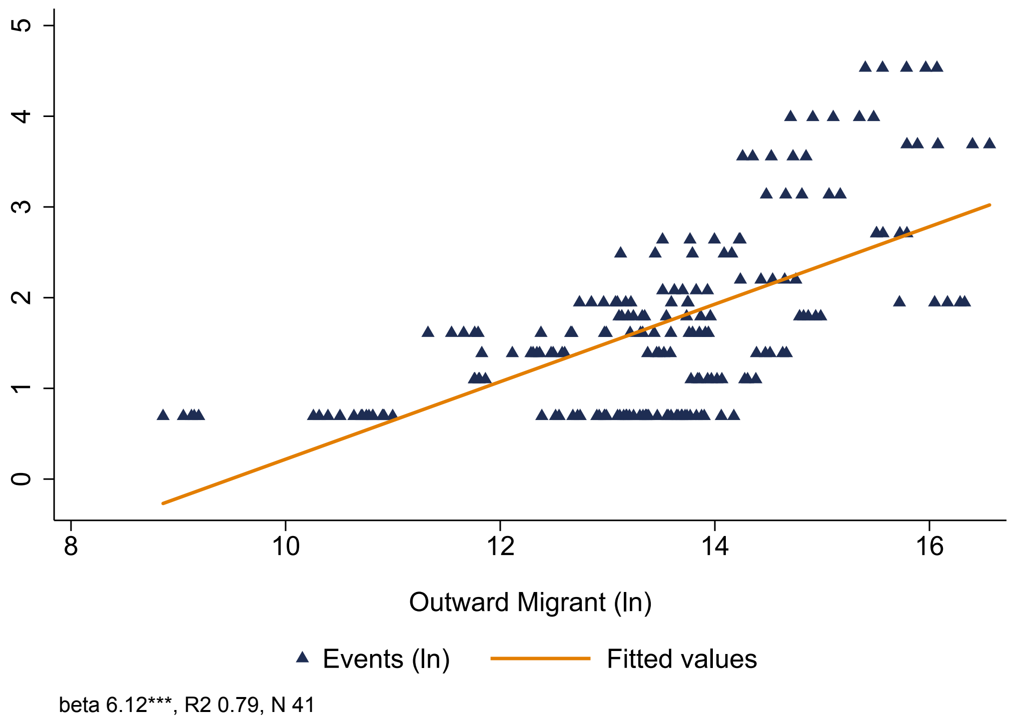
Note: In sample descriptive statistics. 5-year intervals from 1995 to 2015.

Figure 1: Un-conditional correlations for the different channels. Year 2015.



Note: from Equation 1 transaction costs are captured by $ln(Mig_{ijt})$ - top-left scatter; Knowledge diffusion by the interaction $Mig_{ijt} * (TS)_{ijk}$ - top-right scatter; Birthplace diversity by BD_{it} - bottom-right scatter. Scatter plot between aggregate country specific exports and immigrant stocks are reported in the bottom-left scatter.
Source: Authors calculation on BACI (CEPII) and United Nations (2015) data.

Figure 2: Natural Disasters and Outward Migration



Note: regression of outward migration stock on number of natural disasters in the previous decade (log-log specification, 41 counties).
Source: EM-DAT: The Emergency Events Database.

Table 3: Baseline estimation results. Intensive margin. OLS regressions.

Dep var	Ln(export)$\|_{export(t-1)>0}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mig$_{ijt}$	0.134***	0.140***	0.096***	0.062***	0.254***	0.258***	0.259***
	(0.008)	(0.008)	(0.004)	(0.004)	(0.017)	(0.017)	(0.017)
Mig$_{ijt} * TS_{ijk}$	0.058***	0.059***	0.039***	0.040***	0.164***	0.125***	0.109***
	(0.004)	(0.004)	(0.003)	(0.002)	(0.010)	(0.010)	(0.013)
TS$_{ijk}$	-1.983***	-1.984***	-0.927***	-0.972***	-0.741***	-0.706***	-0.692***
	(0.040)	(0.038)	(0.026)	(0.026)	(0.021)	(0.021)	(0.020)
BD$_{it}$	0.563**	0.520*	0.278**	0.255**	0.045**	0.044**	0.044**
	(0.274)	(0.280)	(0.117)	(0.111)	(0.019)	(0.019)	(0.019)
ln(dist)$_{ij}$	-0.609***	-0.608***	-1.131***	-1.057***	-1.057***	-1.062***	-1.067***
	(0.035)	(0.035)	(0.029)	(0.028)	(0.028)	(0.028)	(0.028)
Time Diff$_{ij}$	0.041***	0.043***	0.022***	0.024***	0.024***	0.024***	0.024***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Colony$_{ij}$	0.317***	0.313***	0.401***	0.209***	0.209***	0.203***	0.200***
	(0.090)	(0.090)	(0.065)	(0.057)	(0.057)	(0.057)	(0.057)
Language$_{ij}$	0.020	0.014	0.328***	0.262***	0.262***	0.264***	0.267***
	(0.056)	(0.055)	(0.035)	(0.031)	(0.031)	(0.031)	(0.031)
Border$_{ij}$	0.595***	0.591***	0.334***	0.216***	0.216***	0.216***	0.217***
	(0.050)	(0.050)	(0.041)	(0.042)	(0.042)	(0.042)	(0.042)
RTA$_{ijt}$	0.288***	0.270***	0.403***	0.364***	0.364***	0.366***	0.367***
	(0.040)	(0.043)	(0.031)	(0.030)	(0.030)	(0.030)	(0.030)
ln(1 + Tariff)$_{ijt}$	-1.736***	-1.829***	-2.269***	-2.213***	-2.213***	-2.207***	-2.213***
	(0.155)	(0.160)	(0.121)	(0.119)	(0.119)	(0.118)	(0.119)
ln(Remot)$_{it}$		4.742***	1.527***	1.459***	1.459***	1.421***	1.395***
		(0.824)	(0.357)	(0.343)	(0.343)	(0.340)	(0.339)
ln(Trade)$_{ikt,-jr}^{IIquart}$			1.597***	1.589***	1.589***	1.598***	1.605***
			(0.027)	(0.027)	(0.027)	(0.026)	(0.026)
ln(Trade)$_{ikt,-jr}^{IIIquart}$			2.763***	2.734***	2.734***	2.752***	2.766***
			(0.039)	(0.038)	(0.038)	(0.037)	(0.037)
ln(Trade)$_{ikt,-jr}^{IVquart}$			4.343***	4.269***	4.269***	4.298***	4.326***
			(0.055)	(0.051)	(0.051)	(0.051)	(0.050)
ln(Emig)$_{ijt}$				0.094***	0.094***	0.093***	0.092***
				(0.004)	(0.004)	(0.004)	(0.004)
Definition TS:	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 0.75$	$BI_{ik} < 0.5$
	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1.25$	$BI_{jk} > 1.5$
Std Var	No	No	No	No	Yes	Yes	Yes
jkt FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
rlt FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,706,473	1,706,473	1,706,473	1,706,473	1,706,473	1,706,473	1,706,473
R-squared	0.404	0.408	0.536	0.542	0.542	0.541	0.540

Note: In all regressions standard errors in parentheses are clustered at exporter country-year level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 186 each one observed for 5 time periods (1995, 2000, 2005, 2010, 2015).

Table 4: Baseline estimation results. Extensive margin. OLS regressions.

Dep var	$Dummy = 1$ if $Export_{ijkt} > 0$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mig$_{ijt}$	0.017***	0.017***	0.012***	0.008***	0.035***	0.035***	0.035***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Mig$_{ijt} * TS_{ijk}$	0.006***	0.006***	0.004***	0.004***	0.015***	0.014***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
TS$_{ijk}$	-0.127***	-0.127***	-0.037***	-0.040***	-0.023***	-0.022***	-0.025***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
BD$_{it}$	0.095***	0.083***	0.036**	0.034**	0.006**	0.006**	0.006**
	(0.027)	(0.027)	(0.015)	(0.015)	(0.003)	(0.003)	(0.003)
ln(dist)$_{ij}$	-0.040***	-0.039***	-0.085***	-0.077***	-0.077***	-0.077***	-0.077***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Time Diff$_{ij}$	0.005***	0.005***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Colony$_{ij}$	0.024***	0.023***	0.041***	0.023***	0.023***	0.023***	0.023***
	(0.008)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Language$_{ij}$	0.008	0.009*	0.035***	0.028***	0.028***	0.028***	0.028***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Border$_{ij}$	-0.002	-0.001	-0.014***	-0.028***	-0.028***	-0.028***	-0.028***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
RTA$_{ijt}$	0.033***	0.033***	0.039***	0.035***	0.035***	0.035***	0.035***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ln(1 + Tariff)$_{ijt}$	-0.184***	-0.193***	-0.213***	-0.209***	-0.209***	-0.209***	-0.209***
	(0.014)	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
ln(Remot)$_{it}$		0.456***	0.163***	0.152***	0.152***	0.151***	0.151***
		(0.070)	(0.039)	(0.040)	(0.040)	(0.040)	(0.040)
ln(Trade)$^{IIquart}_{ikt,-jr}$			0.170***	0.167***	0.167***	0.167***	0.167***
			(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
ln(Trade)$^{IIIquart}_{ikt,-jr}$			0.265***	0.260***	0.260***	0.261***	0.261***
			(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
ln(Trade)$^{IVquart}_{ikt,-jr}$			0.359***	0.349***	0.349***	0.350***	0.350***
			(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ln(Emig)$_{ijt}$				0.009***	0.009***	0.009***	0.009***
				(0.000)	(0.000)	(0.000)	(0.000)
Definition TS:	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 0.75$	$BI_{ik} < 0.5$
	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1.25$	$BI_{jk} > 1.5$
Std Var	No	No	No	No	Yes	Yes	Yes
jkt FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
rlt FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,000,682	3,000,682	3,000,682	3,000,682	3,000,682	3,000,682	3,000,682
R-squared	0.173	0.175	0.221	0.223	0.223	0.223	0.223

Note: In all regressions standard errors in parentheses are clustered at exporter country-year level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 195 each one observed for 5 time periods (1995, 2000, 2005, 2010, 2015).

Table 5: Baseline estimation results. 2SLS regressions.

Dep var	$\text{Ln(export)}\|_{export(t-1)>0}$			$Dummy = 1$ if $Export_{ijkt} > 0$		
	(1)	(2)	(3)	(4)	(5)	(6)
Mig_{ijt}	0.104***	0.105***	0.137***	0.011***	0.011***	0.014***
	(0.008)	(0.008)	(0.011)	(0.001)	(0.001)	(0.001)
$\text{Mig}_{ijt} * TS_{ijk}$	0.036***	0.036***	0.030***	0.005***	0.005***	0.004***
	(0.003)	(0.003)	(0.003)	(0.000)	(0.000)	(0.000)
TS_{ijk}	-0.956***	-0.957***	-0.931***	-0.046***	-0.046***	-0.045***
	(0.029)	(0.029)	(0.031)	(0.003)	(0.003)	(0.003)
BD_{it}	0.549***	0.553***	0.651***	0.050**	0.051**	0.089***
	(0.163)	(0.163)	(0.242)	(0.020)	(0.020)	(0.032)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
IV	Predicted Supply	No feedback effect	Natural Disaster	Predicted Supply	No feedback effect	Natural Disaster
IV Mig_{ijt}	0.553***	0.555***	0.411***	0.580***	0.581***	0.436***
IV Mig_{ijt}*TS_{ijt}	0.761***	0.761***	0.750***	0.781***	0.781***	0.774***
IV BD_{it}	0.870***	0.869***	0.551***	0.887***	0.886***	0.549***
F-stat	360.5	362.7	194.3	477.9	482.2	250.7
Observations	1,706,473	1,706,473	1,706,473	3,000,682	3,000,682	3,000,682
R-squared	0.540	0.540	0.537	0.223	0.223	0.223

Note: In all regressions standard errors in parentheses are clustered at exporter country-year level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 195 each one observed for 5 time periods (1995, 2000, 2005, 2010, 2015).

Table 6: Country-sector aggregate results. Total export estimations

Dep var	$Export_{ikt}$ (ln)				
	(1)	(2)	(3)	(4)	(5)
BD	0.466***	0.348***			
	(0.090)	(0.108)			
Mig_{it}	0.050*	0.050*			
	(0.030)	(0.030)			
$RTA^{\#}_{i,t-5}$	0.006***	0.006***			
	(0.001)	(0.001)			
$ln(GDP)_{i,t-5}$	0.329***	0.329***			
	(0.030)	(0.030)			
BD*Differentiated		0.241**	0.233**		
		(0.109)	(0.109)		
BD*Skill Intensive				1.011***	
				(0.114)	
BD*High Tech					0.699***
					(0.148)
Observations	171,028	171,028	171,028	171,028	171,028
R-squared	0.711	0.711	0.722	0.723	0.722
Country FE	Yes	Yes	No	No	No
Country-Year FE	No	No	Yes	Yes	Yes
Sector-Year FE	Yes	Yes	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are double clustered at country-sector and sector-year. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 186.

Table 7: Country-sector aggregate results. Number of destinations estimations

Dep var	# of $Destinations_{ikt}$ (ln)				
	(1)	(2)	(3)	(4)	(5)
BD	0.076***	0.070**			
	(0.028)	(0.034)			
Mig_{it}	-0.015	-0.015			
	(0.010)	(0.010)			
$RTA^{\#}_{i,t-5}$	0.003***	0.003***			
	(0.000)	(0.000)			
$ln(GDP)_{i,t-5}$	0.154***	0.154***			
	(0.009)	(0.009)			
BD*Differentiated		0.012	0.006		
		(0.037)	(0.037)		
BD*Skill Intensive				0.301***	
				(0.040)	
BD*High Tech					0.123**
					(0.058)
Observations	237,864	237,864	237,864	237,864	237,864
R-squared	0.829	0.829	0.842	0.842	0.842
Country FE	Yes	Yes	No	No	No
Country-Year FE	No	No	Yes	Yes	Yes
Sector-Year FE	Yes	Yes	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are double clustered at country-sector and sector-year. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 186.

Table 8: Country-sector aggregate results. Synthetic export performance (RCA) estimations

Dep var	RCA_{ikt} (ln)				
	(1)	(2)	(3)	(4)	(5)
BD	0.539***	0.198**			
	(0.080)	(0.094)			
Mig_{it}	0.031	0.032			
	(0.025)	(0.025)			
$FTA^{\#}_{i,t-5}$	0.008***	0.008***			
	(0.001)	(0.001)			
$ln(GDP)_{i,t-5}$	0.282***	0.283***			
	(0.025)	(0.025)			
BD*Differentiated		0.700***	0.680***		
		(0.091)	(0.089)		
BD*Skill Intensive				0.665***	
				(0.093)	
BD*High Tech					0.822***
					(0.132)
Observations	186,169	186,169	196,858	196,858	196,858
R-squared	0.543	0.544	0.566	0.566	0.566
Country FE	Yes	Yes	No	No	No
Country-Year FE	No	No	Yes	Yes	Yes
Sector-Year FE	Yes	Yes	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are double clustered at country-sector and sector-year. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 186.

Table 9: Country-sector aggregate results. Results by job complexity

Dep var	$Export_{ikt}$	$Export_{ikt}$	$Export_{ikt}$	RCA_{ikt}	$\#\ Dest_{ikt}$
	(1)	(2)	(3)	(4)	(5)
BD	0.257***	0.259***			
	(0.064)	(0.067)			
Mig_{it}	-0.013	-0.013			
	(0.024)	(0.024)			
$FTA^{\#}_{i,t-5}$	0.003***	0.003***			
	(0.001)	(0.001)			
$ln(GDP)_{i,t-5}$	0.329***	0.330***			
	(0.028)	(0.028)			
BD*Job Complexity		0.413***	0.425***	0.203***	0.071***
		(0.046)	(0.045)	(0.040)	(0.020)
Observations	116,268	116,268	116,268	114,262	114,262
R-squared	0.834	0.835	0.848	0.763	0.874
Country FE	Yes	Yes	No	No	No
Country-Year FE	No	No	Yes	Yes	Yes
Sector-Year FE	Yes	Yes	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are double clustered at country-sector and sector-year. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 186.

Table 10: Placebo test: abstract vs manual intensive sectors

Dep var	$Export_{ikt}$	$Export_{ikt}$	RCA_{ikt}	# $Dest_{ikt}$	$Export_{ikt}$	$Export_{ikt}$	RCA_{ikt}	# $Dest_{ikt}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BD	0.279***				0.276***			
	(0.068)				(0.064)			
Mig_{it}	-0.011				-0.013			
	(0.024)				(0.024)			
$FTA^{\#}_{i,t-5}$	0.003***				0.003***			
	(0.001)				(0.001)			
$ln(GDP)_{i,t-5}$	0.341***				0.340***			
	(0.029)				(0.029)			
BD*Abstract Intensive	0.448***	0.456***	0.214***	0.003				
	(0.049)	(0.048)	(0.045)	(0.019)				
BD*Manual Intensity					0.005	-0.002	-0.118**	-0.016
					(0.055)	(0.055)	(0.052)	(0.022)
Observations	114,262	114,262	114,262	114,262	114,262	114,262	114,262	114,262
R-squared	0.831	0.844	0.763	0.874	0.831	0.844	0.763	0.874
Country FE	Yes	No	No	No	Yes	No	No	No
Country-Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Sector-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: In all regressions standard errors in parentheses are double clustered at country-sector and sector-year. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 186.

Bibliography

Alesina, A., Harnoss, J. & Rapoport, H. (2016), 'Birthplace diversity and economic prosperity', *Journal of Economic Growth* **21**(2), 101–138.

Bahar, D. & Rapoport, H. (2018), 'Migration, knowledge diffusion and the comparative advantage of nations', *Economic Journal* **128**(612), F273–F305.

Beine, M. & Parsons, C. R. (2017), 'Climatic factors as determinants of international migration: Redux', *CESifo Economic Studies* **63**(4), 386–402.
URL: *+ http://dx.doi.org/10.1093/cesifo/ifx017*

Berthou, A. & Fontagné, L. (2016), 'Variable Trade Costs, Composition Effects and the Intensive Margin of Trade', *The World Economy* **39**(1), 54–71.
URL: *https://ideas.repec.org/a/bla/worlde/v39y2016i1p54-71.html*

Buono, I. & Lalanne, G. (2012), 'The effect of the uruguay round on the intensive and extensive margins of trade', *Journal of International Economics* **86**(2), 269–283.

Card, D. (2001), 'Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration', *Journal of Labor Economics* **19**(1), 22–64.

Card, D. (2009), 'Immigration and Inequality', *American Economic Review* **99**(2), 1–21.

Costinot, A. (2009), 'On the origins of comparative advantage', *Journal of International Economics* **77**(2), 255–264.

Costinot, A., Donaldson, D. & Komunjer, I. (2012), 'What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas', *Review of Economic Studies* **79**(2), 581–608.

Docquier, F., Lodigiani, E., Rapoport, H. & Schiff, M. (2016), 'Emigration and democracy', *Journal of Development Economics* **120**(C), 209–223.

Docquier, F., Turati, R., Valette, J. & Vasilakis, C. (2018), 'Birthplace diversity and economic growth: Evidence from the us states in the post-world war ii period', *mimeo* .

Felbermayr, G. J. & Toubal, F. (2012), 'Revisiting the Trade-Migration Nexus: Evidence from New OECD Data', *World Development* **40**(5), 928–937.

Fitzgerald, D. & Haller, S. (2018), 'Exporters and shocks', *Journal of International Economics* **113**, 154 – 171.

Frankel, J. A. & Romer, D. H. (1999), 'Does Trade Cause Growth?', *American Economic Review* **89**(3), 379–399.

Goldsmith-Pinkham, P., Sorkin, I. & Swift, H. (2018), Bartik Instruments: What, When, Why, and How, NBER Working Papers 24408, National Bureau of Economic Research, Inc.
URL: *https://ideas.repec.org/p/nbr/nberwo/24408.html*

Gould, D. M. (1994), 'Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows', *The Review of Economics and Statistics* **76**(2), 302–316.

Gray, C. L. & Mueller, V. (2012), 'Natural disasters and population mobility in bangladesh', *Proceedings of the National Academy of Sciences* **109**(16), 6000–6005.

Head, K. & Mayer, T. (2014), Gravity equations: Workhorse, toolkit, and cookbook, Vol. 4, Handbook of International Economics, Gita Gopinath, Elhanan Helpman and Kenneth Rogoff editors, chapter 4.

Head, K. & Ries, J. (1998), 'Immigration and Trade Creation: Econometric Evidence from Canada', *Canadian Journal of Economics* **31**(1), 47–62.

Hong, L. & Page, S. E. (2001), 'Problem Solving by Heterogeneous Agents', *Journal of Economic Theory* **97**(1), 123–163.

Hoogendoorn, S. M. & van Praag, M. C. (2012), Ethnic Diversity and Team Performance: A Field Experiment, IZA Discussion Papers 6731, Institute for the Study of Labor (IZA).

Horwitz, S. & Horwitz, I. (2007), 'The effects of team diversity on team outcomes: A meta-analytic review of team demography', *Journal of Management* **33**(6), 987–1015.

Jaeger, D. A., Ruist, J. & Stuhler, J. (2018), Shift-Share Instruments and the Impact of Immigration, NBER Working Papers 24285, National Bureau of Economic Research, Inc.
URL: *https://ideas.repec.org/p/nbr/nberwo/24285.html*

Kahane, L., Longley, N. & Simmons, R. (2013), 'The Effects of Coworker Heterogeneity on Firm-Level Output: Assessing the Impacts of Cultural and Language Diversity in the National Hockey League', *The Review of Economics and Statistics* **95**(1), 302–314.

Lazear, E. P. (1999), 'Globalisation and the Market for Team-Mates', *Economic Journal* **109**(454), 15–40.

Leromain, E. & Orefice, G. (2014), 'New revealed comparative advantage index: Dataset and empirical distribution', *International Economics* (139), 48–70.

Maggi, G. & Grossman, G. M. (2000), 'Diversity and Trade', *American Economic Review* **90**(5), 1255–1275.

Ortega, F. & Peri, G. (2014), 'Openness and income: The roles of trade and migration', *Journal of International Economics* **92**(2), 231–251.

Ottaviano, G. I. & Peri, G. (2006), 'The economic value of cultural diversity: evidence from US cities', *Journal of Economic Geography* **6**(1), 9–44.

Ozden, C., Parsons, C. R., Schiff, M. & Walmsley, T. L. (2011), 'Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960-2000', *World Bank Economic Review* **25**(1), 12–56.
URL: *https://ideas.repec.org/a/oup/wbecrv/v25y2011i1p12-56.html*

Parrotta, P., Pozzoli, D. & Pytlikova, M. (2014), 'Labor diversity and firm productivity', *European Economic Review* **66**(C), 144–179.

Parrotta, P., Pozzoli, D. & Sala, D. (2016), 'Ethnic diversity and firms' export behavior', *European Economic Review* **89**(C), 248–263.

Parsons, C. & Vezina, P.-L. (2018), 'Migrant networks and trade: The vietnamese boat people as a natural experiment', *Economic Journal* **128**(612), F210–F234.

Peri, G. & Requena-Silvente, F. (2010), 'The trade creation effect of immigrants: evidence from the remarkable case of Spain', *Canadian Journal of Economics* **43**(4), 1433–1459.

Peri, G. & Sparber, C. (2009), 'Task Specialization, Immigration, and Wages', *American Economic Journal: Applied Economics* **1**(3), 135–169.

Rauch, J. E. (2001), 'Business and Social Networks in International Trade', *Journal of Economic Literature* **39**(4), 1177–1203.

Rodríguez-Pose, A. & von Berlepsch, V. (2017), Does population diversity matter for economic development in the very long-term? Historic migration, diversity and county wealth in the US, CEPR Discussion Papers 12347, C.E.P.R. Discussion Papers.
URL: *https://ideas.repec.org/p/cpr/ceprdp/12347.html*

Silva, J. M. C. S. & Tenreyro, S. (2006), 'The Log of Gravity', *The Review of Economics and Statistics* **88**(4), 641–658.

Trax, M., Brunow, S. & Suedekum, J. (2015), 'Cultural diversity and plant-level productivity', *Regional Science and Urban Economics* **53**(C), 85–96.

United Nations, P. D. (2015), Trends in international migrant stock: Migrants by destination and origin. the 2015 revision, Technical report, United Nations Department of Economic and Social Affairs Population Division, New York.

Vivian, H., Dennis, L. & Sara, P. (2015), Diversity matters, Technical report, McKinsey&Company.

White, R. (2007), 'Immigrant-trade links, transplanted home bias and network effects', *Applied Economics* **39**(7), 839–852.

Yotov, Y. V., Piermartini, R., Monteiro, J.-A. & Larch, M. (2017), *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, United Nations publication, World Trade Organization.
URL: *https://books.google.fr/books?id=cWX5MAAACAAJ*

Appendix. Additional Tables and Figures

Appendix A. Natural Disasters

In the following section we detail the countries affected by a severe natural disaster that we use as source of identification in the natural experiment section. We define a disaster as severe if it causes both economic and social disruptions. Using the total damages, number of people affected and total casualties as proxy for the economic and social impact of a catastrophic event, A1 identify the 41 countries affected by a severe event during the pre-sample period, 1985-1990. The data are from the EM-DAT (Emergency Events Database) and consider only natural events: biological (epidemic), climatological (drought, wildfire), geophysical (mass movement, earthquake, volcanic activity), hydrological (flood, landslide), meteorological (storm, fog, extreme temperature).

Table A1: List of countries affected by (severe) natural disasters, 1985-1990

Countries	Events	Damages (US$, Mln)	Affected	Casualties
ARG	6	1640	1416990	44
ATG	1	80	8030	2
AUS	6	265.839	1012	15
BEN	1	4.8	475000	61
BGD	14	2187	57905460	19561
BOL	1	50	310000	29
BRA	13	1886	3752961	884
CAN	2	117	1000	12
CHL	5	1678	1684781	344
CHN	92	13868.94	280067742	11680
COM	1	9	50000	24
CRI	4	88.5	154609	36
DZA	2	1	15000	56
ECU	4	1500	166006	5102
FSM	1	6	203	5
GLP	1	50	11084	5
HKG	4	0.067	3512	12
HND	1	100	48000	5
HTI	5	91.286	873901	81
IDN	22	76.641	285250	832
IND	39	4498.843	21765519	7590
IRN	7	8311.7	884117	40142
ITA	5	2105	2716	27
JAM	1	5.2	300	7
JPN	3	5713	148366	67
KOR	3	547	210000	669
MEX	6	4430.6	2255204	9811
MSR	1	240	12040	11
MWI	3	28	150544	57
NIC	1	400	360278	130
PAN	2	60.35	14732	32
PER	11	60.2	2515946	412
PHL	53	1766.393	22974707	6554
SLV	1	1500	770000	1100
THA	1	452	199000	458
TON	1	2.5	3103	1
TZA	3	0.28	162868	389
USA	34	18574.1	1055222	634
VEN	4	1.8	18029	139
VNM	8	21.725	6929667	1343
YEM	1	33	340000	25

Note: The table reports the total number of (severe) natural disasters over the period 1985-1990 by country, along with the amount of damages, in milions US$, the number of affected residents and total number of casualties. Source: EM-DAT: The Emergency Events Database - Université catholique de Louvain (UCL) - CRED, D. Guha-Sapir - www.emdat.be, Brussels, Belgium.

Appendix B. Exporter country region-income level fixed effects

Table A2: List of countries by region-income cell

Region	Income Level High	Upper-Middle	Lower-Middle	Low	Nes	Total
East Asia & Pacific	7	4	15	7	0	33
Europe & Central Asia	18	12	19	0	0	49
Latin America & Caribbean	4	10	20	2	0	36
Middle East & North Africa	6	6	9	0	0	21
North America	3	0	0	0	0	3
South Asia	0	0	0	8	0	8
Sub-Saharan Africa	0	2	9	31	0	42
Nes	0	0	0	0	3	3
Total	38	34	72	48	3	195

Note: The table reports the total number of countries per region-income cell. Both regions and income levels are from the World Bank. Income levels refers to the first available year reported in the World Bank database: 1987 (151 countries); between 1987-1994 (37 countries); PLW (1996); SRB (2006); TCA and TUV (2009). The category "Nes" includes 3 countries are not included neither in the region nor in the income database: GIB, NRU, VGB.

Appendix C. Data base on synthetic revealed comparative advantage measures.

A side product of the present paper is a new database on the synthetic revealed comparative advantage obtained by estimating the equation 9. The theoretical foundation for considering estimated country-year fixed effects a good proxy for (Ricardian) revealed comparative advantage is provided by Costinot et al. (2012). This database is intended as updating and extending the CEPII RCA comparative advantage index proposed by Leromain & Orefice (2014).

In the dedicated web page the user may download data for the full sample of 195 countries over the period 1990-2015 at both HS 2-digit and 4-digit product level. The user will find two databases, one for each HS product aggregation. Each database contains four variables: (i) the ISO code of the country (variable i); (ii) the year (variable $year$); (iii) the sector of interest (variable hs_code); and (iv) the measure of synthetic revealed comparative advantage (variable RCA).

Appendix D. Robustness of Diversity.

Table A3: Intensive margin. OLS regressions.

Dep var	$\text{Ln(export)}\vert_{export(t-1)>0}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mig_{ijt}	0.134***	0.139***	0.096***	0.061***	0.252***	0.256***	0.257***
	(0.008)	(0.008)	(0.004)	(0.004)	(0.017)	(0.017)	(0.017)
$\text{Mig}_{ijt} * TS_{ijk}$	0.058***	0.059***	0.039***	0.040***	0.164***	0.125***	0.109***
	(0.004)	(0.004)	(0.003)	(0.002)	(0.010)	(0.010)	(0.013)
TS_{ijk}	-1.982***	-1.983***	-0.926***	-0.972***	-0.741***	-0.706***	-0.692***
	(0.040)	(0.038)	(0.026)	(0.026)	(0.021)	(0.021)	(0.020)
BD_{ijt}	0.620**	0.587**	0.337***	0.322***	0.057***	0.056***	0.056***
	(0.263)	(0.268)	(0.116)	(0.111)	(0.020)	(0.020)	(0.020)
$\ln(\text{dist})_{ij}$	-0.615***	-0.614***	-1.135***	-1.061***	-1.061***	-1.066***	-1.071***
	(0.035)	(0.035)	(0.029)	(0.028)	(0.028)	(0.028)	(0.028)
Time Diff_{ij}	0.042***	0.044***	0.023***	0.025***	0.025***	0.024***	0.024***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Colony_{ij}	0.312***	0.308***	0.399***	0.207***	0.207***	0.201***	0.198***
	(0.091)	(0.090)	(0.065)	(0.058)	(0.058)	(0.058)	(0.058)
Language_{ij}	0.021	0.015	0.329***	0.263***	0.263***	0.265***	0.267***
	(0.056)	(0.055)	(0.035)	(0.031)	(0.031)	(0.031)	(0.031)
Border_{ij}	0.590***	0.587***	0.331***	0.214***	0.214***	0.213***	0.215***
	(0.050)	(0.050)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
RTA_{ijt}	0.285***	0.267***	0.401***	0.362***	0.362***	0.363***	0.364***
	(0.040)	(0.043)	(0.031)	(0.030)	(0.030)	(0.030)	(0.030)
$\ln(1 + \text{Tariff})_{ijt}$	-1.745***	-1.837***	-2.274***	-2.218***	-2.218***	-2.212***	-2.218***
	(0.155)	(0.160)	(0.120)	(0.119)	(0.119)	(0.118)	(0.119)
$\ln(\text{Remot})_{it}$		4.747***	1.530***	1.460***	1.460***	1.423***	1.396***
		(0.829)	(0.359)	(0.346)	(0.346)	(0.343)	(0.341)
$\ln(\text{Trade})_{ikt,-jr}^{IIquart}$			1.595***	1.587***	1.587***	1.596***	1.603***
			(0.027)	(0.027)	(0.027)	(0.026)	(0.026)
$\ln(\text{Trade})_{ikt,-jr}^{IIIquart}$			2.762***	2.732***	2.732***	2.750***	2.765***
			(0.039)	(0.038)	(0.038)	(0.037)	(0.037)
$\ln(\text{Trade})_{ikt,-jr}^{IVquart}$			4.342***	4.267***	4.267***	4.297***	4.324***
			(0.055)	(0.051)	(0.051)	(0.051)	(0.050)
$\ln(\text{Emig})_{ijt}$				0.094***	0.094***	0.093***	0.092***
				(0.004)	(0.004)	(0.004)	(0.004)
Definition TS:	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 0.75$	$BI_{ik} < 0.5$
	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1.25$	$BI_{jk} > 1.5$
Std Var	No	No	No	No	Yes	Yes	Yes
jkt FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
rlt FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,706,473	1,706,473	1,706,473	1,706,473	1,706,473	1,706,473	1,706,473
R-squared	0.404	0.408	0.536	0.542	0.542	0.541	0.540

Note: In all regressions standard errors in parentheses are clustered at exporter country-year level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 195 each one observed for 5 time periods (1995, 2000, 2005, 2010, 2015).

Table A4: Extensive margin. OLS regressions.

Dep var	$Dummy = 1$ if $Export_{ijkt} > 0$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mig$_{ijt}$	0.017***	0.017***	0.012***	0.008***	0.035***	0.035***	0.035***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Mig$_{ijt} * TS_{ijk}$	0.006***	0.006***	0.004***	0.004***	0.015***	0.014***	0.014***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
TS$_{ijk}$	-0.127***	-0.127***	-0.037***	-0.040***	-0.023***	-0.022***	-0.025***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
BD$_{ijt}$	0.096***	0.085***	0.036**	0.035**	0.007**	0.007**	0.007**
	(0.026)	(0.026)	(0.014)	(0.015)	(0.003)	(0.003)	(0.003)
ln(dist)$_{ij}$	-0.041***	-0.039***	-0.085***	-0.077***	-0.077***	-0.077***	-0.077***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Time Diff$_{ij}$	0.005***	0.005***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Colony$_{ij}$	0.023***	0.022***	0.040***	0.023***	0.023***	0.023***	0.022***
	(0.008)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Language$_{ij}$	0.007	0.009*	0.035***	0.028***	0.028***	0.028***	0.028***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Border$_{ij}$	-0.002	-0.002	-0.014***	-0.028***	-0.028***	-0.028***	-0.028***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
RTA$_{ijt}$	0.033***	0.032***	0.039***	0.035***	0.035***	0.035***	0.035***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ln(1 + Tariff)$_{ijt}$	-0.185***	-0.194***	-0.214***	-0.210***	-0.210***	-0.210***	-0.210***
	(0.014)	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
ln(Remot)$_{it}$		0.456***	0.163***	0.152***	0.152***	0.152***	0.151***
		(0.070)	(0.039)	(0.040)	(0.040)	(0.040)	(0.040)
ln(Trade)$^{IIquart}_{ikt,-jr}$			0.170***	0.167***	0.167***	0.167***	0.167***
			(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
ln(Trade)$^{IIIquart}_{ikt,-jr}$			0.265***	0.260***	0.260***	0.261***	0.261***
			(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
ln(Trade)$^{IVquart}_{ikt,-jr}$			0.359***	0.349***	0.349***	0.350***	0.350***
			(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ln(Emig)$_{ijt}$				0.009***	0.009***	0.009***	0.009***
				(0.000)	(0.000)	(0.000)	(0.000)
Definition TS:	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 1$	$BI_{ik} < 0.75$	$BI_{ik} < 0.5$
	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1$	$BI_{jk} > 1.25$	$BI_{jk} > 1.5$
Std Var	No	No	No	No	Yes	Yes	Yes
jkt FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
rlt FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,000,682	3,000,682	3,000,682	3,000,682	3,000,682	3,000,682	3,000,682
R-squared	0.173	0.175	0.221	0.223	0.223	0.223	0.223

Note: In all regressions standard errors in parentheses are clustered at exporter country-year level. $***$, $**$, $*$ denotes statistically significance at the 1%, 5% and 10% level, respectively. Total number of countries is 195 each one observed for 5 time periods (1995, 2000, 2005, 2010, 2015).