DISCUSSION PAPER SERIES

Immigrant Entrepreneurs, Diasporas and Exports

Massimiliano Bratti
Luca De Benedictis
Gianluca Santoni

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

Immigrant Entrepreneurs, Diasporas and Exports

Massimiliano Bratti
European Commission JRC, DEMM, IZA and LdA

Luca De Benedictis
Rossi-Doria Center, University Roma Tre, DED, and Luiss

Gianluca Santoni
Cepii

ABSTRACT

Immigrant Entrepreneurs, Diasporas and Exports

In this paper we highlight a new complementary channel to the business and *social network effect à la* Rauch (2001) through which immigrants generate increased export flows from the regions in which they settle to their countries of origin: they can become *entrepreneurs*. Using very small-scale (NUTS-3) administrative data on immigrants' location in Italy, the local presence of immigrant entrepreneurs (i.e. firms owned by foreign-born entrepreneurs) in the manufacturing sector, and on trade flows in manufacturing between Italian provinces and more than 200 foreign countries, we assess the causal relationship going from diasporas and immigrant entrepreneurs towards export flows. Both the size of the diaspora and the number of immigrant entrepreneurs have a positive, significant and economically meaningful effect on exports. We find that increasing the stock of (non-entrepreneur) immigrants by 10% would lead to a 1.7% increase in exports in manufacturing towards immigrants' countries of origin, while increasing the number of immigrant entrepreneurs in manufacturing by 10% would raise exports by about 0.6%. We also show that, besides these bilateral effects, the population of immigrant entrepreneurs raises a province's overall competitiveness and export flows towards all potential destinations.

Corresponding author:
Gianluca Santoni
Cepii
20 avenue de Ségur
TSA 10726
75334 Paris cedex 07
France
E-mail: gianluca.santoni@cepii.fr

1   Introduction

The prevailing explanation for the positive effect of immigrants on exports is that proposed by Rauch (1999, 2001) and Rauch and Trindade (2002), and it is consistent with the significant empirical evidence from several national cases (Combes *et al.*, 2005, Peri and Requena-Silvente, 2010, Bratti *et al.*, 2014): this is the so-called *business and social network effect.* Briefly, when immigrants move from their home country to a new country of residence, they affect not only the latter's labour supply and local demand for goods and services, but also bring to the host country knowledge of their home country's economy and institutions and kinship links that endure in spite of distance and time. This knowledge and these links can be exploited by the host (home) country's entrepreneurs, who can use them to reduce the sunk cost of exporting to and importing from the migrants' country of origin (e.g. information costs[1]) and/or to plane away the marginal cost of exporting.[2] Since the knowledge and the international links embodied in immigrants are largely country specific, the reduction in trade costs acts mainly at the bilateral level.

Although most literature has focused on co-ethnic population networks (namely the effect of immigrants), the individuals who are the natural candidates to explain *business and social network effects* are those who have business-related knowledge and business contacts abroad and who are able to profit from that, having settled down in the host country as migrants: the group of immigrants that establishes firms in the host country and that we call for simplicity *immigrant entrepreneurs.*[3] On top of promoting trade by providing market information, by supplying matching and referral services, and by ensuring contract enforcement through social sanctions when market institutions are weak — like non-entrepreneur migrants can do — immigrant entrepreneurs are also able to directly exploit the advantages deriving from knowledge of their own country, which is better than that of domestic en-

---

[1] Cross-border networks of people sharing the same country of origin can substitute or integrate organized markets in matching international demand and supply. See Rauch (1999) and Felbermayr *et al.* (2015) for a summary of the literature.

[2] As stated by Briant *et al.* (2014): "...immigrant networks may provide contract enforcement through sanctions and exclusions, which substitutes for weak institutional rules and reduces trade costs."

[3] In what follows, we use the term 'immigrant entrepreneurs' to refer to firms owned by immigrants, and not necessarily to those producing and selling 'ethnic goods,' i.e. goods with specific cultural or national connotations.

trepreneurs, to directly sell their goods abroad.

To the best of our knowledge, this is the first paper that, controlling for standard gravity equation covariates and the role played by diasporas, rigorously quantifies the causal effect of *immigrant entrepreneurs* on international trade.[4] This is a timely moment to increase the stock of knowledge on this issue. As stated by Naud *et al.* (2017), in spite of "entrepreneurship and migration [being] at the very top of many national and international agendas," little is still known on the interplay between these two phenomena.

Of course, the increasing relevance of immigrant entrepreneurship goes hand in hand with the increase in migration flows,[5] and the rate of self-employment of immigrants tends to be higher than that of natives or, at any rate, substantial[6] and growing.[7] But precise quantification of the causal effect of immigrant entrepreneurs on exports is uncommon, and we contribute to build up an initial stock of knowledge on this important issue.

Using administrative data on immigrants' location in Italian provinces,[8] i.e. *diasporas* (see, Beine *et al.*, 2011), on the number of *immigrant entrepreneurs* by nationality in the manufacturing sector, and on export flows in the same sector between Italian provinces and more than 200 foreign countries, we assess, though an instrumental variables (IV) strategy, the causal

---

[4] Unlike Aleksynska and Peri (2014), we do not use data on immigrants employed in business-related occupations (e.g managers, sales persons), who may also work in non-tradeable sectors, but we use administrative data on the stock of manufacturing firms owned by immigrants. Our paper also adds to the cross-sectional evidence in Aleksynska and Peri (2014) by providing panel-data evidence for very small geographical units.

[5] Olney (2013) looking at establishments in US cities, shows the positive correlation between immigrants and the number of establishments, especially those of a smaller size.

[6] Internationally, as reported by the OECD (2010), "immigrants in many OECD countries exhibit higher rates of self-employment than natives," notably in Canada, France, the UK, Australia, the Nordic countries and the eastern European countries. Even where immigrants exhibit lower rates of self-employment than natives, as in Turkey, Spain, Portugal, Greece, Switzerland and Italy, those rates are above 10%. In Italy the rate is 17.5%.

[7] In dynamic terms, Fairlie and Lofstrom (2014) observe that "...trends in self-employment rates and new business formation are increasing among immigrants but decreasing among natives." From this point of view Italy is no exception. The overall number of foreign-owned individual (i.e. single owner) firms increased substantially over the past decade, at an annualized rate of 4.4%, countervailing the decrease in domestically owned individual enterprises, -4.5% per year, with annual changes computed as compound annual growth rates.

[8] In the Classification of Territorial Units for Statistics (NUTS; *Nomenclature des units territoriales statistiques*), provinces correspond to NUTS-3 level regions.

relationship going from diasporas and immigrant entrepreneurs towards export flows. The endogeneity of the diaspora and immigrant entrepreneurs is addressed by using both the popular instrument based on immigrant *enclaves* (Altonji and Card, 1991) and instruments built using auxiliary gravity models. Both variables, the diaspora and that of immigrant entrepreneurs, have a positive, significant and economically meaningful effect on exports. In particular, we find that increasing the number of immigrants in a province by 10% (i.e. about 26 immigrants at the sample mean) would lead to a 1.7% increase in exports towards their country of origin, while increasing immigrant entrepreneurship by 10% (i.e. a bit less than one entrepreneur, precisely 0.84 at the sample mean) would raise exports by about 0.6%.

Besides these bilateral trade-creating effects, our analysis demonstrates that the number of immigrant entrepreneurs impacts positively on province-level exports by increasing a province's overall competitiveness, which raises its exports towards all potential destinations. We find that increasing the population of immigrant entrepreneurs by 10% increases competitiveness (i.e. the province-year fixed effect of the gravity model) by 2.4%.

Our paper contributes to the literature on the role of social and business network effects on international trade by capturing the effect of immigrant entrepreneurs over and above that of diasporas. Moreover, it is also related to the recent literature on the impact of immigrants on economic growth and development (see, for instance, Ortega and Peri, 2014), highlighting the effect of immigrant entrepreneurs on exports.

The structure of the paper is as follows. The next section sets the stage for the analysis, reviewing some related literature. Section 3 describes the theoretical setup, which links diasporas and immigrant entrepreneurs to exports. Section 4 describes the main features of the data we use. Section 5 describes the empirical strategy that we follow to investigate the effect of diasporas and immigrant entrepreneurs on trade, and Section 6 comments on the main findings. Finally, Section 7 concludes, summarizing the novelties of the paper and pointing out possible directions for future research.

2 Related literature

As pointed out in the introduction, a wealth of studies on the association between immigrants and trade already exists (see Parsons and Winters, 2014, Felbermayr *et al.*, 2015, for a review of the evidence). Some recent papers

try to rigorously assess the causal effect of immigrants on trade by using IV estimation based either on the shift-share instrument popularized by Altonji and Card (1991) and motivated by the concept of immigrant enclaves,[9] or by exploiting quasi-experiments provided by refugee dispersion policies (e.g. Parsons and Vézina, 2014, Steingress, 2017).

In spite of this abundance of evidence, the channel through which immigrants affect trade remains largely a black box. Drawing on the idea pioneered by Gould (1994) and further developed by Rauch (2001), namely the *business and social network effect*, most authors focus on the bilateral effect of the foreign-born population of a given nationality located in a region (i.e. *diaspora*) on that region's trade with the immigrants' country of origin, and they interpret the positive effect as originating from the knowledge capital (concerning the country of origin's market) and the social network of immigrants. However, none of the extant studies is able to provide smoking-gun evidence that this is really the explanation, or to explore whether and how immigrants' knowledge is transferred to natives, in particular native entrepreneurs, for whom this knowledge flow should be relevant primarily to spurring trade.

Recent studies have sought to add more pieces to this puzzle. The trade-creating effect of immigrants may depend on their characteristics. One study reports that immigrants in skilled occupations have a larger effect on trade (Herander and Saavedra, 2005). Some immigrants could play a pivotal role in establishing business connections, e.g. those employed in managerial and sales jobs. Consistent with this idea, Aleksynska and Peri (2014) show that 'business-related immigrants' have an effect that is double that of non-business-related immigrants. Moreover, after classifying immigrants in managerial and sales jobs by educational level, a statistically significant positive effect of highly educated immigrants on both imports and exports is found. Also in line with the social network effect is evidence that independent immigrants, who presumably retain stronger family ties with their home countries, have a larger impact on trade than family immigrants (Head and Ries, 1998) and the finding that older immigrants, who generally have more connections, have a stronger effect on trade (Herander and Saavedra, 2005).

Another strand of literature, based on firm-level evidence, highlights the trade-enhancing role of immigrants inside the firm. Hatzigeorgiou and Lode-

---

[9] For instance, Peri and Requena-Silvente (2010), Bratti *et al.* (2014).

falk (2016), in an employer-employee panel for Sweden, show that small firms in particular can gain from hiring foreign-born workers who are skilled and recently arrived immigrants and who have a clear positive impact on exports. Similar evidence is found for Denmark (Hiller, 2013) and France (Marchal and Nedoncelle, 2016, Mitaritonna *et al.*, 2017), especially for skilled workers.

A separate and growing literature on immigrant entrepreneurs focuses on the differences between foreign-born and native entrepreneurs and among foreign-born entrepreneurs.[10] Sahin *et al.* (2014) reports that skilled immigrants in the US are "more likely to start firms with more than 10 employees than comparable natives;" in the UK, the probability of starting a firm is higher for those who initially arrived on a study visa or a work visa, compared with those who arrived via family reunification (Clark and Drinkwater, 2000). Immigrant entrepreneurs appear to specialize in a narrower range of industries or occupations than native entrepreneurs (Patel and Vella, 2013).

However, surprisingly enough, none of the existing studies have investigated the hypothesis that a great deal of the trade-creating effect of immigrants may be accounted for, over and above diasporas, by immigrant entrepreneurs.

3 Conceptual framework

The idea, however, is very simple and intuitive. If diasporas bring knowledge capital and social networks to the host country, reducing both the sunk and the marginal cost of exporting, immigrant entrepreneurs are in the position of directly exploiting this knowledge capital and these links for their own businesses. Let us frame this idea in a general heterogeneous firms' set-up of international trade (Melitz, 2003, Arkolakis *et al.*, 2012), as summarized in Hsieh *et al.* (2016).

We focus on country ι. The country is composed of $\mathcal{I}$ provinces (with $i \in [1, \mathcal{I}]$). Firms located in each province are (potential) exporters to each j foreign country (with $j \in [1, \mathcal{J}]$), while immigrants from (potentially) each j foreign country can settle in (potentially) every province i of country $\mathcal{I}$. At the aggregate level, and for every unit of time considered, we can express total exports from province i to country j as $X_{ijt} \geq 0$; the total number of firms located in i and exporting to j as $n_{ijt} \geq 0$; and the total number of

---

[10] See Fairlie and Lofstrom (2014) for a comprehensive review of the literature and Kerr and Kerr (2017) for recent evidence from the US.

immigrants from country j living in province i as $D_{jit} \geq 0$. Therefore, both D_{jit} and X_{ijt} can be zero if there is no diaspora from country j to province i and if there are no firms located in i that export to country j.

The *trade equation* relates X_{ijt} to n_{ijt}, to the total consumer expenditure in country j, Y_{jt}, and to the relative price of the average productivity i-firm selling variety ω to consumers in country j:

$$X_{ijt} = n_{ijt} \left(\frac{\bar{p}_{ijt}}{P_{jt}} \right)^{1-\sigma} Y_{jt}, \tag{1}$$

where P_{jt} is the price index dual to the CES utility function of consumers in country j, over the Ω_{ijt} varieties produced in i and sold in j, and $\sigma > 1$ is the constant elasticity of substitution between varieties, so that the income elasticity of demand is also constant.

The price $\bar{p}_{ijt}$ depends on production costs (w_{it}), trade costs (τ_{ijt}) and exporting firms' average productivity ($\bar{\phi}_{ijt}$). A constant mark up is assumed, for instance,

$$\bar{p}_{ijt} = \frac{\sigma}{\sigma - 1} \left(\frac{w_{it}\tau_{ijt}}{\bar{\phi}_{ijt}} \right) = \frac{\sigma}{\sigma - 1} \left(\frac{w_{it}\tau_{ijt}}{k_{ijt}\bar{\phi}_{it}} \right), \tag{2}$$

where $\bar{\phi}_{ijt}$ is an average computed over the set of all productivities of firms in province i serving country j (Φ_{ijt}). We define a variable $k_{ijt} = \bar{\phi}_{ijt}/\bar{\phi}_{it}$, i.e. the ratio between the average productivity of firms in i serving market j and the average productivity computed over all exporters in province i.

Finally, to close the model we equate expenditure with labour income in country j

$$Y_{jt} = w_{jt}L_{jt}. \tag{3}$$

The system of equations (1)-(3) constitutes the conceptual framework that we use to discuss the possible effects of immigrants on exports. Let us focus on each equation in turn.

Equation (3) highlights the first mechanical effect of migration on trade: the contraction in L_{jt} due to emigration reduces foreign demand. Essentially, the migration balance in country j will have a positive effect on X_{ijt}. On the other hand, the interplay between labour demand and labour supply may lead to positive effects of emigration on labour incomes in the home country (Dustmann *et al.*, 2015). Y_{jt} can also be influenced by demographic factors other than migration, by labour market conditions in country j and by

many social and political factors that go beyond the scope of our conceptual framework. We cannot account for the effect of these factors on X_{ijt}, but this calls for the need to take these unspecified factors under control. We will undertake this by exploiting the panel dimension of the data and using *country j-year fixed effects* in the empirical analysis.

Equation (2) indicates the role of prices ($\bar{p}_{ijt}$) in determining export flows (X_{ijt}). This will depend in turn on wages (w_{it}), trade costs (τ_{ijt}) and average exporting firms' productivity ($\bar{\phi}_{ijt}$).

As far as w_{it} is concerned, the literature on the wage effect of immigration is large (Borjas, 1994, 2015, Card, 2001, 2009) and provides different predictions (Ottaviano and Peri, 2012, Bratsberg *et al.*, 2014) depending on labour market and worker characteristics, on whether markets are fully integrated or dual/segmented, and on the degree of substitutability or complementarity between domestic workers and immigrant workers. In general, we expect either a negative or a zero effect on average wages (including both natives and immigrants) at the province level, with an effect on exports that may be either positive or null (Cortes, 2008, Balkan and Tumen, 2016). We expect such effect, if any, to act on X_{ijt} but also on X_{it}, boosting provincial exports towards *all* destinations. The limited information we have on local labour markets at the province level again raises the need to control for the wage effect through the use of *province i-year fixed effects* in the empirical analysis.

As regards $\bar{\phi}_{ijt}$, Ottaviano and Peri (2006) and Sparber (2008) show how the productivity of firms may be affected by the presence of immigrants. The channels can be many. If complementarity between workers of different ethnic groups exists, an increase in *ethnic diversity* may have a positive effect on average firms' productivity.[11] Immigrant workers may also be more productive if they are likely to be positively selected in terms of ability, especially if they become entrepreneurs. Immigrants might also be more motivated (Sahin *et al.*, 2014) and work longer or non-standard hours than natives (Zhang and Sanders, 1999, Giuntella, 2012), and increase production per worker. High-skilled immigrants also have a positive influence on innovation (Lachenmaier and Woessmann, 2006, Hunt and Gauthier-Loiselle, 2010, Parrotta *et al.*, 2014, Jahn and Steinhardt, 2016), which in turn has a positive impact on trade (Lachenmaier and Woessmann, 2006, Becker and Egger, 2013). It is, however, difficult to conceive that the positive effect of

---

[11] Although Alesina and La Ferrara (2005) document that the effect can be the opposite if the integration of different ethnic groups implies extra communication costs for firms.

employing immigrants from country j would operate for the specific trading pair i-j, and we expect diasporas to increase firms' productivity and create a positive effect on exports to *all* destinations. In such a case, the effect of $\bar{\phi}_{it}$ would be well captured by the *province i-year fixed effects* mentioned previously. The effect could be different for each destination country if, in the spirit of Melitz (2003), the increase in productivity affects the destination-specific productivity thresholds. In that case, the increase in productivity would be enough to export to certain countries but not to others. This will emerge both through prices, affecting the intensive margin of firm trade, and also through the effect on n_{ijt}, i.e. the number of exporting firms, which we will discuss shortly.

The effect of immigrants on trade costs is at the heart of the literature on migration and trade reviewed in Section 2. Trade costs can be fixed and sunk, $\mathcal{T}_{ijt}$, or proportional to the value of the good exported, as in the case of τ_{ijt}. Diasporas can reduce the marginal costs of exporting from i to j through the establishment of an enforcement channel, which operates as an insurance mechanism (Rauch, 2001, Briant *et al.*, 2014). Using the iceberg cost metaphor, since less of the shipped good melts away during the journey between i and j, the reduction in τ_{ijt} due to diasporas operates at the intensive margin of exports.

$\mathcal{T}_{ijt}$ would also be affected by diasporas if the sunk cost of exporting is related to the knowledge of the market of country j that is embedded in immigrants from country j located in i. This channel of export promotion operates at the extensive margin, lowering the productivity threshold of exporting to country j and allowing more firms to become active in such a market. This would result in an increase in n_{ijt} in equation (1).

The level of n_{ijt} in equation (1) can also influenced by the number of immigrants that become entrepreneurs. The number of immigrant entrepreneurs active in province i makes n_{ij} grow if immigrant entrepreneurs directly use their country-specific knowledge or their personal links to export. This would result in an increase in X_{ijt}.

The population of immigrant entrepreneurs and diasporas are, however, positively correlated,[12] and it becomes crucial to isolate the effects of the

---

[12] Immigrant entrepreneurs may be more likely to hire foreign workers at every occupational level. The presence of immigrant entrepreneurs could, for instance, be the reason why in a region or in a firm one can find more (highly educated) immigrants working as sales persons or managers, especially when (skilled) foreign-born workers are discriminated against by native employers.

former and the latter to gain a better understanding of what really drives the immigrants-trade link.[13]

The above conceptual framework is used to derive the following export equation, obtained by replacing in equation (1) the price equation (2) and taking logarithms,

$$\ln X_{ijt} = (1 - \sigma) \ln(\frac{\sigma}{\sigma - 1}) + \ln n_{ijt} + (1 - \sigma) \ln \tau_{ijt} - (1 - \sigma) \ln k_{ijt} +$$
$$+ \underbrace{(1 - \sigma) \ln w_{it} - (1 - \sigma) \bar{\phi}_{it}}_{\text{province-year fixed effects}} + \underbrace{(\sigma - 1) \ln P_{jt} + \ln Y_{jt}}_{\text{country-year fixed effects}}. \tag{4}$$

We do not observe n_{ijt}, τ_{ijt} and k_{ijt}, but for the reasons outlined above they can be affected by the diaspora and the stock of immigrant entrepreneurs. Thus we estimate the following reduced form gravity equation,[14] which is the stochastic version of equation (4):

$$\ln X_{ijt} = \boldsymbol{\delta}_{it} + \boldsymbol{\theta}_{jt} + \alpha_1 \ln(1 + D_{ijt}) + \alpha_2 \ln(1 + IE_{ijt}) +$$
$$+ \alpha_3 \ln d_{ij} + \alpha_4 Border_{ij} + \epsilon_{ijt} \tag{5}$$

where, to recap, i is the subscript for Italian provinces (NUTS-3), j indicates the foreign country (i.e. the country of origin of immigrants), and t stands for time. X_{ijt} is trade (exports) between province i and country j at time t (excluding zero-trade observations). D_{ijt} and IE_{ijt} are the stocks of (non-entrepreneur) immigrants and immigrant entrepreneurs, respectively, from country j located in province i, potentially acting as a trade-enhancing force, in contrast to $\ln(d_{ij})$, which is the logarithm of the great-circle distance between province i and country j. $Border_{i,j}$ is a border dummy, which is included to take into account possible non-linearities in the effect of distance.

---

[13] This is also relevant for policymaking. If a genuine positive effect on exports is found for the population of immigrant entrepreneurs, policies aiming to simplify the administrative burden (e.g. creating offices to help immigrants, who often do not speak the host country's language, with administrative requirements) and to reduce the costs required to set up firms may have positive returns in terms of propelling international trade.

[14] The Anderson and van Wincoop (2003) specification of the gravity equation can be derived from micro-foundations, and it results from an expenditure function that takes into account the fundamental role of general equilibrium effects in trade, i.e. the multilateral resistance index. See De Benedictis and Taglioni (2011), Anderson (2011) and Head and Mayer (2015) on the theoretical foundation of the gravity equation, and Beine *et al.* (2016) for an application to migration.

The province-year fixed effects $\boldsymbol{\delta}_{it}$s capture the effect of the average level of wages and firm productivity (i.e. w_{it} and $\bar{\phi}_{it}$, respectively) and can be considered as indexes of the competitiveness of province i in year t. They also absorb the effect of factors varying along the same dimensions (e.g., the number of native firms, province i's GDP). The country of origin-year fixed effects $\boldsymbol{\theta}_{jt}$s absorb the effect of variables such as country j's GDP, or participation in trade agreements. Finally, ϵ_{ijt} is an error term clustered at the province-country level to account for serial correlation in trade.

In the following section, we describe the main features of the data used in the empirical analysis, including the definition of immigrants and immigrant entrepreneurs, while identification and estimation issues are discussed in Section 5.

4   Data and descriptive statistics

The empirical analysis in this paper is carried out by combining three publicly available datasets on province-level export flows in manufacturing, foreign-born residents without Italian citizenship[15] and foreign-born entrepreneurs in manufacturing for the period 2002-2011, respectively.[16] Export flows report the value, originally recorded in euros, of custom transactions between Italian provinces and around 210 destination countries, while data on foreign-born residents without Italian citizenship cover 187 nationalities.

Concerning foreign-born entrepreneurs, we use data produced by the National Chamber of Commerce (*Infocamere*). The responsibilities of the Chamber of Commerce are defined by Law n. 580 of 1993.[17] The most relevant duty of *Infocamere* for our analysis is that provincial offices are in charge of producing and maintaining the registry of all firms active in Italy. We focus on individually owned firms (*impresa individuale*, individual firms or individual enterprises hereafter), a form of business whose entire legal and financial representation is vested in a single individual.[18] For individual firms, we can associate firm ownership with a unique person and national-

---

[15] Data are collected by the Italian National Institute of Statistics (ISTAT).

[16] The period 2002-2011 is one in which the spatial classification of administrative units remained invariant in Italian national statistics.

[17] Further reformed on February 2010 by Law Decree DL23.

[18] Individual entrepreneurs are different from self-employed workers, who do not have employees.

ity, and such data can be used to analyze bilateral trade flows using gravity models.

Individual firms are the most common legal form of firms in Italy. According to *Infocamere*, at the end of 2013 around 54% of all active firms in Italy were individual firms. The overall number of foreign-owned firms has increased substantially over the past decade — at an annualized rate of 4.4% — and accounted for 10.9% of overall individual firms in 2011 (see Figure 1). Taking a closer look at the evolution of the time series, it emerges that the greatest contribution to the sharp rise in foreign entrepreneurs comes from eastern European countries and from countries outside the EU. The number of individual firms whose owners were EU residents in 2002 declined at an annual rate of -2.45%, whereas for extra-EU countries it increased at an annualized rate of 6.04% over the same period (2002-2011). At the geographical level, the distribution of individual firms is extremely highly correlated with the overall distribution of firms — see Figure 2.
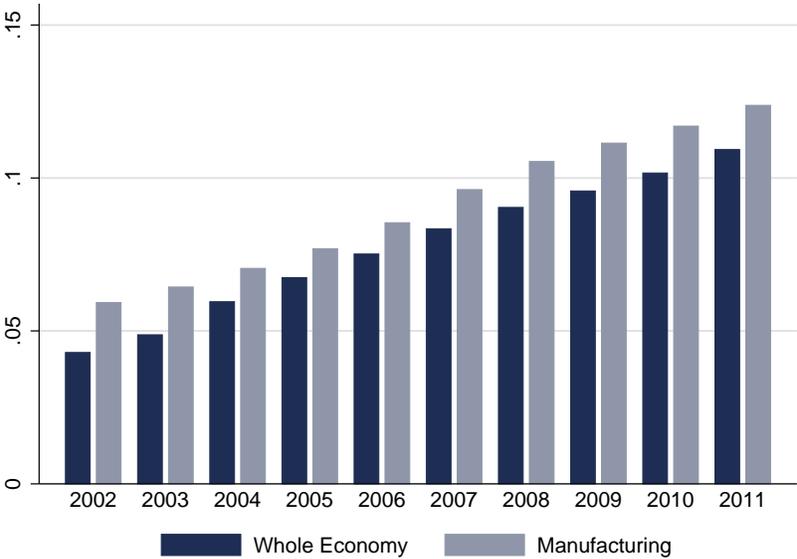
Table 1 reports the evolution of the foreign presence as both residents and individual entrepreneurs for the 20 most represented nationalities' in 2011 (i.e. the final year of our analysis). Concerning the distribution of countries of origin, immigrant entrepreneurs are significantly more concentrated than diasporas. Entrepreneurs from China alone account for 47% of the total number of individual firms in manufacturing, followed by Romania and Switzerland.[19] As expected, the evolution of foreign residents and immigrant entrepreneurs time series is strongly if not perfectly correlated — the Pearson (unconditional) correlation is 76%. Finally, Table 2 reports descriptive statistics of the estimation sample for our main variables of interest. For the average province, the number of foreign residents is about 260, whereas there are fewer than nine foreign individual entrepreneurs.

5   Identification of the effects of diasporas and immigrant entrepreneurs on exports

A well-known problem in the trade-migration literature is that unobservable variables affecting immigrants' location choices may be correlated with those influencing trade, determining an endogeneity bias. The common solution

---

[19] All the main results reported in the next section are robust if China is excluded from the estimation sample.

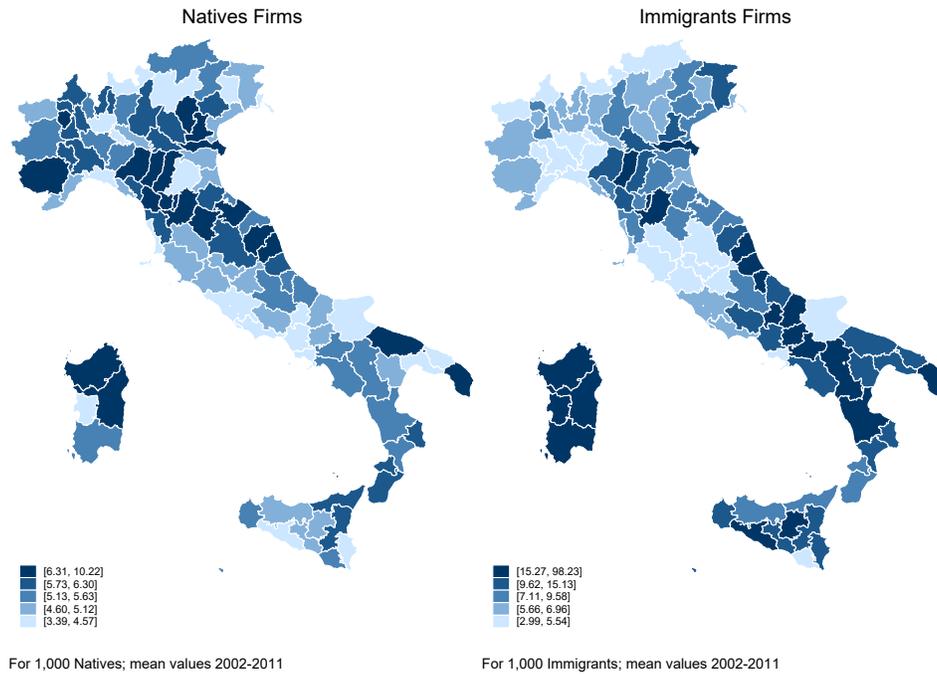Figure 1: Immigrant entrepreneurs (proportion of total individual firms)

Note. The figure plots the proportion of immigrant entrepreneurs on the total number of individual firms in the manufacturing sector.

Table 1: Individual foreign-born entrepreneurs in manufacturing and foreign residents: top 20 origin countries in 2011

ISO 3	$IE_{\#}$	IE_{Share}	$\%\Delta IE_{\#}$	$D_{\#}$	D_{Share}	$\%\Delta D_{\#}$	IE/D_{000}
CHN	14733	47.33	12.35	192867	4.87	7.60	76.39
ROU	1873	6.02	30.63	815197	20.60	19.43	2.30
CHE	1499	4.82	-2.09	7720	0.20	-1.32	194.17
MAR	1418	4.56	7.32	397448	10.05	9.34	3.57
DEU	1119	3.59	0.70	34075	0.86	-1.40	32.84
ALB	1076	3.46	8.53	440365	11.13	11.29	2.44
TUN	682	2.19	4.13	81611	2.06	3.10	8.36
FRA	646	2.08	-0.18	23366	0.59	-3.64	27.65
ARG	522	1.68	-1.98	7647	0.19	-2.26	68.26
SRB	515	1.65	5.24	85306	2.16	-0.52	6.04
EGY	440	1.41	8.61	64751	1.64	4.38	6.80
SEN	428	1.38	7.91	71688	1.81	3.46	5.97
BGD	403	1.29	16.64	79722	2.01	9.64	5.06
PAK	327	1.05	13.47	67206	1.70	8.86	4.87
VEN	308	0.99	4.10	4645	0.12	-1.27	66.31
BRA	283	0.91	8.30	36484	0.92	1.12	7.76
BEL	268	0.86	0.40	4708	0.12	-3.20	56.92
GBR	228	0.73	1.99	22105	0.56	-2.37	10.31
NGA	221	0.71	9.97	47725	1.21	3.20	4.63
UKR	200	0.64	53.88	175383	4.43	17.78	1.14
Top 20	27189	87.35	9.50	2660019	67.23	4.16	10.2
Total	31127	100.00	10.27	3956454	100.00	0.78	7.9

Note. D refers to foreign residents (diaspora) while IE refers to foreign-owned individual firms in manufacturing (immigrant entrepreneurs). Countries are ordered according to the individual firm's ranking in the top 20 nationalities in 2011. $\%\Delta IE_{\#}$ and $\%\Delta D_{\#}$ denote average annual growth rates between 2002 and 2011 in immigrant entrepreneurs and diaspora, respectively. IE/D_{000} represents the number of immigrant entrepreneurs per 1000 foreign residents. The description of ISO 3 country codes is reported in Appendix A.

Figure 2: Individual firms per 1000 inhabitants

Natives Firms

Immigrants Firms

[6.31, 10.22]
[5.73, 6.30]
[5.13, 5.63]
[4.60, 5.12]
[3.39, 4.57]

For 1,000 Natives; mean values 2002-2011

[15.27, 98.23]
[9.62, 15.13]
[7.11, 9.58]
[5.66, 6.96]
[2.99, 5.54]

For 1,000 Immigrants; mean values 2002-2011

Note. The figure plots the incidence of individual firms against the total number of inhabitants, natives (left panel), or immigrants (right panel) by province. Immigrant and native firm distributions correlate positively at the province-by-country level. The log of foreign-owned individual firms is regressed on the log of native firms given a significant coefficient of 0.053 (standard error = 0.012) and an R-squared of 0.135, after controlling for origin-by-time and province fixed effects.

15

Table 2: Descriptive statistics

Variable	$N^{>0}$	$Zeros^{\%}$	Mean	p25	p50	p75
$Exports^{Mln}_{US\$}$ (X)	141091	29.8	27.86	0.09	0.95	8.82
Diasporas (D)	119729	40.4	260.58	3	14	74
Immigrant entrepreneurs (IE)	32889	83.6	8.49	1	2	5
$Dist^{Km}$ (d)	200850		6068.71	2576.56	5125.04	8484.74

Note. $Exports^{Mln}_{US\$}$ are exports in manufacturing in US\$ millions, *diaspora* refers to foreign residents, *immigrant entrepreneurs* refers to foreign-owned individual firms in manufacturing, $N^{>0}$ stands for the number of non-zero cells at province-country-year level, and $Dist^{Km}$ is distance in kilometers. All variables are in levels (estimation period 2002-2011). Mean, p25, p50, p75 stand for the average, first quartile, second quartile, and third quartile, respectively.

to this problem is to leverage a presumably exogenous source of variation in immigrants' locations using an IV strategy. When the independent variable of interest is the stock of immigrants, this variation is generally provided by immigrant *enclaves*. The idea is to use the past geographical distribution of immigrants by ethnicity to apportion annual nationwide flows of immigrants to different regions. This was the instrument proposed by Altonji and Card (1991) in their study of the effect of immigrants on natives' labour market outcomes, which has been widely used in trade-migration studies (see, for instance, Peri and Requena-Silvente, 2010, Bratti *et al.*, 2014).[20] The underlying idea is that the presence in Italy of individuals from a given foreign country may provide useful information about the host country to new potential immigrants from the same country, reduce relocation costs, and increase the potential benefits of migration. It must be noted, however, that endogeneity concerns may be fewer in migration-trade studies than in studies addressing the effects of immigrants in the host country's labour market, since it might be much easier for migrants to observe (or predict) the

---

[20] Some recent studies exploit quasi-natural experiments provided by the (presumably) random allocation of refugees across US states (Parsons and Vézina, 2014, Steingress, 2017). Even if considering specific episodes of migration generally allows for a convincing identification, refugees are only a fraction of total immigrants, often very different from economic immigrants (e.g. often they cannot work before their refugee status has been recognized and have fewer contacts with natives), and their effect on trade may not be easily generalizable to the latter.

state of a province's labour market, and to locate themselves in high-demand markets, than to predict that a given province may provide the ideal environment to set up an exporting firm. Endogeneity concerns cannot, however, be completely ruled out, and therefore we resort to an IV estimation strategy.

In more detail, we use the distribution of immigrants' requests for residence permits[21] (*permessi di soggiorno*) provided by the Ministry of Interior in 1995 to apportion to provinces the flows of immigrants by ethnicity at the nationwide level, obtaining a *predicted stock of immigrants*, which is used as an instrument for the observed stock. Let us define D_{ijt} as the diaspora (i.e. the number of immigrants) from country j located in province i at time t, and D_{jt} the total stock of immigrants from country j at time t in Italy. Then the proportion of total immigrants of nationality j residing in province i at time t can be defined as:

$$wh_{ijt} = \frac{D_{ijt}}{D_{jt}}. \tag{6}$$

After defining D_{j0} as the total stock of immigrants from country j in Italy in the first year of the time interval (time zero, i.e. 2002), the predicted stock of immigrants is:

$$\widehat{D}_{ijt} = wh_{ij95}D_{j0} + wh_{ij95}\sum_{q=0}^{t} F_{jq} = wh_{ij95}(D_{j0} + \sum_{q=0}^{t} F_{jq}) = wh_{ij95}D_{jt} \tag{7}$$

where wh_{ij95} is the lagged distribution of immigrants by nationality across provinces computed using residence permits, and F_{jq} is the total net inflow of immigrants from country j to Italy at time q. The instrument is then given by the product of two terms; the first (wh_{ij95}) exhibits trading-pair variation, and the second (D_{jt}) country by time variation. The validity of this instrument is generally argued stressing the lagged nature of the weights used (wh_{ij95}) and the aggregate nature of nationwide immigration flows, which should ensure their orthogonality to province-country-year demand and supply shocks, which may also affect trade during the estimation period. From equation (7) it is also clear that, when estimating a linear-in-logs specification (i.e. double-log specification) of the gravity model, the enclave instrument is not compatible with the inclusion of both trading pair (ij) and country-year (jt) fixed effects, which would absorb the whole instrument's variation.[22]

---

[21] A residence permit, issued by the Italian Ministry of the Interior, is required for all foreign nationals (non-EU citizens) who plan to stay more than 3 months in the country.

[22] Indeed, $\ln(wh_{ij95}D_{jt}) = \ln(wh_{ij95}) + \ln(D_{jt})$.

We use a similar idea to build an instrument for the stock of immigrant entrepreneurs. In particular, we use the province-level distribution of immigrant entrepreneurs by country of origin in 2000 (the first year for which we have data on immigrant entrepreneurs from *Infocamere*) and apportion the time-varying nationwide stocks of firms by ethnicity to provinces according to these weights. This instrument should be valid by the same arguments used for the immigrant enclave instrument. As to relevance, there are three main reasons why the predicted stock of immigrant entrepreneurs should correlate well with the observed stock: (i) some of the firms operating in the year the weights are computed will still be active during the estimation period; (ii) production linkages between immigrant entrepreneurs may induce a co-location of entrepreneurs, e.g. immigrant entrepreneurs may prefer suppliers of the same nationality; (iii) co-location may also be induced by imitation behavior, i.e. new immigrant entrepreneurs may set up firms after observing that their co-nationals are running successful businesses.

A possible concern with the shift-share instrument is that it attributes a zero value to all stocks of immigrants or immigrant entrepreneurs that were not present in a province in the base year. Thus the instrument cannot affect the stock of immigrants from those communities during the estimation period. The issue becomes more severe the more the base year is lagged in time. This is less of a problem where the effects are not heterogeneous by either ethnicity or province of location, but it may affect the IV estimates if the effects are heterogeneous along these dimensions. The compliers with the instrument are indeed those trading pairs (country-province) whose stocks of immigrants and immigrant entrepreneurs are moved by the enclave mechanism. This cannot happen, for instance, for recently established communities, i.e. communities that were not present in a given province in the base year. For this reason, we check the robustness of our findings using an alternative instrument, which should be less sensitive to the issue just described. Namely, we estimate gravity models for diaspora and the stock of immigrant entrepreneurs using Poisson pseudo maximum likelihood (PPML):[23]

$$D_{ijt} = \exp(\boldsymbol{\delta}_{it} + \boldsymbol{\theta}_{jt} + \gamma_1 \ln(1 + D_{ij0} * D_{jt}) + \gamma_2 \ln d_{ij} + \gamma_3 Border_{ij})v_{ijt} \quad (8)$$

$$IE_{ijt} = \exp(\boldsymbol{\delta}_{it} + \boldsymbol{\theta}_{jt} + \phi_1 \ln(1 + IE_{ij0} * IE_{jt}) + \phi_2 \ln d_{ij} + \phi_3 Border_{ij})u_{ijt} \quad (9)$$

<hr>

[23] Ortega and Peri (2014) use a similar strategy to estimate the causal effect of migrants and trade on income per capita.

where the control variables have the same meaning as above. After estimating the two gravity equations, we compute the predicted values excluding from the linear predictor *demand pull factors*, i.e. the province-year fixed effects. Identification is based on one exclusion restriction for each equation (namely $\ln(1 + D_{ij0} * D_{jt})$ and $\ln(1 + IE_{ij0} * IE_{jt})$ for the immigrants and immigrant entrepreneurs equations, respectively), and the same exogeneity assumption as the shift-share instrument. However, one advantage of using the two auxiliary regressions is that, while the value of the instrument is zero, whenever D_{ij0} or IE_{ij0} is zero for the shift-share instrument, the additional regressors used in the PPML models can predict non-zero stocks of immigrants and immigrant entrepreneurs even if a given ethnicity was not present in a province in the base year. Of course this comes at a cost. Note that, given that the auxiliary regressions include the same controls as the second stages of two-stage least squares regression analysis (2SLS), for the ethnicities that were not present in a province in the base year, identification is based only on the non-linearity of PPML.[24]

As we mentioned, where there are heterogeneous effects the two instruments may produce quite different results. If the pro-trade effect, for instance, is higher for older communities of immigrants, i.e. those who established themselves early in Italy, then the shift-share instrument, which weights these communities more, may return higher IV estimates of the effect of immigrants than the PPML instrument.

6 Results

6.1 Dyadic effects

A first set of results, in which endogeneity is not addressed, is reported in Table 3. Column (1) reports the specification commonly used in the gravity equations augmented with the diaspora, which does not distinguish between entrepreneurs and non-entrepreneurs. In what follows, we will often refer to non-entrepreneur immigrants simply as 'immigrants,' or *diaspora*. The ordinary least squares (OLS) estimates on the sample of observations with positive exports return a coefficient of 0.115, which is in line with the findings of the existing literature (see, for instance, Figure 1 in Bratti *et al.*, 2014).

---

[24] Even if the instruments are generated by a regression, 2SLS standard errors do not need any further adjustment (Wooldridge, 2010).

Column (2) reports the OLS results where the stocks of immigrants and of immigrant entrepreneurs are entered separately in the regression. The estimated elasticities of exports with respect to immigrants and immigrant entrepreneurs are 0.096 and 0.086, respectively, in both cases statistically significant at the 1% level. According to these results, the elasticity of exports to immigrant entrepreneurs is similar in magnitude to that of exports to non-entrepreneur immigrants. Estimating the log-log model only on non-zero export observations may induce a bias in our estimates. Instead of adding a constant to exports before taking the logarithm, and using the log-log specification, in column (3) we use the PPML estimator proposed by Silva and Tenreyro (2006).[25] The estimated elasticities remain positive, statistically significant at the 1% level, and of a similar order of magnitude: 0.069 for immigrants and 0.054 for immigrant entrepreneurs, respectively. Despite the incidence of zeros in the sample (29.7% of export flows), PPML results on positive observations (column 5) are almost indistinguishable from those on the full sample.[26] On the grounds of the similarity between the estimates including and excluding zeros, and the convergence problems we encountered estimating an IV-PPML model with a high number of fixed effects, we use a log-log specification estimated on strictly positive export observations as our preferred model for the remainder of the analysis.[27]

Back-of-the-envelope calculations from the estimates in column (2) indicate that, for the median exporting province, a 1% increase in the population of non-entrepreneur immigrants (D) would increase exports by roughly US\$ 652 (keeping immigrant entrepreneurs constant),[28] whereas an inflow of 1% immigrant entrepreneurs (IE) would increase exports by US\$ 510 (keeping

---

[25] The debate on the most appropriate non-linear estimator to be applied when zeros are a relevant proportion of trade flows is still wide open. See De Benedictis and Taglioni (2011), Baltagi *et al.* (2014) and Head and Mayer (2015) on this specific point of the gravity literature. If the stock of immigrants and immigrant entrepreneurs are included in the logarithm in the gravity equation estimated with PPML, their coefficients can be interpreted as elasticities.

[26] As noted in Silva and Tenreyro (2006), this suggests that the difference between OLS and PPML may be driven by heteroscedasticity rather than truncation.

[27] See Head and Mayer (2015) on that. In a recent paper, Aleksynska and Peri (2014) proceed similarly, and after comparing PPML and linear estimates use the linear-in-logs as their preferred econometric specification.

[28] This figure is obtained by multiplying the median export flow (US\$ 0.95 Million) by the estimated elasticity (0.069) and dividing by 100. All the following back-of-the-envelope computations are done in the same way.

non-entrepreneur immigrants constant). On the other hand, keeping constant the overall number of immigrants $(D + IE)$, but allowing one migrant to become an entrepreneur, would increase the median export flow by roughly US\$ 5769.[29]

Table 3: Baseline estimates, sample 2002-2011

Dependent variable:	OLS		PPML		
	$\ln(X_{ijt}) > 0$		$X_{ijt} \geq 0$		$X_{ijt} > 0$
	(1)	(2)	(3)	(4)	(5)
$\ln(1+ D_{ijt}+ IE_{ijt})$	0.115***		0.091***		
	(0.009)		(0.020)		
$\ln(1+ D_{ijt})$		0.096***		0.069***	0.069***
		(0.009)		(0.020)	(0.019)
$\ln(1+ IE_{ijt})$		0.086***		0.054***	0.057***
		(0.016)		(0.020)	(0.020)
$\ln(d_{ij})$	-1.109***	-1.125***	-0.697***	-0.709***	-0.721***
	(0.072)	(0.072)	(0.123)	(0.127)	(0.125)
Border_{ij}	0.445**	0.386*	-0.247**	-0.216**	0.209*
	(0.200)	(0.200)	(0.110)	(0.108)	(0.107)
Observations	141,091	141,091	200850	200850	141,091
R-squared	0.796	0.796	0.890	0.891	0.890
Fixed effects	it; jt	it; jt	it; jt	it; jt	it; jt

Note. X are export flows. D refers to foreign residents while IE refers to foreign-owned individual firms in manufacturing. In all regressions, standard errors in parentheses are clustered at province-by-country and province-by-year level. Double clustering in PPML estimates has been performed using `clus nway` (Kleinbaum *et al.*, 2013). ∗∗∗, ∗∗, ∗ statistically significant at the 1%, 5% and 10% level, respectively. i, j, t are province, country and time subscripts.

As we argued in the previous section, the OLS estimates may be subject to an endogeneity bias, and we seek to tackle this issue by resorting to an IV (2SLS) estimation strategy. Table 4 reports the first-stage results of the

---

[29] Note that one migrant represents 7% of the median stock of (non-entrepreneur) immigrants (14 immigrants) but 50% of the median stock of immigrant entrepreneurs (2 entrepreneurs). This effect is derived by multiplying the median export flow (US\$ 0.95 Million) by the estimated elasticity (0.054) times the change in the stock of immigrant entrepreneurs (0.118), minus the product of the median export flow (US\$ 0.95 Million) times the elasticity (0.069) times the change in the stock of immigrants (-0.004). Moreover, US\$ 5769 corresponds to an increase of 0.61% for the median export flow.

Table 4: First-stage 2SLS estimates for immigrants and immigrant entrepreneurs, estimation sample 2002-2011

Dependent variable:	$\ln(1 + \text{D}_{ijt}+ \text{IE}_{ijt})$ (1)	$\ln(1 + \text{D}_{ijt})$ (2)	$\ln(1 + \text{IE}_{ijt})$ (3)	$\ln(1+ \text{D}_{ijt})$ (4)	$\ln(1+\text{IE}_{ijt})$ (5)
$\ln(1 + \text{D}_{ijt}+ \text{IE}_{ijt})^{Card}_{1995}$	0.500***				
	(0.009)				
$\ln(1 + \text{D}_{ijt})^{Card}_{1995}$		0.484***	0.047***		
		(0.010)	(0.003)		
$\ln(1 + \text{IE}_{ijt})^{Card}_{2000}$		0.131***	0.671***		
		(0.015)	(0.012)		
$\ln(1 + \text{D}_{ijt})^{\widehat{PPML}}$				0.772***	0.079***
				(0.013)	(0.004)
$\ln(1 + \text{IE}_{ijt})^{\widehat{PPML}}$				0.063***	0.956***
				(0.022)	(0.013)
$\ln(\text{d}_{ij})$	-0.027	-0.165***	0.162***	-0.233***	-0.034*
	(0.044)	(0.046)	(0.020)	(0.046)	(0.020)
Border_{ij}	0.674***	0.836***	0.243**	0.534	0.086
	(0.169)	(0.162)	(0.107)	(0.165)	(0.130)
Observations	141,091	141,091	141,091	141,091	141,091
R-squared	0.894	0.891	0.806	0.854	0.721
F-statistic	2916		1159		186.7
Fixed effects	it; jt	it; jt	it; jt	it; jt	it; jt

Note. D refers to foreign residents while IE refers to foreign-owned individual firms. Standard errors in parentheses are double clustered at province-by-country and province-by-year level. Double clustering in PPML estimates has been performed using `clus nway` (Kleinbaum *et al.*, 2013). $***$, $**$, $*$ statistically significant at the 1%, 5% and 10% level, respectively. i, j, t are province, country and time subscripts.

Table 5: Second-stage 2SLS estimates, sample 2002-2011

Dependent variable:	2SLS $\ln(X_{ijt} > 0)$		
	(1)	(2)	(3)
$\ln(1 + D_{ijt} + IE_{ijt})$	0.172***		
	(0.017)		
$\ln(1 + D_{ijt})$		0.157***	0.174***
		(0.019)	(0.017)
$\ln(1 + IE_{ijt})$		0.062**	0.059**
		(0.024)	(0.023)
$\ln(d_{ij})$	-1.086***	-1.090***	-1.081***
	(0.071)	(0.073)	(0.073)
$Border_{ij}$	0.356*	0.300	0.272
	(0.196)	(0.197)	(0.195)
Observations	141,091	141,091	141,091
Fixed effects	it ; jt	it ; jt	it ; jt
IV	$Card_{95}^{D+IE}$	$Card_{95}^{D}$	$PPML^{\widehat{D}}$
		$Card_{00}^{IE}$	$PPML^{\widehat{IE}}$

Note. X are export flows. D refers to foreign residents while IE refers to foreign-owned individual firms. Standard errors in parentheses are double clustered at province-by-country and province-by-year level. Double clustering in PPML estimates has been performed using `clus nway` (Kleinbaum *et al.*, 2013). ***, **, * statistically significant at the 1%, 5% and 10% level, respectively. i, j, t are province, country and time subscripts.

2SLS estimates on positive-export observations using the shift-share instrument. As with the OLS estimates, we start with the specification pooling all immigrants irrespective of their entrepreneur status. The first-stage coefficient is 0.5 and statistically significant at the 1% level. The F-statistic shows no sign of a weak instrument problem. Columns (2) and (3) refer to the specification including immigrants and immigrant entrepreneurs separately, showing the two first stages using the shift-share instrument. The elasticity of the stock of immigrants to the predicted stock of immigrants is 0.484 and to the predicted stock of immigrant entrepreneurs is 0.131, in both cases statistically significant at the 1% level. Hence, immigrants tend not only to follow compatriots' past location choices, but also to settle where there are more immigrant entrepreneurs, as presumably they find better employment opportunities. Similarly, the elasticities of the stock of immigrant entrepreneurs to the predicted stock of immigrant entrepreneurs and the predicted stock of immigrants are 0.671 and 0.047, respectively. There is, therefore, evidence that immigrant entrepreneurs tend to set up firms where there are larger communities from their country of origin, and where there is already a larger presence of immigrant business created by co-nationals. The joint F-statistics for the two instruments are very high, and again there is no evidence of a weak instrument problem. Finally, columns (4) and (5) show the first stages using the instruments built with the auxiliary gravity PPML model for immigrants and immigrant entrepreneurs. Again, the instruments are highly statistically significant, and the PPML instrument seems to do a better job of separating the effects of immigrants and immigrant entrepreneurs in each first stage. Indeed, both the coefficient on the predicted stock of immigrants in the immigrants equation and the coefficient of the predicted stock of foreign entrepreneurs in the entrepreneurs equation increase.

Table 5 reports the second-stage of 2SLS. Column (1), pooling immigrants and immigrant entrepreneurs, shows a 0.17 elasticity of exports to diaspora, almost 50% larger than that obtained with the OLS estimates. The increase in the elasticity when using 2SLS seems to be at odds with an endogenous location story, i.e. more immigrants and immigrant entrepreneurs locate where there are more opportunities for international trade. By contrast, the results may be consistent with a 'negative selection' story. That is immigrants may decide to become entrepreneurs especially in less competitive provinces, which also tend to export less. Column (2) reports the second stage of the specification, including diaspora and immigrant entrepreneurs

separately; their estimated elasticities are 0.157 and 0.062, respectively. Finally, column (3) shows the second stage of the estimates using the instruments based on the auxiliary PPML regressions,[30] where the elasticity of exports to immigrants is 0.174 and to immigrant entrepreneurs 0.059. Thus, the two IV strategies give very similar results. Using the second strategy, for instance, the estimated effect of a 1% increase in migrants is relatively higher with respect to OLS estimation and corresponds to US$ 1644 for the median export flows, whereas the estimated effect of a 1% increase in immigrant entrepreneurs shrinks to US$ 557. The latter implies that the expected pro-trade effect of a migrant becoming an entrepreneur (keeping $D+IE$ constant) amounts to US$ 5946 at the median export flow.

6.1.1 Robustness checks

A potential threat to our identification strategy may come from the omission of country-pair unobservables that are correlated with both exports and immigrant entrepreneurs' locations.[31]

Although we do not have any obvious omitted variable that can severely bias our estimates in mind, as province-level unobservables acting on all ethnicities (e.g. quality of province institutions) and country-level unobservables affecting all provinces (e.g. cultural proximity) are controlled for in our regressions by province-year and country-year fixed effects, respectively, we cannot completely discard the possibility of omitted-variable bias.

For this reason, in Table 6 we report some robustness checks that aim to control for potential ij confounders. The idea is to cluster foreign countries according to some similarity criteria, create bilateral fixed effects according to these criteria (iJ), and include them in the 2SLS estimations. Column (1) defines the cluster by quartiles of distance (in kilometers) and column (2) by quartiles of language proximity (Melitz and Toubal, 2014).[32] These estimates are meant to capture confounders such as cultural proximity acting at the trading-pair level, e.g. the level of trust that culturally proximate immigrants enjoy in different provinces. Albeit with some slight changes in

---

[30] The results of these regressions are reported in Table B1 in Appendix B.

[31] As we have mentioned, the shift-share estimation strategy proposed by Altonji and Card (1991) is incompatible with including country-pair fixed effects and country-year fixed effects in gravity models, which absorb the whole instrument's variability.

[32] In particular, we use the unadjusted proximity index based on language trees, but the results are robust to the use of alternative indexes.

Table 6: IV Estimation: Robustness Checks, sample 2002-2011

Dependent Variable:	2SLS $\ln(X_{ijt} > 0)$				
	J: quartile of distance km (1)	J: quartile of language proximity (2)	Sample excludes main emigration countries (3)	Sample excludes China (4)	Accounts for spatial dependence (5)
$\ln(1 + \mathrm{D}_{ijt})$	0.185***	0.179***	0.178***	0.173***	0.174***
	(0.016)	(0.017)	(0.018)	(0.017)	(0.017)
$\ln(1 + \mathrm{IE}_{ijt})$	0.062***	0.051**	0.051**	0.064**	0.059**
	(0.022)	(0.023)	(0.026)	(0.026)	(0.027)
$\ln(\mathrm{d}_{ij})$	-0.937***	-0.878***	-1.331***	-1.084***	-1.081***
	(0.073)	(0.078)	(0.079)	(0.073)	(0.069)
Border_{ij}	0.267	0.404**	-0.023	0.276	0.272
	(0.189)	(0.189)	(0.322)	(0.196)	(0.206)
Observations	141,091	141,091	132,915	140,051	141,091
F-statistic	1564	1512	1464	1427	569.1
Fixed effects	it, jt, iJ	it, jt, iJ	it, jt	it, jt	it, jt
Cluster	ij, it	ij, it	ij, it	ij, it	ij, Rt

Note. X are export flows. D refers to foreign residents while IE refers to foreign-owned individual firms. J identifies an aggregation of nationalities of the same quartile of geographical distance to i (column 1) and quartile of language proximity to i (column 2). Column 3 excludes the top destination for Italian emigrants, covering 75% of total emigrants in year 2012 (in decreasing order: Argentina, Germany, Switzerland, France, Brazil, Belgium, the US, and the UK); column 4 excludes China from the estimation sample. Finally, column 5 controls for cross-sectional dependence in the error terms by clustering standard errors at the province-by-country and area-by-year level (macro areas: North-East, North-West, Centre, South). Double clustering in PPML estimates has been performed using `clus nway` (Kleinbaum *et al.*, 2013). The set of instruments includes $\mathrm{PPML}^{\widehat{D}}$ and $\mathrm{PPML}^{\widehat{IE}}$. Standard errors in parentheses. $***$, $**$, $*$ statistically significant at the 1%, 5% and 10% level, respectively. i, j, t are province, country and time subscripts.

the magnitude of the coefficients, the results are very robust in qualitative terms. Both the diaspora and the population of immigrant entrepreneurs have significant positive effects on exports, with elasticities in the ranges of 0.179-0.185 and 0.051-0.062, respectively.

Another example of omitted variables potentially biasing our estimates is the stock of Italian emigrants living in immigrants' countries of origin. Indeed, the trade-generating effect may originate from the Italian diaspora providing potential immigrants in the home country with useful information on suitable locations in Italy to live in or to set up businesses (Bastos and Silva, 2012, Lodefalk, 2016), and at the same time helping Italian firms to overcome informational barriers related to export activities. We check the robustness to this potential threat by excluding from the estimation sample the top destinations of Italian emigrants (accounting for up to 75% of Italian emigrant stock)[33] and the results are presented in column (3). The estimates are robust, and the coefficients both on immigrants and on immigrant entrepreneurs remain virtually unchanged.

In column (4), we check the sensitivity of our estimates to excluding China from the estimation sample, which accounts for a large proportion of both the total population of immigrants and, in particular, entrepreneurs. The coefficient on immigrant entrepreneurs becomes larger, reaching 0.064 compared with 0.059 in the baseline estimates (column (3) of Table 5). Thus, we conclude that our results are not driven by trade with China.

Finally, in column (5) we check the robustness of our results to allow for cross-sectional dependence in the error terms and cluster standard errors at the province-by-country and area-by-year level (four macro-areas are considered: North-East, North-West, Centre, South). Allowing for spatial dependence at the Italian macro-area level has virtually no effect on the coefficients' statistical significance.

6.2 Aggregate province-level effects

We conclude our analysis by attempting to throw light on other potential effects of diasporas and immigrant entrepreneurs, which could affect export of province i towards all destinations, i.e. the effects running through labour costs and average firm productivity (w_{it} and $\bar{\phi}_{it}$, respectively) described in

---

[33] The year or reference is 2012 but countries of destination have been fairly stable in the last sixty years; data source: the Registry of Italians Resident Abroad (AIRE).

Table 7: Aggregate analysis: main results

Dependent variable:				Estimated Province-by-Year Fixed Effect: $\hat{\delta}_{it}$		
	(1)	(2)	(3) WLS	(4)	(5)	(6) 2SLS
ln(D_{it})	1.100***	0.951***	0.279*	0.170	0.224	0.128
	(0.137)	(0.146)	(0.157)	(0.170)	(0.204)	(0.298)
ln(IE_{it})	0.238**	0.313***	0.292***	0.268***	0.286***	0.237**
	(0.100)	(0.105)	(0.084)	(0.090)	(0.089)	(0.116)
ln(TFP)$_{i2002}$		4.466**	3.444**	2.697**	2.581**	2.560*
		(1.846)	(1.428)	(1.282)	(1.300)	(1.314)
ln(MNE)$_{i2002}$			0.683***	0.677***	0.678***	0.721***
			(0.097)	(0.092)	(0.092)	(0.107)
(DIV)$_{i2002}$				6.980**	5.695	7.638*
				(3.373)	(3.886)	(4.447)
City$^{Pop>pc75}$					-0.153	-0.041
					(0.178)	(0.229)
Observations	918	918	918	918	918	918
R-squared	0.694	0.719	0.817	0.825	0.826	0.825
F-statistic						12.82
Stock-Yogo 10%						7.03
IV						Card$_{95}^{M}$; Card$_{00}^{EF}$

Note. Total factor productivity (TFP) is estimated using Wooldridge (2009) generalized methods of moments approach using AIDA data (Bureau van Dijk). MNE represents the (log) of the number of multinational enterprises in a given province i at the beginning of the period (source: ICE-Reprint). DIV represents the log of the inverse of the Herfindahl-Hirschman index measuring diversity in the ethnic composition of residents. Observations are weighted by the inverse of the standard error of $\hat{\delta}_{it}$. Milan is the excluded province. The estimation sample covers the period 2003-2011. Year fixed effects are included in all regressions. Standard errors in parentheses are clustered at province level. ***, **, * statistically significant at the 1%, 5% and 10% level, respectively.

Section 3. We attempt this by regressing the province-year fixed effects estimated using 2SLS ($\widehat{\delta_{it}}$) on the populations of immigrants and immigrant entrepreneurs, immigrant diversity, and some additional controls. Namely, we estimate the following equation:

$$\widehat{\delta_{it}} = \beta_1 \ln(D_{i,t}) + \beta_2 \ln(IE_{i,t}) + \beta_3 \ln(TFP_{i2002}) + \beta_4 \ln(MNE_{i2002}) \quad (10)$$
$$+ \beta_5 \ln(DIV_{i2002}^N) + City_{i2002} + \gamma_t + \varepsilon_{i,t}$$

where i is the subscript for Italian provinces (NUTS-3) and t stands for time. $D_{i,t}$ and $IE_{i,t}$ are the total stocks of immigrants and immigrant entrepreneurs located in province i at time t, respectively. In the light of the fact that the dependent variable is nothing more than the 'outward' multilateral resistance term and captures the relative 'competitiveness' of exports from province i, we add some additional controls at the beginning of the period, which account for different initial competitiveness levels of Italian provinces. Namely, we include total factor productivity (TFP) of manufacturing firms,[34] the proportion of foreign multinational enterprises (MNEs),[35] a measure of the diversity of the local labour force (DIV),[36] and a dummy variable equal to one for provinces with above-median populations ($City$).[37] All these control variables are evaluated in the first available year, i.e. 2002, to make them predetermined with respect to the estimation period.[38] Accordingly, we limit the estimation to the period 2003-2011. Finally, γ_t is a year fixed effect, whereas ϵ_{ijt} is an error term clustered at the province level to account for serial correlation in trade. Since the fixed effects that represent our dependent variables have been estimated from a regression, observations are weighted by the inverse of the standard errors of the $\widehat{\delta_{it}}$s.[39]

Results are reported in Table 7. The weighted least square (WLS) estimates in columns (1)-(5) show that the positive association between immigrants and the dependent variable, which we consider a proxy of provinces' overall competitiveness, is reduced in size and loses statistical significance

---

[34] TFP is estimated using Wooldridge (2009) GMM approach on firm-level balance sheets from AIDA (Bureau van Dijk).

[35] This measures the number of MNEs in a given province i. Source: ICE-Reprint.

[36] Computed as the inverse of the Herfindahl-Hirschman index on total residents of a given province (including natives).

[37] Changing the population threshold does not affect the main results.

[38] Since diaspora and immigrant entrepreneurs during the estimation period may affect them.

[39] Results are fully robust to the use of FGLS; see Table B2 in Appendix.

when the controls are progressively included in the regression. A big drop in the coefficient is caused by the inclusion of the stock of MNEs and the population diversity index, both of which are positively associated with competitiveness. In contrast, the coefficient on immigrant entrepreneurs is very stable across all specifications, and varies in the range of 0.237-0.313. In the last column, we present the 2SLS estimates using the shift-share instrument.[40] Again, the effect of immigrant entrepreneurs on province competitiveness is positive and significant, with an estimated elasticity of 0.237, while the elasticity to immigrants is positive but statistically insignificant. All in all, the results in this section point to a non-negligible positive effect of the population of immigrant entrepreneurs on a province's capacity to export towards all destinations, which adds to the dyadic effects estimated in Section 6.1.

7 Concluding remarks

When establishing themselves in a region, immigrants bring knowledge about their countries of origin and retain long-lasting relationships with their co-nationals who are left behind. Such knowledge and contacts may partly spill over to native entrepreneurs and help them to export their products abroad. Indeed, immigrants help natives to overcome the informational barriers that make it costly to enter foreign markets, or they may substitute poor market institutions, for instance by helping with contract enforcement. A number of studies have already reported evidence consistent with this hypothesis by demonstrating a positive causal effect of immigrants on exports.

The current paper adds to this evidence by investigating the trade-creating effect of immigrant entrepreneurs. After outlining various reasons why immigrant entrepreneurs can spur export activities in the regions in which they settle, primarily the fact that they can directly exploit their superior knowledge of the country of origin by setting up a business, we test this hypothesis using data from Italy.

Using a 10-year panel dataset on Italian provinces (i.e. NUTS-3 regions), and gathering administrative data on exports, the diaspora, and the population of entrepreneurs by country of origin, we estimate augmented gravity

---

[40] Indeed, the PPML instrument cannot be used when the predicted values are aggregated at the same level of the multilateral-resistance terms, since the fitted values will be equal to those observed (Fally, 2015).

models for exports. The potential endogeneity of both immigrants and immigrant entrepreneurs is addressed using an instrumental variables estimator based on a shift-share instrument à la Altonji and Card (1991). As is well known, this instrument exploits the immigrant *enclaves* that were present in domestic regions in a base year and apportion to regions the annual nation-wide waves of immigrants (shift) according the relative weight of the enclaves (share). To overcome some potential weaknesses of such an instrument, which does not use for identification new immigrant or entrepreneur nationalities that were not present in the host country in the base year, we also use a second instrumental variable based on auxiliary PPML gravity models for non-entrepreneur migrants and immigrant entrepreneurs, which overcomes this limitation. The IV estimates obtained with both strategies give very similar results, and they point to large causal effects of both immigrants and immigrant entrepreneurs on exports. Using the second IV strategy (based on PPML models), the estimated effect of a 1% increase in the population of non-entrepreneur immigrants corresponds to a US$ 1644 rise in yearly manufacturing exports for the host province towards the immigrants' country of origin at the median export flows, whereas the estimated effect of a 1% increase in the population of immigrant entrepreneurs on yearly province exports in manufacturing is US$ 557. According to these estimates, transforming a (non-entrepreneur) migrant into an entrepreneur would raise yearly province exports in manufacturing by US$ 5946 at the median export flow. Our results are robust to a number of sensitivity checks, addressing potential confounding factors that may bias our estimates, such as communities of Italian emigrants living abroad or the fact that the whole effect may be driven by larger and more entrepreneurial immigrant communities (e.g. the Chinese communities).

Finally, we propose that diasporas and immigrant entrepreneurs may impact positively on province-level exports not only via the dyadic effects postulated by the business and social network mechanism but also by increasing a province's overall competitiveness, which raises its exports towards all possible destinations. We investigate this hypothesis by regressing the province-year fixed effects estimated in the export gravity equations on the diaspora, entrepreneurs, and a set of control variables. We find that increasing the population of immigrant entrepreneurs by 10% increases competitiveness (i.e. the province-year fixed effect of the gravity model) by 2.4%.

Although the regional nature of our data does not enable us to ultimately say whether the observed effects are of a private nature only, i.e. if they

are fully accounted for by the superior export performance of immigrant entrepreneurs, or are partly of a public nature, also capturing spillover effects on native firms — a question that could only be answered using firm-level data — our paper is, to the best of our knowledge, the first to document rigorously a positive causal effect of the presence of immigrant entrepreneurs on regions' export performance.

References

Aleksynska M. and Peri G. (2014) Isolating the network effect of immigrants on trade, *World Economy*, 37, 3, 434–455.

Alesina A. and La Ferrara E. (2005) Ethnic diversity and economic performance, *Journal of Economic Literature*, 43, 3, 762–800.

Altonji J.G. and Card D. (1991) The effects of immigration on the labor market outcomes of less-skilled natives, in: *Immigration, Trade and the Labor Market*, National Bureau of Economic Research, Inc, NBER Chapters, 201–234.

Anderson J.E. (2011) The gravity model, *Annual Review of Economics*, 3, 133–160.

Anderson J.E. and van Wincoop E. (2003) Gravity with gravitas: A solution to the border puzzle, *American Economic Review*, 93, 1, 170–192.

Arkolakis C., Costinot A. and Rodríguez-Clare A. (2012) New trade models, same old gains?, *American Economic Review*, 102, 1, 94–130.

Balkan B. and Tumen S. (2016) Immigration and prices: Quasi-experimental evidence from Syrian refugees in Turkey, *Journal of Population Economics*, 29, 3, 657–686.

Baltagi B.H., Egger P. and Pfaffermayr M. (2014) Panel data gravity models of international trade, in: *The Oxford Handbook of Panel Data*, Baltagi B.H., ed., Oxford Handbooks, chapter 20.

Bastos P. and Silva J. (2012) Networks, firms, and trade, *Journal of International Economics*, 87, 2, 352–364.

Becker S.O. and Egger P.H. (2013) Endogenous product versus process innovation and a firms propensity to export, *Empirical Economics*, 1–26.

Beine M., Bertoli S. and Fernández-Huertas Moraga J. (2016) A practitioners guide to gravity models of international migration, *The World Economy*, 39, 4, 496–512.

Beine M., Docquier F. and Özden Ç. (2011) Diasporas, *Journal of Development Economics*, 95, 1, 30–41.

Borjas G.J. (1994) The economics of immigration, *Journal of Economic Literature*, 32, 4, 1667–1717.

Borjas G.J. (2015) Immigration and globalization: A review essay, *Journal of Economic Literature*, 53, 4, 961–974.

Bratsberg B., Raaum O., Red M. and Schne P. (2014) Immigration wage effects by origin, *Scandinavian Journal of Economics*, 116, 2, 356–393.

Bratti M., De Benedictis L. and Santoni G. (2014) On the pro-trade effects of immigrants, *Review of World Economics (Weltwirtschaftliches Archiv)*, 150, 3, 557–594.

Briant A., Combes P.P. and Lafourcade M. (2014) Product complexity, quality of institutions and the protrade effect of immigrants, *World Economy*, 37, 1, 63–85.

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.

Clark K. and Drinkwater S. (2000) Pushed out or pulled in? Self-employment among ethnic minorities in england and wales, *Labour Economics*, 7, 5, 603–628.

Combes P.P., Lafourcade M. and Mayer T. (2005) The trade-creating effects of business and social networks: Evidence from France, *Journal of International Economics*, 66, 1, 1–29.

Cortes P. (2008) The effect of low-skilled immigration on U.S. prices: Evidence from CPI Data, *Journal of Political Economy*, 116, 3, 381–422.

De Benedictis L. and Taglioni D. (2011) The gravity model of international trade, in: *The Trade Impact of European Union Preferential policies: An Analysis through Gravity Models*, De Benedictis L. and Salvatici L., eds., Springer, chapter 4.

Dustmann C., Frattini T. and Rosso A. (2015) The effect of emigration from Poland on Polish wages, *Scandinavian Journal of Economics*, 117, 2, 522–564.

Fairlie R.W. and Lofstrom M. (2014) Immigration and entrepreneurship, in: *Handbook of the Economics of International Migration*, Chiswick B.R. and Miller P.W., eds., North-Holland, 877.

Fally T. (2015) Structural gravity and fixed effects, *Journal of International Economics*, 97, 1, 76–85.

Felbermayr G., Grossmann V. and Kohler W. (2015) Migration, international trade, and capital formation: Cause or effect?, in: *Handbook of the Economics of International Migration*, Chiswick B.R. and Miller P.W., eds., North-Holland, volume 1, chapter 18, 913 – 1025.

Giuntella G. (2012) Do immigrants squeeze natives out of bad schedules? Evidence from Italy, *IZA Journal of Migration*, 1, 1, 1–21.

Gould D.M. (1994) Immigrant links to the home country: Empirical implications for U.S. bilateral trade flows, *Review of Economics and Statistics*, 76, 2, 302–16.

Hatzigeorgiou A. and Lodefalk M. (2016) Migrants influence on firm-level exports, *Journal of Industry, Competition and Trade*, 16, 4, 477–497.

Head K. and Mayer T. (2015) Gravity equations: Workhorse, toolkit, and cookbook, in: *Handbook of international economics, Vol.4*, Gopinath, Helpman and Rogoff, eds., Elsevier.

Head K. and Ries J. (1998) Immigration and trade creation: Econometric evidence from Canada, *Canadian Journal of Economics*, 31, 1, 47–62.

Herander M.G. and Saavedra L.A. (2005) Exports and the structure of immigrant-based networks: The role of geographic proximity, *Review of Economics and Statistics*, 87, 2, 323–335.

Hiller S. (2013) Does immigrant employment matter for export sales? Evidence from Denmark, *Review of World Economics (Weltwirtschaftliches Archiv)*, 149, 2, 369–394.

Hsieh C.T., Li N., Ossa R. and Yang M.J. (2016) Accounting for the new gains from trade liberalization, Technical Report 22069, National Bureau of Economic Research.

Hunt J. and Gauthier-Loiselle M. (2010) How much does immigration boost innovation?, *American Economic Journal: Macroeconomics*, 2, 2, 31–56.

Jahn V. and Steinhardt M.F. (2016) Innovation and immigration — Insights from a placement policy, *Economics Letters*, 146, C, 116–119.

Kerr W.R. and Kerr S.P. (2017) Immigrant entrepreneurship, in: *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, Haltiwanger J., Hurst E., Miranda J. and Schoar A., eds., University of Chicago Press, Chicago, chapter 11.

Kleinbaum A.M., Stuart T.E. and Tushman M.L. (2013) Discretion within constraint: Homophily and structure in a formal organization, *Organization Science*, 24, 5, 1316–1336.

Lachenmaier S. and Woessmann L. (2006) Does innovation cause exports? Evidence from exogenous innovation impulses and obstacles using German micro data, *Oxford Economic Papers*, 58, 2, 317–350.

Lodefalk M. (2016) Temporary expats for exports: Micro-level evidence, *Review of World Economics (Weltwirtschaftliches Archiv)*, 152, 4, 733–772.

Marchal L. and Nedoncelle C. (2016) How migrant workers foster French exports, Technical Report 1028, FREIT.

Melitz J. and Toubal F. (2014) Native language, spoken language, translation and trade, *Journal of International Economics*, 93, 2, 351 – 363.

Melitz M.J. (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity, *Econometrica*, 71, 6, 1695–1725.

Mitaritonna C., Orefice G. and Peri G. (2017) Immigrants and firms outcomes: Evidence from France, *European Economic Review*, 96, C, 62–82.

Naud W., Siegel M. and Marchand K. (2017) Migration, entrepreneurship and development: Critical questions, *IZA Journal of Migration*, 6, 5, 2–16.

OECD (2010) Entrepreneurship and migrants. Report by the OECD working party on SMEs and entrepreneurship, Technical report, OECD, Paris, France.

Olney W.W. (2013) Immigration and firm expansion, *Journal of Regional Science*, 53, 1, 142–157.

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

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

Ottaviano G.I. and Peri G. (2012) Rethinking the effect of immigration on wages, *Journal of the European Economic Association*, 10, 1, 152–197.

Parrotta P., Pozzoli D. and Pytlikova M. (2014) The nexus between labor diversity and firm's innovation, *Journal of Population Economics*, 27, 2, 303–364.

Parsons C. and Vézina P.L. (2014) Migrant networks and trade: The Vietnamese boat people as a natural experiment, Economics Series Working Papers 705, University of Oxford, Department of Economics. Forthcoming in The Economic Journal.

Parsons C.R. and Winters L.A. (2014) International migration, trade and aid: A survey, *International Handbook on Migration and Economic Development*, 65.

Patel K. and Vella F. (2013) Immigrant networks and their implications for occupational choice and wages, *Review of Economics and Statistics*, 95, 4, 1249–1277.

Peri G. and 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.

Rauch J.E. (1999) Networks versus markets in international trade, *Journal of International Economics*, 48, 1, 7–35.

Rauch J.E. (2001) Business and social networks in international trade, *Journal of Economic Literature*, 39, 4, 1177–1203.

Rauch J.E. and Trindade V. (2002) Ethnic Chinese networks in international trade, *Review of Economics and Statistics*, 84, 1, 116–130.

Sahin M., Nijkamp P. and Suzuki S. (2014) Contrasts and similarities in economic performance of migrant entrepreneurs, *IZA Journal of Migration*, 3, 1, 7.

Silva J.S. and Tenreyro S. (2006) The log of gravity, *Review of Economics and Statistics*, 88, 4, 641–658.

Sparber C. (2008) A theory of racial diversity, segregation, and productivity, *Journal of Development Economics*, 87, 2, 210–226.

Steingress W. (2017) The causal impact of migration on US trade: Evidence from political refugees, Staff Working Papers 17-49, Bank of Canada.

Wooldridge J.M. (2009) On estimating firm-level production functions using proxy variables to control for unobservables, *Economics Letters*, 104, 3, 112–114.

Wooldridge J.M. (2010) *Econometric Analysis of Cross Section and Panel Data*, volume 1 of *MIT Press Books*, The MIT Press.

Zhang P. and Sanders J. (1999) Extended stratification: Immigrant and native differences in individual and family labor, *The Sociological Quarterly*, 40, 4, 681–704.

A   Appendix: ISO 3 country codes' description

Table A1: ISO 3 country codes and country names

ISO 3 code	Country name
CHN	China
ROU	Romania
CHE	Switzerland
MAR	Morocco
DEU	Germany
ALB	Albania
TUN	Tunisia
FRA	France
ARG	Argentina
SRB	Serbia
EGY	Egypt
SEN	Senegal
BGD	Bangladesh
PAK	Pakistan
VEN	Venezuela
BRA	Brazil
BEL	Belgium
GBR	United Kingdom
NGA	Nigeria
UKR	Ukraine

Note. Codes refer to the countries listed in Table 1 in the main text.

B Appendix: Additional Tables

Table B1: PPML IV auxiliary regression, estimation sample 2002-2011

Dependent variable:	D_{ijt}	IE_{ijt}
	(1)	(2)
$\ln(d_{ij})$	0.165*	0.600***
	(0.100)	(0.077)
$Border_{ij}$	0.431**	0.703***
	(0.180)	(0.171)
$\ln(1 + Card_{95}^{D})$	0.694***	
	(0.025)	
$\ln(1 + Card_{00}^{IE})$		0.890***
		(0.021)
Observations	200850	200850
R-squared	0.81	0.97
FEs	it ; ij	it ; jt

Note. D refers to foreign residents while IE refers to foreign owned individual firms. Standard errors in parentheses are clustered at province-by-country. $***$, $**$, $*$ statistically significant at the 1%, 5% and 10% level, respectively. i, j, t are province, country and time subscripts.

Table B2: Aggregate Analysis: Main results under different weighting schemes

Dependent variable:	Estimated Province-by-Year Fixed Effect: $\hat{\delta}_{it}$		
	(1) OLS	(2) WLS	(3) FGLS
ln(D_{it})	0.259	0.224	0.253
	(0.206)	(0.204)	(0.205)
ln(IE_{it})	0.308***	0.286***	0.307***
	(0.092)	(0.089)	(0.092)
ln(TFP)$_{i2002}$	3.045**	2.581**	2.987**
	(1.313)	(1.300)	(1.307)
ln(MNE)$_{i2002}$	0.712***	0.678***	0.711***
	(0.092)	(0.092)	(0.092)
(DIV)$_{i2002}$	4.112	5.695	4.274
	(4.031)	(3.886)	(4.009)
City$^{Pop>pc75}$	-0.181	-0.153	-0.177
	(0.188)	(0.178)	(0.187)
Observations	918	918	918
R-squared	0.835	0.826	0.834

Note. The estimates above replicate column (5) of Table 7 in the main text under different weighting schemes. Total factor productivity (TFP) is estimated using Wooldridge (2009) GMM approach on AIDA data (Bureau van Dijk). MNE represents the (log) of the number of multinational enterprises in a given province i at the beginning of the period (source: ICE-Reprint). DIV is the log of the inverse of the Herfindahl-Hirschman index measuring diversity in the ethnic composition of residents. Milan is the excluded province. The estimation sample covers the period 2003-2011. Year fixed effects are included in all regressions. Standard errors in parentheses are clustered at province level. ***, **, * statistically significant at the 1%, 5% and 10% level, respectively.