Immigrant entrepreneurs, diasporas and exports

Massimiliano Bratti* Luca De Benedictis† Gianluca Santoni‡

June 10, 2019

Abstract

This paper demonstrates the positive effect of immigrant entrepreneurs on manufacturing exports over and above that of diasporas. Using small-scale regional administrative data, our instrumental variables estimates of export gravity models imply that ceteris paribus, i.e. holding constant the total number of immigrants, the expected pro-trade effect of a migrant becoming an entrepreneur amounts to an average increase of US$ 5,946 in the export flows towards her country of origin. Besides these dyadic effects, immigrant entrepreneurs unlike non-entrepreneurial immigrants raise a region's overall competitiveness and export flows towards other destinations as well.

Keywords: Exports, gravity model, immigrant entrepreneurs, Italy.
JEL Classification: F10, F14, F22, R10.

Forthcoming in *Journal of Regional Science*. doi: 10.1111/jors.12455

*European Commission Joint Research Centre (Ispra, Italy); DEMM, Università degli Studi di Milano (Milan, Italy); *IZA* (Bonn, Germany); *LdA* (Milan and Turin, Italy). massimiliano.bratti@unimi.it

†*Rossi-Doria Center*, University Roma Tre (Rome, Italy); DED, University of Macerata (Macerata, Italy); Luiss (Rome, Italy). luca.debenedictis@unimc.it

‡Corresponding Author. Cepii (Paris, France). gianluca.santoni@cepii.fr

1 INTRODUCTION

The prevailing explanation for the positive effect of immigrants on exports is the one proposed by Rauch (1999, 2001) and Rauch and Trindade (2002): 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 with them the knowledge of their home country's economy and institutions, as well as kinship links that endure in spite of distance and time. This knowledge and these social ties can be exploited by the host country's entrepreneurs, who can use them to reduce the sunk cost of exporting to the immigrants' countries of origin (e.g. information costs[1]) and/or to cut the marginal cost of exporting.[2] Since the knowledge and international links embodied in immigrants are largely country-specific, the reduction in trade costs acts mainly at the bilateral level. This narrative is fully consistent with the empirical evidence coming from several national cases (Combes *et al.*, 2005, Peri and Requena-Silvente, 2010, Bratti *et al.*, 2014).

Most of the literature has focused on co-ethnic population networks (namely the effect of *diasporas*, a.k.a. the size of local ethnic communities) and has interpreted their significant effect on trade in terms of an information flow, or a spillover effect, from immigrants to natives. Gould (1994, p. 314), for instance, concludes his seminal paper by stating that "Immigrants convey knowledge spillovers that can reduce information costs to economic agents who do not migrate. These spillovers reveal value-creating production and trade opportunities and utility-increasing consumption opportunities for the non-migrants in both countries."[3] However, the individuals who are the natural candidates to exploit the migration-trade nexus are those who have business-related knowledge and business contacts abroad and who are able to benefit directly from it: the immigrants who establish firms in the host country and who, for the sake of simplicity, we refer to as *immigrant entrepreneurs*.[4] This intuition was already spelled out in Gould (1994), who mentions a study by Min (1990) that found the most frequent occupation of Korean immigrant entrepreneurs in the Los Angeles area was trading activities with Korea. However, this evidence was largely disregarded in subsequent contributions.[5]

In the current paper, we further elaborate on this idea. 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-entrepreneurial migrants can do—immigrant entrepreneurs are also able to directly exploit advantages deriving from knowledge of their own country, which is superior to that of domestic entrepreneurs, to directly *sell to their home countries the goods they produce where they settle down*.

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.[6] 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,[7] and the rate of self-employment of immigrants tends to be higher than that of natives, or at any rate, substantial and growing (OECD, 2010).

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, this work 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.

Using administrative data on immigrants located 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, through an instrumental variables (IV) strategy, the causal relationship going from diasporas and immigrant entrepreneurs towards export flows.[9] The endogeneity of the diaspora and immigrant entrepreneurs is addressed by using both the popular instrument based on immigrant *enclaves* (Card, 2001) and instruments built using auxiliary gravity models.[10] Both variables, the size of the diaspora and the number of immigrant entrepreneurs, have a positive, significant and economically meaningful effect on exports. In particular, we find that increasing the number of non-entrepreneurial 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. Increasing immigrant entrepreneurship by 10% (i.e. a bit less than one entrepreneur, or precisely 0.84, at the sample mean) would raise exports by about 0.6%. Our estimates imply that holding constant the total number of immigrants, the expected pro-trade effect of a migrant becoming an entrepreneur amounts to an increase of US$ 5,946 at the median export flow.

We carefully discuss the validity of the 'past immigrant settlement' instrument in our specific context (in Appendix C), in light of the two most common criticisms to shift-share instruments, which are related to the presence of long-run trends affecting the location of immigrants[11] (McCaig, 2011) or the strong serial correlation in the composition of immigrant inflows by nationality (Jaeger *et al.*, 2018).

In the sensitivity analysis (in Appendix C), we also check the robustness of our estimates to a number of other potential threats, such as the omission of unobservable variables acting at the trade-pair level, or the whole effect being accounted for by Chinese immigrants, who exhibit very high rates of entrepreneurship. Importantly, we also assess the sensitivity of the estimates to the stock of inward and outward bilateral foreign direct investment (FDI), which, to the best of our knowledge, has never been done in the past.

Besides these dyadic 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 effects of the gravity model) by 2.4%.

Our paper makes three main contributions to the existing literature. First, it introduces and tests a new channel through which immigrants may increase export flows, namely by becoming entrepreneurs and setting up firms that export abroad. Since the populations of immigrant entrepreneurs and diasporas are positively correlated, it becomes crucial to isolate the effects of each, to gain a better understanding of what really drives the immigrants-trade link. In this respect, our analysis demonstrates that the effects estimated in previous papers are not fully accounted for by immigrant entrepreneurs. Indeed, both non-entrepreneurial immigrants and immigrant entrepreneurs contribute to increasing a province's exports towards their countries of origin (dyadic effects mainly relating to the informational channel). Second, unlike other papers that have only focused on dyadic effects (see, for instance, Bratti *et al.*, 2014), our paper also uncovers a generalized positive effect of immigrant entrepreneurs on export flows towards all destinations. In line with our conceptual framework, this last effect might be produced by an increase in firm productivity, which overcomes the cut-off necessary to bear the fixed costs of exporting. Last, our paper is informative for the on-going debate on countries' migration policies. Indeed, many countries are switching to more restrictive immigration policies, both towards forced (e.g. refugees) and economic immigrants, and we may expect these countries to benefit less in the future from the economic gains from immigration, which have been documented in the economic literature. Our paper is only concerned with the trade (namely export) losses and shows that reducing immigration is likely to have negative effects on exports, but that even with a constant population of immigrants some of these negative effects could be at least partly countervailed by stimulating immigrant entrepreneurship.

The structure of the paper is as follows. The next section sets the stage for the analysis, reviewing the 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 used. Section 5 describes the empirical strategy and identification. In Section 6, we comment on the main findings. A discussion of the most common criticism to the 'past immigrant settlement' instrument and some robustness checks are reported in the appendix. Section 7 investigates some 'aggregate' (non-dyadic) effects of the stock of immigrants and immigrant entrepreneurs on trade at the province level. Section 8 concludes, summarizing the main findings of the paper. Finally, the Appendices available online, include information on the countries listed in Table 1 of the main text (Appendix A); on heterogeneous effects by quartiles of non-entrepreneur immigrants, entrepreneur immigrants and export values (Appendix B); robustness checks (Appendix C) and weighted regressions (Appendix D).

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 Card (2001) and motivated by the concept of immigrant enclaves (Bartel, 1989),[12] or by exploiting quasi-experiments provided by migrant or refugee dispersion policies (e.g. Cohen *et al.*, 2017, Steingress, 2018, Parsons and Vézina, 2018).

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 are able to provide smoking-gun evidence that this is really the explanation, nor do they explore whether and how immigrants' knowledge is transferred to natives, in particular native entrepreneurs, for whom this knowledge flow should be primarily relevant 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 Lodefalk (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. Similar evidence of a positive effect of foreign employees on trade is found for Denmark (Hiller, 2013), France (Marchal and Nedoncelle, 2016, Mitaritonna *et al.*, 2017) and Germany (Andrews *et al.*, 2017), especially for skilled or senior workers.

A separate and growing literature on immigrant entrepreneurs focuses on the differences between foreign-born and native entrepreneurs and among foreign-born entrepreneurs.[13] 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 or 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 terms of a general set-up of international trade involving heterogeneous firms (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 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, to 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 constant elasticity of substitution (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 CES 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 in turn depend on wages (w_{it}), trade costs (τ_{ijt}) and the average productivity of exporting firms ($\bar{\phi}_{ijt}$).

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 zero effect on average wages (including both natives and immigrants) through this channel 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 an effect, if any, to act on X_{ijt} but also on X_{it}, boosting provincial exports towards *all* destinations. By contrast, an increase in the number of immigrant entrepreneurs may increase the local demand for labour, countervailing the negative effect of increased labour supply on wages, and reduce firms' export performance. 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 firm productivity.[14] 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 more 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 employing immigrants from country j would operate for the specific trading pair ij, 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 destination-specific productivity thresholds[15] 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 native firms' marginal costs of exporting 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 a 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 and that is transferred to native entrepreneurs. 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 the number of exporting firms n_{ijt} and X_{ijt} in equation (1).

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

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} (i.e. the number of firms in province i exporting to country j, the trade cots τ_{ijt} and the ratio between the average productivity of firms in i serving market j and the average productivity computed over all exporters in province i in each period t, respectively) 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,[16] 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-entrepreneurial) 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. $Border_{ij}$ is a border dummy that is included to take into account possible non-linearities in the effect of distance.[17] The province-year fixed effects $\boldsymbol{\delta}_{it}$ 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 effects 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 $\boldsymbol{\theta}_{jt}$ 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 non-entrepreneurial 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[18] and foreign-born entrepreneurs in manufacturing for the period 2002-2011.[19] 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*).[20] 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.[21] For individual firms, we can associate firm ownership with a unique person and nationality (i.e. for every province in Italy, we grouped firms according to the country of origin of the entrepreneur), and such data can be used to analyse 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.

[Figure 1 about here]

[Figure 2 about here]

Table 1 reports the evolution of the foreign presence in terms of both residents and individual entrepreneurs for the 20 most represented nationalities in 2011 (i.e. the final year of our analysis). Concerning the distribution by country 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.[22] As expected, the evolution of the foreign residents and immigrant entrepreneurs time series are strongly 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.

[Table 1 about here]

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 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 Card (2001) in his study of the effect of immigrants on the labour market outcomes of natives, which has been widely used in trade-migration studies (see, for instance, Peri and Requena-Silvente, 2010, Bratti *et al.*, 2014).[23] 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 origin 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 on the host country's labour market, since it might be much easier for migrants to observe (or predict) the 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; therefore, we resort to an IV estimation strategy.

More specifically, we use the distribution of immigrants' requests for residence permits[24] (*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.[25] Let us define D_{ijt} as the diaspora (i.e. the number of non-entrepreneurial immigrants) from country j located in province i at time t and D_{jt} as 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:

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

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:

$$\widehat{D}_{ijt}^{\text{Card}} = wh_{ij95}D_{j0} + wh_{ij95}\sum_{q=0}^{t}F_{jq} = wh_{ij95}(D_{j0} + \sum_{q=0}^{t}F_{jq}) = wh_{ij95}D_{jt} \quad (6)$$

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). On the one hand, from equation (6) 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.[26] On the other hand, it is often the case that when using yearly data, there is not enough within-trading pair (ij) year-to-year variation in the nationwide ethnic composition of immigrants to use the enclave instrument, which is based on the idea of a strong autocorrelation in migrants' settlement decisions (Jaeger *et al.*, 2018).

We use a similar idea to build an instrument for the stock of immigrant entrepreneurs ($\widehat{IE}_{ijt}^{\text{Card}}$). 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 immigrant entrepreneurs 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 behaviour, 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):[27]

$$D_{ijt} = \exp(\boldsymbol{\delta}_{it} + \boldsymbol{\theta}_{jt} + \gamma_1 \ln(1 + \widehat{D}^{\text{Card}}_{ijt}) + \gamma_2 \ln d_{ij} + \gamma_3 Border_{ij}) v_{ijt}$$

$$IE_{ijt} = \exp(\boldsymbol{\delta}_{it} + \boldsymbol{\theta}_{jt} + \phi_1 \ln(1 + \widehat{IE}^{\text{Card}}_{ijt}) + \phi_2 \ln d_{ij} + \phi_3 Border_{ij}) u_{ijt}$$

where the control variables have the same meanings as above. After estimating the two gravity equations, we compute the predicted values excluding *demand pull factors* (i.e. province-year fixed effects) from the linear predictor. Identification is based on one exclusion restriction for each equation (namely $\ln(1 + \widehat{D}^{\text{Card}}_{ijt})$ and $\ln(1 + \widehat{IE}^{\text{Card}}_{ijt})$ for the non-entrepreneurial 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 shift-share instruments is zero whenever non-entrepreneurial immigrants or immigrant entrepreneurs from j were not present in province i in the base year, the additional regressors used in the PPML models can also predict non-zero stocks for those ethnic groups. Of course, this comes at a cost. Note that, given that the auxiliary regressions include the same controls as the second stage of two-stage least squares regression analysis (2SLS), for the nationalities that were not present in a province in the base year, identification is based only on the non-linearity of the PPML-predicted values in the covariates.[28]

As previously 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 settled earlier in Italy, then the shift-share instrument, which weights these communities more, may deliver higher IV estimates of the effect of immigrants than the PPML instrument.

6 RESULTS

6.1 DYDADIC 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-entrepreneurial immigrants (NE immigrants for brevity) simply as 'immigrants,' or *diasporas*. 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). 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 NE immigrants and immigrant entrepreneurs are 0.096 and 0.086, respectively, in both cases statistically significant at the 1% level.[29] According to these results, the elasticity of exports to immigrant entrepreneurs is similar in magnitude to that of exports to non-entrepreneurial 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).

The estimated elasticities remain positive, statistically significant at the 1% level and of a similar order of magnitude: 0.069 for NE immigrants and 0.054 for immigrant entrepreneurs. 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.[30] 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.[31]

[Table 3 about here]

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 NE immigrants (D) would increase exports by roughly US$ 652 (keeping immigrant entrepreneurs constant),[32] whereas a 1% inflow of immigrant entrepreneurs (IE) would increase exports by US$ 510 (keeping 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$ 5,769.[33]

Although our baseline specification omits trading pair fixed effects because the simultaneous inclusion of these fixed effects and country-year fixed effects would completely absorb the variation in the *enclave* instrument in the IV estimates (see the previous section), it may still be important to assess how the estimated effects change when ij fixed effects are included in the OLS models. Indeed, as stressed by Baier and Bergstrand (2007) in the context of the impact evaluation of trade agreements on bilateral trade, controlling for these fixed effects may help attenuate potential endogeneity concerns. For this reason, Table B1 in Appendix B reports the OLS estimates of the models including trading pair and province-year fixed effects. Column (1) reports our baseline estimates of Table 3. Column (2) reports the estimates of the model including ij and it fixed effects, while in column (3) we also add interactions between world macro regions and year fixed effects.[34] In both columns (2) and (3), the estimated coefficients are smaller than those in column (1), but still economically significant and statistically significant at the 1% level. The elasticities are 0.060 for non-entrepreneur immigrants and 0.053 for immigrant entrepreneurs in column (2), and 0.058 for NE immigrants and 0.044 for immigrant entrepreneurs in column (3). The smaller magnitude of the coefficients compared with Table 3 can be explained by the fact that these estimates only exploit within-trading pair variation over time and not the large variation existing across Italian provinces in the stocks of immigrants coming from each origin country.

[Table 4 about here]

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 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 NE 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, 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 and where 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 especially where there is already a large 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. As for the other covariates included in the first stages, it is worth noting the opposite signs for distance on the stock of NE immigrants and immigrant entrepreneurs, negative and positive, respectively. While the first effect is fully consistent with a gravity model in which immigrant inflows are negatively affected by distance, the second effect could be explained by discrimination in paid employment (Clark and Drinkwater, 2000): immigrants who are geographically and culturally distant from the natives might suffer more employment discrimination and choose to become self-employed or start their own business.[35] Finally, columns (4) and (5) show the first stages using the instruments built with the auxiliary gravity PPML models for NE immigrants and immigrant entrepreneurs. Again, the instruments are highly statistically significant. The PPML instruments seem to do a better job of separating the effects of NE immigrants and immigrant entrepreneurs in each first stage. Both the coefficient on the predicted stock of NE immigrants in the immigrants equation and the coefficient of the predicted stock of foreign entrepreneurs in the entrepreneurs equation increase.

[Table 5 about here]

Table 5 reports the second-stage of 2SLS. Column (1), pooling NE 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 NE 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 particularly in less competitive provinces (e.g. less exposed to native firms' competition) which also tend to export less. Column (2) reports the second stage of the specification, which includes diaspora and immigrant entrepreneurs separately; their estimated elasticities are 0.157 and 0.062, respectively. The 2SLS effect of immigrant entrepreneurs is slightly lower than that found with OLS, while the 2SLS estimate for diaspora is larger. These results may suggest that diaspora is negatively selected while entrepreneur immigrants are positively selected.

Another related explanation is that, in the presence of heterogeneous effects, the IV estimates refer to the 'compliers' with the enclave instruments (i.e. Local Average Treatment Effects, LATE), and that these local effects are larger (smaller) than those of a random immigrant (immigrant entrepreneur) in the population. Unfortunately, we do not have individual data on the characteristics of entrepreneur and NE immigrants, and we cannot use them to assess positive or negative selection, or to characterize the compliers. One last explanation is measurement error. Indeed, the stock of immigrants is measured through municipalities' population registers, but if immigrants are very geographically mobile (Schündeln, 2014), these stocks may be affected by a substantial measurement error. Immigrants may, for instance, be registered in different municipalities from those in which they work or currently live (however the OLS attenuation bias for NE immigrants would imply that the instrument is less affected by measurement error).

The same is less likely to occur for ethnic entrepreneurs, who are more geographically tied to the location of their business.

Finally, column (3) shows the second stage of the estimates using the instruments based on the auxiliary PPML regressions,[36] where the elasticity of exports to NE 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 an increase of US$ 1,644 at the median export flow, whereas the estimated effect of a 1% increase in immigrant entrepreneurs shrinks this to US$ 557. The lower elasticity of exports to immigrant entrepreneurs compared to NE immigrants in reality hides a much larger effect of the former when variations are computed in 'heads,' because the average number of immigrant entrepreneurs per origin country/province is much lower. Indeed, our estimates imply that the *ceteris paribus* expected pro-trade effect of a migrant becoming an entrepreneur (holding the total number of immigrants $D + IE$ constant) amounts to a US$ 5,946 increase computed at median export flow.

It is worth noting that by excluding trading pair fixed effects but including it and jt fixed effects, identification in our IV estimates mostly stems from the different location choices of immigrants from a given country across Italian provinces, i.e. cross-sectional variation. Our identification strategy resembles that used in Card (2001), who studies the effect of immigrants on wages by exploiting variation in their presence between occupations at the city level, whereas we exploit variation in the stock of immigrants by country of origin at the province level. The main threat to identification is therefore that the lagged location choice of immigrants, which is one key component of the 'enclave' instrument, is not 'as good as randomly assigned' with respect to the trade potential of trading pairs. We seek to address this concern in two ways: 1) by considering location in a year sufficiently back in time (in the estimates in the main text); and 2) by including lagged export in a base year in the estimated equation. The latter model seeks to only exploit variation in the predicted stock of immigrants that is orthogonal to the export potential of a given trading pair ij, as captured by the level of export of this pair in a base year, which might be known to the older waves of immigrants when settling in the host country. The estimates reported in Table C1 of Appendix C show that the elasticities are only marginally affected. In Appendix C, we also discuss some weaknesses of the shift-share (or *enclave*) instrument stressed in the recent literature.

7 AGGREGATE PROVINCE-LEVEL EFFECTS

We conclude our analysis by attempting to shed light on other potential effects of diasporas and immigrant entrepreneurs that could affect the exports 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 Section 3. We attempt this by regressing the province-year fixed effects estimated using 2SLS ($\widehat{\delta_{it}}$) on the populations of NE immigrants and immigrant entrepreneurs, immigrant diversity and some additional controls. Namely, we estimate the following equation:

$$\begin{aligned}\widehat{\delta_{it}} =& \beta_1 \ln(D_{it}) + \beta_2 \ln(IE_{it}) + \beta_3 \ln(TFP_{i2002}) + \beta_4 \ln(MNE_{i2002}) \\ &+ \beta_5 \ln(DIV^N_{i2002}) + City_{i2002} + \gamma_t + \varepsilon_{it}\end{aligned}$$

where i is the subscript for Italian provinces (NUTS-3) and t stands for time. D_{it} and IE_{it} are the total stocks of NE immigrants and immigrant entrepreneurs located in province i at time t, respectively. In light of equation (4), the dependent variable can be interpreted as an estimate of the logarithm of province competitiveness (i.e. $\ln(w_{it}/\bar{\phi_{it}})^{1-\sigma}$), and we interpret the coefficients on the regressors measured in logarithm as elasticities. Since 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 to account for different initial competitiveness levels of Italian provinces. Namely, we include total factor productivity (TFP) of manufacturing firms,[37] the proportion of foreign multinational enterprises (MNEs),[38] a measure of the diversity of the local labour force (DIV)[39] and a dummy variable equal to one for provinces above the 75th percentile of population ($City$).[40] All of these control variables are evaluated in the first available year (i.e. 2002) to make them predetermined with respect to the estimation period.[41] Thus, we limit the estimation to the period 2003-2011. This specification is preferred to the one including year and province fixed effects, which explains 98% of the variance in the dependent variable. This is not surprising given that the dependent variable is a time-varying measure of a province's competitiveness and is very persistent over time. This implies that, like in our dyadic estimates, by excluding province fixed effects, the effects of interest are mostly identified by cross-sectional variation. Our specification would be equivalent to modeling the province fixed effects as

$u_i = \pi_1 \ln(TFP_{i2002}) + \pi_2 \ln(MNE_{i2002}) + \pi_3 \ln(DIV^N_{i2002}) + City_{i2002} + \upsilon_i$, where υ_i is a white-noise error term, uncorrelated with the potentially endogenous variables.[42] Finally, γ_t is a set of year fixed effects (one of which acts as the regression intercept), whereas ε_{it} is an error term clustered at the province level to account for serial correlation in trade.

TFP measured at the beginning of the estimation period is a proxy for the initial level of a province's productivity, which may affect both NE immigrants' and immigrant entrepreneurs' location choices and simultaneously increase a province's export performance. The controls MNEs and *City* are included for the same reason. In light of the extensive literature on the productivity- and export-enhancing effects of labour force diversity (see, among others, Parrotta *et al.*, 2016, Bombardini *et al.*, 2012), $\ln(DIV^N_{i2002})$ is included in the regression to assess whether the gains in export performance are mainly to be attributed to the initial level of population diversity rather than to the sizes of the immigrant and immigrant entrepreneur populations.[43] Since the fixed effect that represents our dependent variables has been estimated from a regression, observations are weighted by the inverse of the standard errors of the $\widehat{\delta_{it}}$ using weighted least square (WLS) estimates.[44].

[Table 6 about here]

Results are reported in Table 6. The WLS estimates in columns (1)-(5) show that the positive association between NE immigrants and the dependent variable, which we consider a proxy of provinces' overall competitiveness, is reduced in size and loses statistical significance when the controls are progressively included in the regression. A large 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.[45] 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.

8 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 export their products abroad. Indeed, immigrants help natives 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.

The current paper adds a complementary hypothesis: the trade-creating effect of immigrant entrepreneurs. After outlining various reasons why immigrant entrepreneurs can spur export activities from the regions in which they settle towards their country of origin, we test this hypothesis using provincial manufacturing data from Italy.

Using a 10-year panel dataset on Italian provinces (i.e. NUTS-3 regions) and gathering administrative data on exports in manufactures, the diaspora (i.e. non-entrepreneurial immigrants) and the population of entrepreneurs by country of origin, we estimate a set of augmented gravity models for exports. The potential endogeneity of both non-entrepreneurial immigrants and immigrant entrepreneurs is addressed using an instrumental variables estimator based on a shift-share instrument à la Card (2001). To overcome some of the potential weaknesses of such an instrument, which does not use new immigrant or entrepreneur nationalities that were not present in the host country in the base year for identification, we also use a second instrumental variable based on auxiliary PPML gravity models for non-entrepreneurial migrants and immigrant entrepreneurs that overcomes this limitation. The IV estimates obtained with both strategies give very similar results and point to large causal effects of both non-entrepreneurial immigrants and immigrant entrepreneurs on exports. The estimated effect of a 1% increase in the population of non-entrepreneurial immigrants corresponds to a US$ 1,644 rise in yearly manufacturing exports from the host province towards the immigrants' country of origin, 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-entrepreneurial) migrant into an entrepreneur would raise yearly province exports in manufacturing by US$ 5,946 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, the fact that the whole effect may be driven by larger and more entrepreneurial immigrant communities (e.g. Chinese communities), or controlling for 'initial conditions' that may affect non-entrepreneurial immigrants' and immigrant entrepreneurs' locations.

Finally, we propose that diasporas and immigrant entrepreneurs may positively impact 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, the stock of immigrant 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%.

The policy implications of our findings are straightforward: the presence of immigrant diasporas allows provinces to reach out to international markets more easily, and the effect is magnified when immigrants become entrepreneurs. Provinces interested in increasing their export competitiveness must implement policies that make the establishment of immigrant-led firms easier. Linguistic, financial or legal barriers to immigrant entrepreneurship result in lower competitiveness and lost exports. An effective *startup visa* policy (e.g. allowing immigration conditional on starting a business and after satisfying specific financial or other requirements) may foster the presence of footloose foreign entrepreneurs. On the other hand, local policies to encourage entrepreneurship among immigrants, reducing 'red tape' barriers, could complement more traditional export promotion policies.

In spite of the results obtained in the present study, several issues still remain open to future research. The possibility of tracing the mechanisms linking the presence of diasporas and of immigrant entrepreneurs to bilateral exports and to competitiveness crucially depends on moving the analysis from aggregate data to employer–employee firm-level data. Only that level of disaggregation could make possible to disentangle the effect of immigrant workers from that of immigrant entrepreneurs on exports (and on the performance of firms in general), entering the black box of *business and network effects*. Such data could allow a progression beyond the effect on exports in manufacturing and enable the investigation of the effect of immigrants and immigrant entrepreneurs in the service sector as well. The role of migration in trade in services is indeed an under-explored issue that is growing in relevance (see Ottaviano *et al.*, 2018) and requires further research to corroborate the scant extant evidence. Thus, the availability of firm-level data could propel research on the pro-trade effect of immigrants to a much-needed new stage of inquiry.

Acknowledgements. We thank three anonymous reviewers, the journal's co-editor Prof. Steven Brakman, participants in ITSG (Italian Trade Study Group) in Cagliari, DEGIT (Dynamics, Economic Growth and International Trade) in Nottingham, SIE (Italian Economic Association) annual conference in Milan, ETSG (European Trade Study Group) in Helsinki, the 8th International Conference "Economics of Global Interactions: New Perspectives on Trade, Factor Mobility and Development" in Bari, and those at seminars and workshops at Cepii-Paris, Roma Sapienza, and the University of Milan, and especially Giuseppe de Arcangelis, Giovanni Bruno, Nicola Coniglio, Rosario Crinò, David Dorn, Tommaso Frattini, Marco Leonardi, Marina Murat, Elena Meschi, Gianmarco I.P. Ottaviano and Anna Rosso for insightful comments and suggestions. We also thank Rosalia Alessi for her generous help with the data. The scientific output reported and views expressed in this paper do not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. The usual disclaimer applies.

Notes

[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] To explain the trade-creating effect of migrants, the author also mentions in that paper that "the native language of the immigrants can become known, or used more often, by the host country residents" or that "the importance of these immigrant information effects, of course, would depend on the initial amount of foreign market information in the host country and the ability of immigrants to relay information and to integrate their communities into the host country" (Gould, 1994, p.303).

[4] 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.

[5] An exception is Cohen *et al.* (2017), which using a sample of US firms, shows that the ethnic composition of the local population is positively correlated with the ethnic composition of a firm's board of directors (including top management) and that 'more connected' firms trade significantly more with the countries to which they are 'connected'. Some recent papers focus instead on the role of skilled/unskilled migrants or on those employed in business-related occupations, but none focus on the role of immigrant entrepreneurs (see the following section).

[6] Unlike Aleksynska and Peri (2014), we do not use data on immigrants employed in business-related occupations (e.g. managers, salespersons), who may also work in non-tradable 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.

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

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

[9] Our analysis focuses only on exports in manufacturing. It would of course be interesting to assess the role of immigrants on exports in other sectors as well. However, the notorious problems in measuring international trade in services at the country level (Lipsey, 2009, Steingress, 2018) get exacerbated at the regional level. As an example, the Italian National Statistical Institute does not provide data on bilateral trade in services at the NUTS-3 level of disaggregation. The only paper we are aware of that deals with the role of immigration on trade in services is Ottaviano *et al.* (2018), which uses UK pre-crisis firm-level data. However, unlike in the UK, the bulk of immigrants in Italy work in manufacturing and non-tradable services (see, for instance, Mocetti and Porello, 2010).

[10] The enclave instrument was first introduced in Altonji and Card (1991), which uses the past settlement of immigrants irrespective of their nationality, while Card (2001) also exploits immigrant ethnic composition to build an instrument for the stock of immigrants.

[11] e.g. induced by serial correlation in local demand shocks.

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

[13] See Fairlie and Lofstrom (2014) for a comprehensive review of the literature and Kerr

and Kerr (2017) for recent evidence from the US.

[14] 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.

[15] In models with heterogeneous firms, such as Melitz (2003), Arkolakis *et al.* (2012), Hsieh *et al.* (2016), in order to export firms have to bear a fixed cost, hence only those with productivity above a given cutoff will manage to export.

[16] The Anderson and van Wincoop (2003) specification of the gravity equation can be derived from micro-foundations and 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.

[17]Equation (5) is obtained from (4) by assuming that $\ln n_{ijt} + (1-\sigma)\ln \tau_{ijt} - (1-\sigma)\ln k_{ijt} = \alpha_1 \ln(1 + D_{ijt}) + \alpha_2 \ln(1 + IE_{ijt}) + \alpha_3 \ln d_{ij} + \alpha_4 Border_{ij}$. For a similar constant-elasticity specification (but only for immigrants and not distinguishing between entrepreneurs and non-entrepreneurs), see Combes *et al.* (2005).

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

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

[20] Provincial offices of *Infocamere* are in charge of producing and maintaining the registry of all firms active in their territory.

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

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

[23] Some recent studies exploit quasi-natural experiments provided by the (presumably) random allocation of refugees across US states (Steingress, 2018, Parsons and Vézina, 2018) or World War II internment camps (Cohen *et al.*, 2017). Even if considering specific episodes of migration generally allows for a convincing identification, refugees are only a fraction of total immigrants, are probably 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 generalisable to all kinds of immigrants.

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

[25] 1995 is the earliest year for which data on residence permits by province and country of origin are publicly available and for which the structure of Italian provinces was similar to that of the estimation period (2002-2011).

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

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

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

[29] Although the constant elasticity specification implied by the log-log form is usually adopted in trade empirical models mainly on theoretical grounds, there might be some concerns that the effects of immigrant entrepreneurs and NE immigrants are not constant

along the distribution of these independent variables or of export values. For instance, doubling the number of immigrants may have different effects on exports if there is a low initial number of immigrants, a high initial number of immigrants, low initial exports or high initial exports. For this reason, we estimated models in which the independent variables of interest were interacted with quartile dummies (of D, IE and X). The results are reported in Appendix B.1. The estimates show a lower effect of D in the first quartile (of D) and of IE in the fourth quartile (of IE), although the confidence intervals generally overlap across quartiles and with the pooled estimates obtained with the constant elasticity form (Figure B1 in Appendix B.1). The estimated effects of D tend to decrease by quartile of export values while the effect of IE is constant across quartiles of exports (Figure B2 in Appendix B.1).

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

[31] 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.

[32] 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 of the following back-of-the-envelope computations are done in the same way.

[33] Note that one migrant represents 7% of the median stock of (non-entrepreneurial) 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$ 5,769 corresponds to an increase of 0.61% for the median export flow.

[34] The latter specification is estimated to control for trends in migration and trade common to all countries belonging to the same macro-region. The simultaneous inclusion of ij, it and jt fixed effects, instead, absorbs most variation in the stock of immigrants and immigrant entrepreneurs, and the estimated OLS effects are small and not statistically significant. This is also consistent with the findings in Aydemir and Borjas (2011). In particular, since longitudinal estimates rely on immigrant shares computed in small samples (small populations in our case) at the province-origin country level, there might be a large measurement error in these variables, and controlling for a large set of fixed effects (ij, it and jt) may leave too little 'true' variation in the variables of interest and cause an important attenuation bias.

[35] Indeed, the stock of immigrant entrepreneurs is composed of two types of individuals: 1) individuals who were already entrepreneurs in their home countries; and 2) individuals who decided to become entrepreneurs in the host country. However, an alternative interpretation may also be possible. If only the ablest and most productive (or less risk-averse) immigrants become entrepreneurs, the positive effect of distance on the number of immigrant entrepreneurs may suggest that immigrants who move despite the higher distance-related costs (higher risk of a longer journey) are more positively selected.

[36] The results of these regressions are reported in Table B2 in Appendix B.

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

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

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

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

[41] Since diaspora and immigrant entrepreneurs during the estimation period may have an effect on them.

[42] We also estimated a specification including province fixed effects but excluding year fixed effects, exploiting therefore only time variation, and obtained similar, albeit less precise, estimates. The estimated coefficients are equal to 0.051 (s.e. 0.046) for $\ln(1+D_{it})$ and 0.162 (s.e. 0.091) for $\ln(1+IE_{it})$, the latter is statistically significant at 10%.

[43] The included regressors explain 85% of the variation in ln(1+ D_{ijt}) and 64% in ln(1+ IE_{ijt}).

[44] Results are fully robust to the use of feasible generalized least squares (FGLS) (see Table D1 in Appendix D.)

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

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.

Andrews M., Schank T. and Upward R. (2017) Do foreign workers reduce trade barriers? Microeconomic evidence, *World Economy*, 40, 9, 1750–1774.

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

Aydemir A. and Borjas G.J. (2011) Attenuation bias in measuring the wage impact of immigration, *Journal of Labor Economics*, 29, 1, 69–112.

Baier S.L. and Bergstrand J.H. (2007) Do free trade agreements actually increase members' international trade?, *Journal of International Economics*, 71, 1, 72–95.

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.

Bartel A.P. (1989) Where do the new U.S. immigrants live?, *Journal of Labor Economics*, 7, 4, 371–91.

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, *World Economy*, 39, 4, 496–512.

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

Bombardini M., Gallipoli G. and Pupato G. (2012) Skill Dispersion and Trade Flows, *American Economic Review*, 102, 5, 2327–2348.

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.

Cohen L., Gurun U.G. and Malloy C. (2017) Resident networks and corporate connections: Evidence from World War II internment camps, *Journal of Finance*, 72, 1, 207–248.

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.

Jaeger D.A., Ruist J. and Stuhler J. (2018) Shift-share instruments and the impact of immigration, NBER Working Papers 24285, National Bureau of Economic Research, Inc.

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.

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.

Lipsey R.E. (2009) Measuring international trade in services, in: *International Trade in Services and Intangibles in the Era of Globalization*, Reinsdorf M. and Slaughter M.J., eds., University of Chicago Press, chapter 1, 27–60.

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

McCaig B. (2011) Exporting out of poverty: Provincial poverty in Vietnam and U.S. market access, *Journal of International Economics*, 85, 1, 102–113.

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

Min P.G. (1990) Korean immigrants in Los Angeles, Institute for social science research, working paper series, Institute for Social Science Research, UCLA.

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

Mocetti S. and Porello C. (2010) How does immigration affect native internal mobility? New evidence from Italy, *Regional Science and Urban Economics*, 40, 6, 427–439.

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

Ottaviano G.I.P., Peri G. and Wright G.C. (2018) Immigration, trade and productivity in services: Evidence from UK firms, *Journal of International Economics*, 112, 88–108.

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.

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

Parsons C. and Vézina P.L. (2018) Migrant networks and trade: The Vietnamese boat people as a natural experiment, *The Economic Journal*, 128, 612, 210–234.

Parsons C. and Winters A.L. (2014) International migration, trade and aid: A survey, in: *International Handbook on Migration and Economic Development*, Edward Elgar Publishing, chapter 4, 65–112.

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.

Schündeln M. (2014) Are immigrants more mobile than natives? Evidence from Germany, *Journal of Regional Science*, 54, 1, 70–95.

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. (2018) The causal impact of migration on US trade: Evidence from political refugees, *Canadian Journal of Economics*, 51, 4, 1312–1338.

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*, 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.

Table 2: Descriptive statistics

Variable	$N^{>0}$	$Zeros^{\%}$	Mean	p25	p50	p75
$Exports^{Mln}_{US\$}$ (X)	141,091	29.8	27.86	0.09	0.95	8.82
$diaspora$ (D)	119,729	40.4	260.58	3	14	74
$Immigrant\ entrepreneurs$ (IE)	32,889	83.6	8.49	1	2	5
$distance^{Km}$ (d)	200,850		6,068.71	2,576.56	5,125.04	8,484.74

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

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	14,733	47.33	12.35	192,867	4.87	7.60	76.39
ROU	1,873	6.02	30.63	815,197	20.60	19.43	2.30
CHE	1,499	4.82	-2.09	7,720	0.20	-1.32	194.17
MAR	1,418	4.56	7.32	397,448	10.05	9.34	3.57
DEU	1,119	3.59	0.70	34,075	0.86	-1.40	32.84
ALB	1,076	3.46	8.53	440,365	11.13	11.29	2.44
TUN	682	2.19	4.13	81,611	2.06	3.10	8.36
FRA	646	2.08	-0.18	23,366	0.59	-3.64	27.65
ARG	522	1.68	-1.98	7,647	0.19	-2.26	68.26
SRB	515	1.65	5.24	85,306	2.16	-0.52	6.04
EGY	440	1.41	8.61	64,751	1.64	4.38	6.80
SEN	428	1.38	7.91	71,688	1.81	3.46	5.97
BGD	403	1.29	16.64	79,722	2.01	9.64	5.06
PAK	327	1.05	13.47	67,206	1.70	8.86	4.87
VEN	308	0.99	4.10	4,645	0.12	-1.27	66.31
BRA	283	0.91	8.30	36,484	0.92	1.12	7.76
BEL	268	0.86	0.40	4,708	0.12	-3.20	56.92
GBR	228	0.73	1.99	22,105	0.56	-2.37	10.31
NGA	221	0.71	9.97	47,725	1.21	3.20	4.63
UKR	200	0.64	53.88	175,383	4.43	17.78	1.14
Top 20	27,189	87.35	9.50	2,660,019	67.23	4.16	10.2
Total	31,127	100.00	10.27	3,956,454	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 for immigrant entrepreneurs and the diaspora, respectively. IE/D_{000} represents the number of immigrant entrepreneurs per 1,000 foreign residents. The description of ISO 3 country codes is reported in Appendix A.

Table 3: Baseline estimates, sample 2002-2011

	OLS		PPML		
Dependent variable:	$\ln(X_{ijt})$		$X_{ijt} \geq 0$		$X_{ijt} > 0$
	(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	200,850	200,850	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. d is distance. Estimates in columns (1) and (2) only include observations with $X > 0$. In all regressions, standard errors in parentheses are clustered at the province-by-country and province-by-year level. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively. i, j and t are province, country and time subscripts. The total number of observations is 200,815 (province $\times$ origin countries $\times$ years). Excluding zero trade flows reduces the estimation sample to 141,091 observations.

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

Dependent variable:	$\ln(1 + D_{ijt} + IE_{ijt})$ (1)	$\ln(1 + D_{ijt})$ (2)	$\ln(1 + IE_{ijt})$ (3)	$\ln(1 + D_{ijt})$ (4)	$\ln(1 + IE_{ijt})$ (5)
$\ln(1 + \widehat{D}_{ijt}^{\text{Card}} + \widehat{IE}_{ijt}^{Card})$	0.500***				
	(0.009)				
$\ln(1 + \widehat{D}_{ijt}^{\text{Card}})$		0.484***	0.047***		
		(0.010)	(0.003)		
$\ln(1 + \widehat{IE}_{ijt}^{\text{Card}})$		0.131***	0.671***		
		(0.015)	(0.012)		
$\ln(1 + \widehat{D}_{ijt}^{\text{PPML}})$				0.772***	0.079***
				(0.013)	(0.004)
$\ln(1 + \widehat{IE}_{ijt}^{\text{PPML}})$				0.063***	0.956***
				(0.022)	(0.013)
$\ln(d_{ij})$	-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	2,916	1,159		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. d is distance. i, j and t are province, country and time subscripts. Standard errors in parentheses are double clustered at the province-by-country and province-by-year level. The F-statistic refers to the excluded instruments. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

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

	(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 ($D + IE$)	Card (D and IE)	PPML (D and IE)

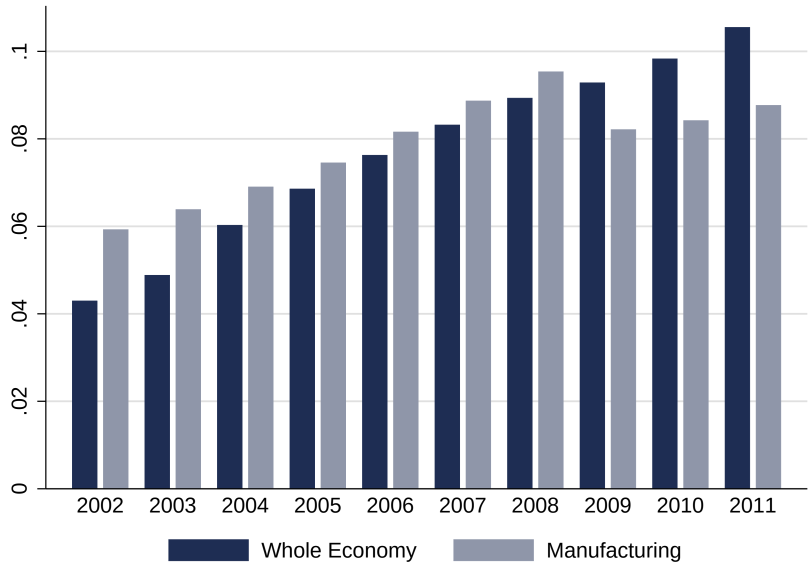
Note. The dependent variable is $\ln(X_{ijt} > 0)$, where X are export flows. D refers to foreign residents, while IE refers to foreign-owned individual firms. d is distance. i, j and t are province, country and time subscripts. Standard errors in parentheses are double clustered at the province-by-country and province-by-year level. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Table 6: Aggregate analysis: main results

Dependent variable:	Estimated Province-by-Year Fixed Effects: $\hat{\delta}_{it}$					
	(1)	(2)	(3)	(4)	(5)	(6)
			WLS			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 (D and IE)

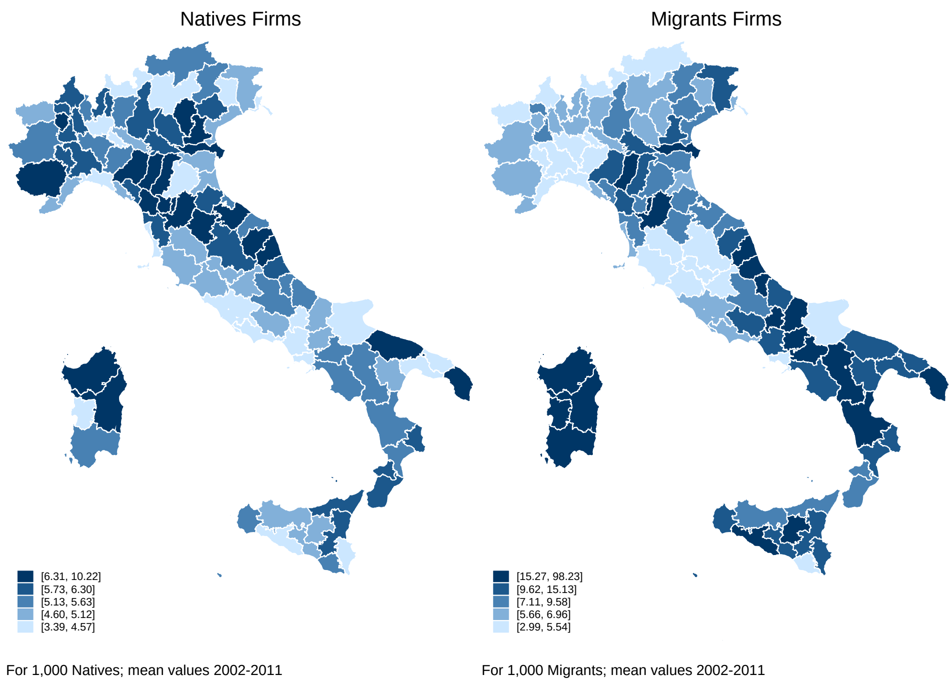
Note. The dependent variable is the province-year fixed effect $\hat{\delta}_{it}$ estimated in the (bilateral) export gravity equation. Total factor productivity (TFP) is estimated using the GMM approach (Wooldridge, 2009) on AIDA data (Bureau van Dijk). MNE represents the (log of the) number of multinational enterprises in a given province i at the beginning of the period (source: ICE-Reprint). DIV represents the log of the inverse of the Herfindahl–Hirschman index measuring diversity in the ethnic composition of residents. City$^{Pop>pc75}$ is a dummy variable equal to one for provinces above the 75th percentile of province population. 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. The F-statistic refers to the excluded instruments. Standard errors in parentheses are clustered at the province level. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

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



Note. The figure plots the proportion of immigrant entrepreneurs on the number of individual firms for the whole economy and the manufacturing sector.

Figure 2: Individual firms per 1,000 inhabitants



Note. The figure plots the incidence of individual firms against the total number of inhabitants, native (left panel) or immigrant (right panel), by province. Immigrant firm and native firm distributions correlate positively at the province-by-country level. When the log of foreign-owned individual firms is regressed on the log of native firms, we obtain 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.