Immigrant Networks, Trade Creation, and Trade Diversion

Tomáš Konečný[†]

CERGE-EI[‡]

Abstract

I derive a simple gravity model with matching and use data on the foreign-born population located in 19 OECD-member countries to estimate the impact of immigrant links on trade. The relative impact on trade of immigrant networks declines with the GDP of source country and is generally smaller than estimates from preceding studies. The theoretical framework explicitly models forgone bilateral trade and shifts in trade flows due to immigrant links. There is some empirical evidence that immigrant networks shift trade flows between countries. The net effect on total trade of a 10-percent increase in the overall immigrant stock varies between -0.12 and 1.18 percent for host countries and -6.99 and 2.58 percent for source countries in the sample.

Keywords: international trade, immigration, informal trade barriers
JEL classification: F22, O24

[†] I would like to thank CERGE-EI and the Czech Academy of Sciences for generous financial support during my stay at the University of California, San Diego (UCSD). Special acknowledgment goes to James Rauch, UCSD, who helped substantially to formulate the ideas in the paper. I would also like to thank Libor Dušek and Štěpán Jurajda for their invaluable comments. During the course of my work on this paper, I benefited from the feedback provided by Julian Betts, Peter Katuščák, Andreas Ortmann, Fusako Tsuchimoto, Robert Menkyna and Krešimir Žigić. The usual disclaimer applies.

[‡] CERGE-EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of Academy of Sciences of the Czech Republic.
Address: CERGE-EI, PO Box 882, Politických vězňů 7, 11121 Prague, Czech Republic.
E-mail: tomas.konecny@cerge-ei.cz.

1 Introduction

Informal trade barriers have become one of the central points in the debate launched by McCallum's "mystery of the missing trade" (McCallum, 1995) i.e., the finding that nations tend to trade too much intranationally and too little internationally. Particular attention has been directed towards insufficient information on available trading opportunities and imperfect contract enforcement. Insufficient information about foreign partners seems to be pronounced especially in more differentiated industries where product characteristics vary along multiple dimensions and price happens to be only one of several decision criteria. The resulting higher search costs can then make otherwise efficient cross-border matches unprofitable (Rauch and Trindade, 2003; Casella and Rauch, 2003). Similarly, in the absence of efficient contract enforcement when trade parties originate from different jurisdictions, potential contract renegation and losses accrued by the afflicted party decrease the incentives to engage in trade and, again, might prevent otherwise successful international matches (Greif, 1994).

Some social networks seem to be well equipped to deal with both kinds of informal trade barriers. These networks, often defined by common ethnicity or religion, can provide useful information and trade contacts to their members and/or employ some sort of collective punishment mechanism that could substitute for inadequate enforcement institutions. In particular, numerous studies on informal barriers examine the impact on trade of immigrant networks (e.g., Head and Ries, 1998; Gould, 1993; Girma and Yu, 2002). The results of these studies consistently support the notion that immigrant links indeed facilitate bilateral trade between host and source countries.

The present paper offers two extentions to the existing literature. First, it evaluates the potential role of immigrants in trade diversion, i.e., shifts in trade flows due to immigrant links to country of origin. Second, it derives a simple matching framework relating trade, immigrant links and the output of their country of origin, and calculates the GDP-adjusted estimates of the immigrants' overall impact on trade by host and source country.

The paper argues that in a situation when exporters decide between several competing

destination markets, the combination of pervasive informal trade barriers and country-specific knowledge possessed by immigrants might actually lead to a diversion of trade. Consider a German machinery producer who wishes to export to either Vietnam or Thailand. Other things being equal, if the informal trade barriers are uniform across both countries and trade is still profitable, the producer will be indifferent as to where to export. If, on the other hand, the producer is of Vietnamese ancestry, or perhaps employs Vietnamese officers in its trade department, the contacts and knowledge of local conditions might bias the export choice in favor of Vietnam. [1] Now assume such a decision has been made by a larger number of firms. While from the perspective of Germany the total exports do not change (or they increase somewhat if immigrants are more efficient in finding suitable matches), its bilateral trade with Thailand becomes lower than it would have been in the absence of immigrant networks. In this case, trade diversion from Thailand occurs due to a lost fraction of transactions that would have been realized by otherwise indifferent exporters.

A study on offshoring in the apparel industry (Gereffi, 1999) provides a related empirical observation; it describes the case of Taiwanese firms channeling large portions of their offshore investment into Malaysia and Thailand despite markedly lower wages in other parts of the region. A large part of both economies is, however, controlled by ethnic Chinese who maintain extensive social networks. The author argues that these networks shape many investment decisions.[2] Within the present context, the trade diversion would take the form of unrealized offshoring projects in countries such as Bangladesh or Sri Lanka, i.e., in destinations with very low wages but insufficient links to Chinese networks.

The following section reviews the existing empirical research on the role of immigrant links in international trade. Section 3 presents the empirical model and Section 4 discusses

[1] Heerander and Saavedra (2006) cite Peng's (1998) survey on the characteristics of trade intermediaries located in the U.S. According to this survey, 40 percent of U.S. intermediary officers or managers were foreign-born.

[2] Rauch and Trindade (2002) find that for trade between Southeast Asian countries with high population shares of ethnic Chinese, the smallest average portion of trade in differentiated products attributable to ethnic Chinese networks reaches nearly 60%.

the data employed. The following sections cover econometric issues results and sensitivity analysis. Section 7 concludes.

2 Evidence on trade and immigrant links

A number of country-specific studies exist that estimate the relationship between trade and immigrant links. For example, Gould (1993) analyzes migration inflows in the U.S. using panel data from 1970 to 1986 and predicts a 10-percent increase in immigrant stock to increase U.S. exports by 4.7 percent and U.S. imports by 8.3 percent. The U.S. trade data have been analyzed also by, e.g., Tadesse and White (2007, 2008a, 2008b, 2009); Bandyopadhyay, Coughlin and Wall (2008); Co, Euzent and Martin (2004). An exercise using Canadian data has been produced by Head and Ries (1998). The authors employ two different measures of immigrant links, namely the cumulative sum of immigrant inflows after 1970 and the imputed immigrant populations using census data, and report a 10-percent increase in the immigrant stock to raise Canadian bilateral exports by 1.0-1.3 percent and imports by 3.1-3.9 percent. The link between immigration and Canadian trade has also been studied by Helliwell (1997) and Wagner, Head and Ries (2003). Other more recent country studies include the U.K. (Girma and Yu, 2002); Denmark (White 2007a); Greece (Piperakis, 2003); New Zealand (Law and Bryant, 2005); Spain (Blanes, 2005); Rauch and Trindade (2002) used data on Chinese minorities in South-East Asia. These studies will form a useful benchmark for the trade creation estimates discussed in Section 5.

A number of studies focus on the characteristics of immigrants' country of origin influencing immigrant-driven trade. The Canadian study by Head and Ries (1998) finds that a trade contribution of more recent immigrant cohorts from East Asian and Latin American countries tends to exceed that of traditional migrant communities from within the European continent.[3] The U.S. study by White (2007b) divides source countries into four

[3] Recent shifts in the structure of immigrants' countries of origin for OECD member states have been documented in OECD (2004).

income groups and estimates the immigrant-link effect for each distinct group. His results indicate that immigrant networks from low income economies exert stronger influence on trade than their higher income counterparts. On the contrary, White (2007a) finds the opposite result for the Danish data. Of course, these contrasting results might be driven by a number of distinct channels that would ultimately correlate with the income level of a source country. Besides different immigration histories emphasized by Head and Ries (1998), trade activities of immigrant networks could select into a relatively small number of sectors within the source economy, so that their differential contribution to trade would partially reflect the source countries' sectoral dynamics. In that case, the less developed economies with a larger share of traditional sectors might observe correspondingly larger shares of immigrant-driven trade.[4] Even without the selective focus on a subset of industries, however, some networks might have limited capacity to exploit all available trade opportunities given the time, skill or logistic constraints, which would again translate into their lower relative contribution to trade. Despite the current inability to disentangle the individual mechanisms at work, White's estimates provide at least some idea on the actual magnitude of these effects.

Moving towards the potential trade-diverting role of immigrant networks, the concurrent studies by Dolman (2007) and Konečný (2007) use the OECD data from the Statistics Portal on Demography and Population and attempt to quantify the overall effect of immigration on trade. Dolman (2007) argues that while immigrant networks indeed seem to facilitate bilateral trade between host country and country of origin, they also reduce trade with other countries so that the overall effect is close to zero. The study by Konečný (2007) represents an earlier version of the present article and contains a detailed analysis of the two mechanisms within an explicit matching framework.

The research by Herander and Saavedra (2005) explores the spatial dimension of immigrant networks. Focusing on trade-creation effects of immigrant networks operating within

[4] Production of cultural goods could serve as one of examples. At a more general level, the role of cultural differences in immigrant-trade links has been investigated in a number of insightful studies by Tadesse and White (see Tadesse and White, 2007, 2008a, 2008b, 2009; White and Tadesse, 2007, 2008a, 2008b).

and between the U.S. states,[5] the results show a consistently stronger impact on U.S. state export volumes to a source country for local as compared to out-of-state populations. In particular, their results qualitatively conform to previous estimates in that a 10-percent increase in the local state immigration should on average increase the state's exports by 1.6 percent. The estimated impact of the out-of-state population, i.e. of the immigrant network geographic spillovers, then raises the states' export volumes by 0.7 percent only.

The present study aims to estimate a rather different dimension of network spillovers. While Herander and Saavedra (2005) deal with trade facilitating spillovers generated by immigrant networks of the *same* nationality located in different U.S. states, I instead focus on the relevance of potential trade-diverting spillovers by immigrant networks from *different* countries of origin across many host economies.

The following section presents the estimation framework.

3 Empirical model

For the empirical evaluation of the trade creation and diversion hypotheses I use a simple gravity framework that explicitly allows for matching in trade. The gravity relationship proportionally linking trade flows to the output of trading economies can be derived from a wide range of international trade models.[6] The present section, however, shows that the gravity relationship might be consistent even with a very simple world economy with matching and no differences in productivity, endowments or preferences across countries.

Assume the world population N is distributed across $J + I$ countries, where J are labelled host and I source economies that differ in size and in the structure of their population. Each agent regardless of location and status has linear preferences and is endowed with x units of indivisible input normalized to zero, which can be used either for local production or as an input into a joint venture with a foreign partner. Local production technology transforms the normalized input into 1 unit of output. Within the

[5] Another study focusing on trade-immigration link at the U.S. state level is Dunlevy (2006).

[6] Examples include Anderson (1979); Bergstrand (1990); Deardorf (1998); and Helpman and Krugman (1985).

joint venture, each of the participating parties has to invest their whole endowment to produce a) either $2a$ in case the venture has been formed between native agents from i and j, or b) $2b$ once both parties originate from i, yet one has the status of immigrant in host j. Different productivities between joint ventures. $\{a, b\} > 1$ are measures of match quality. The present model assumes that $\{a, b\} > 1$ is a result of the combination of host country and source country's specific knowledge. Distinct productivities of joint ventures reflect, e.g., different outside options of immigrants versus native agents in host country j. Furthermore, assume that differences between joint venture productivities a and b are relatively small so that $b > (a - s)$, where s is a search cost parameter described below. Finally, agents within one country or agents from two host countries cannot form a joint venture.

The total N_j population in each host economy j consists of $\sum_i m_{ij}$ immigrants from source countries i and $N_j - \sum_i m_{ij}$ native agents, where m_{ij} equals the immigrant population from i residing in j. Source economies i consist of native agents only. Native agents in i and j have to incur search costs s in case they opt for foreign investment and look for potential trade partners.

During their random search for joint venture, native agents in j might fail to meet foreign agents with a probability $(1 - p_j)$, where p_j equals the probability that a searching native agent from country j forms a joint venture. Immigrants in j coming from source countries i are identical to native agents, but they know identity of agents from source country i without having to incur search costs s. Note that given this absence of search costs and the level of uniform match quality b, immigrants never choose to produce locally or to form a joint venture with agents from other than their source country i. Instead, they contact native agents in source economy i and set up a joint venture. Native agents in i always accept the immigrants' offer, because $b > 1$, the productivity differences a and b are relatively small, and the agents do not have to incur search costs, given that they were contacted by the foreign party.

The remaining populations in each country anticipate the choices of immigrants and

of contacted native agents in source economies and select local production if and only if net expected profits exceed gains from a joint venture and/or uncontacted native agents in source i would not accept the potential offer. The participation constraints of native agents in host country j are:[7]

$$\text{produce locally iff} \quad 1 > [(1-p_j)+p_j a] - s \text{ or}$$
$$1 \leq [(1-p_j)+p_j a] - s \text{ and } 1 > [(1-p_i)+p_i a] - s$$
$$\text{invest in other countries iff} \quad 1 \leq [(1-p_j)+p_j a] - s \text{ and } 1 \leq [(1-p_i)+p_i a] - s,$$

where p_j corresponds to

$$p_j = \frac{\sum^i \left[N_i - \sum^j m_{ij}\right]}{N_I} \min\left[1, \frac{N_I}{N_J}\right],$$

and p_i equals

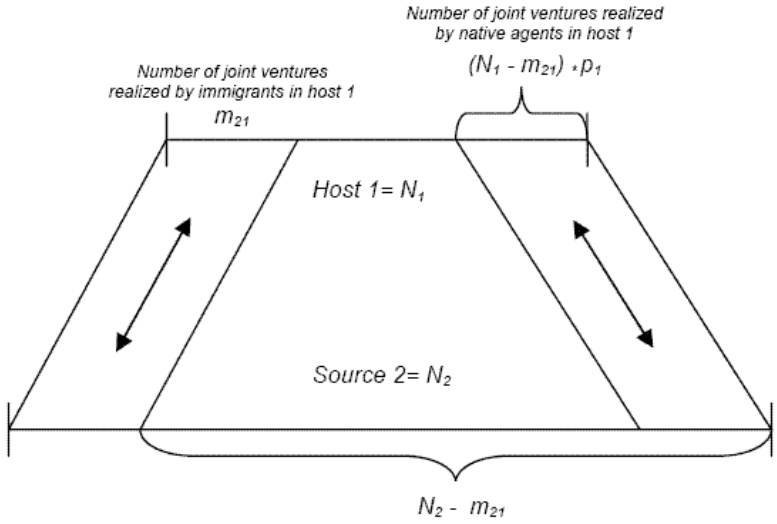
$$p_i = \frac{\sum^j \left[N_j - \sum^i m_{ij}\right]}{N_J} \min\left[1, \frac{N_J}{N_I}\right].$$

The participation constraints of uncontacted native agents in i are the same except that p_i replaces p_j. Figure 1 outlines an example with the world economy consisting of host country 1 and source country 2. The picture shows that immigrants m_{21} coming from source country 2 and residing in host country 1 match with native agents in 2 and set up joint ventures. The remaining native population $N_1 - m_{21}$ in country 1 and $N_2 - m_{21}$ in country 2 decide to either produce locally or to search for a foreign partner. Figure 1 represents a situation in which all agents try to form a joint venture. Nonetheless, only a fraction in each of the two economies succeeds in finding a foreign partner, the rest produce locally.

I take an approximation and assume the shares of overall immigrant populations in

[7] I assume both investors in a joint venture play Nash bargaining solution and split the resulting joint surplus $2a$ equally. Searching parties cover their costs s individually.

Figure 1: Matching in world economy with one host and one source country.



host countries and the size of immigrant communities with respect to their source country populations are sufficiently small, i.e., $h_j = \frac{\sum^i m_{ij}}{N_j} \to 0$, $\forall j$ and $d_i = \frac{\sum^j m_{ij}}{N_i} \to 0$, $\forall i, j$.[8] Then $p_j \to 1$, $p_i \to 1$, and country j's share in the aggregate output of all host countries equals

$$\frac{GDP_j}{GDP_J} = \frac{N_j \left[(1-h_j)\left((1-p_j)(1-s) + p_j(a-s)\right) + bh_j\right]}{\sum^j N_j \left[(1-h_j)\left((1-p_j)(1-s) + p_j(a-s)\right) + bh_j\right]} \approx \frac{N_j}{\sum^j N_j}, \tag{1}$$

where the terms in the brackets correspond to the contributions of local production, immigrant joint ventures, and joint ventures of native agents.

Similarly, a source country i's share in output of all source countries corresponds to

$$\frac{GDP_i}{GDP_I} = \frac{N_i \left[(1-d_i)\left((1-p_j)(1-s) + p_j(a-s)\right) + bd_i\right]}{\sum^i N_i \left[(1-d_i)\left((1-p_j)(1-s) + p_j(a-s)\right) + bd_i\right]} \approx \frac{N_i}{\sum^i N_i}. \tag{2}$$

For $N_J \leq N_I$ and using (1), trade volume T_{ij}^N generated by host j natives' joint ventures

[8] The average immigrant share in host countries $\frac{\sum^i m_{ij}}{N_j}$ in the sample is 0.026 and the average size of immigrants relative to source country populations $\frac{\sum^j m_{ij}}{N_i}$ equals 0.033.

equals[9][10]

$$T_{ij}^N = (a-s)\, N_J \frac{GDP_i GDP_j}{GDP_I GDP_J} \left(1 - \frac{\sum_{j=1}^J m_{ij}}{N_i}\right)\left(1 - \frac{\sum_{i=1}^I m_{ij}}{N_j}\right) \tag{3}$$

and trade volume T_{ij}^I generated by the immigrants from i residing in j is

$$T_{ij}^I = bN_J \frac{m_{ij}}{N_j} \frac{GDP_j}{GDP_J}, \tag{4}$$

where use was made of (1). Summing the last two expressions, one obtains the relationship for bilateral trade:

$$T_{ij} = T_{ij}^N + T_{ij}^I \tag{5}$$

$$= (a-s)\, N_J \frac{GDP_i GDP_j}{GDP_I GDP_J} \left[\left(1 - \frac{\sum_{j=1}^J m_{ij}}{N_i}\right)\left(1 - \frac{\sum_{i=1}^I m_{ij}}{N_j}\right) + \frac{b}{a-s} \frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP_I}}\right]. \tag{6}$$

Pre-multiplying by $\left(1 - \sum_{j=1}^J m_{ij}/N_i\right)\left(1 - \sum_{i=1}^I m_{ij}/N_j\right)$, taking logarithms and approximating $\ln(1+x) \sim x$ for x small, one obtains

$$\ln T_{ij} = \ln\left((a-s)\, N_J \frac{GDP_i GDP_j}{GDP_I GDP_J}\right) - \frac{\sum_{j=1}^J m_{ij}}{N_i} - \frac{\sum_{i=1}^I m_{ij}}{N_j} + \phi_{ij} \frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP_I}}, \tag{7}$$

where

$$\phi_{ij} = \frac{b}{(a-s)} \left(1 - \frac{\sum_{j=1}^J m_{ij}}{N_i}\right)^{-1} \left(1 - \frac{\sum_{i=1}^I m_{ij}}{N_j}\right)^{-1}$$

Finally, for estimation purposes, I use the general version of (7):

$$\ln T_{ij} = b_0 + b_1 \ln GDP_i GDP_j + b_2 \frac{\sum_{j=1}^J m_{ij}}{N_i} + b_3 \frac{\sum_{i=1}^I m_{ij}}{N_j} + b_4 \frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP_I}} + b_5 \left(\frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP_I}}\right)^2$$

[9] The case where $N_I \leq N_J$ does not change the line of argument.

[10] It might happen that the middle term in brackets, and hence predicted trade, can turn negative. The situation corresponds to a hypothetical country with its overseas diaspora larger than the country's domestic population. As all observations in the present sample are positive, I assume such a situation does not occur.

$$+a'z + \alpha_j + \varepsilon_{ij}, \tag{8}$$

where $\ln T_{ij}$ corresponds to the natural logarithm of either exports or imports flowing between countries i and j.

The coefficients b_2 and b_3 indicate the indirect impact on *native*-driven bilateral trade between i and j that has been caused by the immigrants' choice to trade with their source countries (see Equation 3) and are expected to be equal to minus one. The coefficient b_2 captures the effect on bilateral trade of source country diasporas located in other countries. The larger is the overall diaspora relative to the population of the country of origin, the lower are the chances of a host's *native* agents to find a match in the concerned source country. Since b_2 relates to the population of a source country N_i and approximates the potentially negative impact on *native*-driven bilateral trade, in the following I call the relative size of the diaspora $\frac{\sum_{j=1}^{J} m_{ij}}{N_i}$ the source country trade diversion term.

The coefficient b_3 captures the role of the overall share of immigrants in host j's population. Using the logic of the present empirical model, the more immigrants in a given host country match with agents in their countries of origin, the lower will be the probability of the host's *native* agents to trade with a given trade partner. b_3 connects to the population of a host country N_j and similarly to the coefficient b_2, which estimates the negative impact on *native*-driven bilateral trade. For these reasons, I label the overall immigrant share in host j's population $\frac{\sum_{i=1}^{I} m_{ij}}{N_j}$ the host country trade diversion term.

Being an empirical counterpart of ϕ_{ij} in Equation 7, the coefficient b_4 reflects the direct trade contribution by immigrants from i located in j (see also Equation 4) and is expected to be positive.[11] Note that the corresponding term differs from the commonly used natural logarithm of immigrant stock[12] as well as other commonly employed measures of immigrant links and has the source country GDP_i in its denominator. While the natural

[11] I assume the parameter ϕ_{ij} in Equation 7 to be constant across all pairs ij, i.e., $\phi_{ij} = \phi$. This certainly leads to a measurement error in the right-hand-side variable and a subsequent coefficient bias towards zero. On the other hand, the estimates explicitly accounting for $\left(1 - \sum_{j=1}^{J} m_{ij}/N_i\right)\left(1 - \sum_{i=1}^{I} m_{ij}/N_j\right)$ practically do not differ from the simplifed output with ϕ replacing ϕ_{ij}. The estimation results are available upon request.

[12] The natural logarithms have been used by Head and Ries (1998); Girma and Yu (2002); and Heerander and Saavedra (2006).

logarithm formulation remains intuitively appealing and easy to interpret, it suffers from the lack of theoretical justification and zero predicted trade in the absence of immigrant networks. The immigrant terms derived within the present framework rely on an explicit model and emphasize relative rather than absolute measures of immigrant networks. As the coefficient b_4 reflects direct positive immigrant effects on trade, the corresponding term will be referred to as the trade creation term.

The emphasis on the relative number of immigrants derives from the model's assumptions of different populations across host and source countries, and the possibility to form a joint venture with one agent only. Other things being equal, the higher the fraction of host j's population represented by immigrants from i, the more joint ventures will be formed with agents in the immigrants' source country i. Similarly, the larger the economy is of the immigrants' country of origin, the higher agent j's probability of forming a joint venture will be with an agent from i, and the smaller will be the immigrants' relative contribution to bilateral trade between i and j. The relationship between the absolute measures of immigrant links (such as the natural logarithm of immigrant stock) and the relative measure derived within the present framework will be discussed in Section5.

Larger immigrant communities might tend to trade with each other instead of trading with their country of origin. To accomodate a possible trade substitution, I add a quadratic approximation of the trade creation term with a negative expected sign of the coefficient b_5.[13] It should be remembered that in order to obtain the net effect of immigrants on bilateral trade between i and j, one should take into account both the trade-creation and trade-diversion effects of immigrant links.

z is a $k \times 1$ vector of additional explanatory variables that vary either at the level of host j, source i, or at the level of country pairs ij. The former two groups include export shares in the GDP as a proxy for openess and institutional quality measures. The country-pair ij variables consist of the natural logarithm of distance, the product of GDP's

[13] While the immigrant ties introduced by the present matching model shift the geographical pattern of trade, they should not influence the total volume of trade between a given host country and its trading partners. If one is willing to accept the assumption of a more efficient matching technology by immigrant joint ventures, the total trade effect would be positive.

per capita (expressed in natural logarithms), and dummies for shared colonial past and common language.

Colonial past and common language are often used as proxies for informal trade barriers. As for the colonial dummy, entrepreneurs from a former colonial power, e.g., traders or specialized information agencies, might have extended business links from colonial times and thus possess valuable information and contacts. Furthermore, a former colonial power often played a key role in the design of local institutions in the source country. The resulting institutional proximity would then translate into relatively lower demands on the understanding of the local market environment. A common language dummy should capture lower search costs for all agents using the same mother language and again facilitate the matching process.

I divide the colony and language dummy variables by the GDP of a source country i, so that the resulting variables are non-increasing in the source economy's size. The expected signs of coefficient estimates for both variables are positive, resulting in a larger predicted trade impact of common language and/or colonial past for smaller source economies. Intuitively, had all the trading partners shared a colonial past (or language), the relative trade enhancing role of both would be zero. As the trading partner gets smaller in size, however, their relevance should tend to increase as a smaller open economy tends to be relatively more sensitive to trade barriers.

The error term has two components. ε_{ij} is a random term specific to individual country pairs ij and independent of other errors. α_j correponds to an error term that is correlated within host country j. If common group errors α_j have not been controlled for, the resulting standard error estimates might suffer from a notable downward bias (Moulton, 1986). I thus allow for a more general covariance structure and heteroscedasticity of α_j as proposed by Liang and Zeger (1986). As an alternative form of adjustment for common-group errors, I employ the 2-step estimation approach by Donald and Lang (2007) that generates more reliable estimates in case the number of groups is small.

The advantage of the latter procedure is its robustness in case the number of groups

is small, so that researchers do not have to rely on the asymptotics along the number of groups necessary for the `cluster` command.

The two-step procedure starts with the OLS regression of the natural logarithm of bilateral exports/imports on variables differing across country pairs ij, country j- and i-fixed effects:

$$\text{1st stage: } \ln T_{ij} = b_0 \ln GDP_i GDP_j + \mathbf{x}'_{ij} b + a_0 \frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP_I}} + d'_{ij} + \varepsilon_{ij},$$

where the term following the coefficient a_0 is the newly added share in the host population of a given immigrant stock relative to the country of origin GDP_i.

In the second stage, I run a feasible GLS with the relevant fixed effect coefficient estimates from the first stage as dependent variables and country i- (or j-) level variables on the right-hand side of the regression:

$$\text{2nd stage:} \quad \hat{d}_j = c(J) + \mathbf{x}'_j z + a_1 \frac{\sum_{i=1}^{I} m_{ij}}{N_j} + u_j, \; \hat{var}(u_j) = \hat{\sigma}^2 I(J) + \Sigma_{\hat{d}_j} \tag{9}$$

$$\text{and} \quad \hat{d}_i = c(I) + \mathbf{x}'_i w + a_2 \frac{\sum_{j=1}^{J} m_{ij}}{N_i} + u_i, \; \hat{var}(u_i) = \hat{\sigma}^2 I(I) + \Sigma_{\hat{d}_i}, \tag{10}$$

where Equation 9 estimates the coefficient on the host trade diversion term, Equation 1.10 estimates the coefficient on the source trade diversion term, and $\hat{var}(u_{\{j,i\}})$ stands for the variance of the respective 2nd-stage error term $u_{\{j,i\}}$. The vectors of country-specific terms $\mathbf{x}_i$ and $\mathbf{x}_j$ include the natural logartihms of real GDP and GDP per capita, the corresponding relative measure, share of exports in GDP, and the Heritage Foundation measure of institutional quality. $\frac{\sum_{i=1}^{I} m_{ij}}{N_j}$ stands for the population share of the overall immigrant stock (regardless of origin) within a given host country, $\frac{\sum_{j=1}^{J} m_{ij}}{N_i}$ represents the size of the overseas diaspora relative to the population in the diaspora's country of origin. The GLS procedure uses fixed effect covariance estimates $\Sigma_{\{\hat{d}_j, \hat{d}_i\}}$ from the 1st stage for

the construction of weights.[14]

4 Data

4.1 Immigrants

The cross-country information on the numbers of foreign-born persons over 15 years of age for 19 OECD member countries was retrieved from the OECD Statistics Portal on Demography and Population.[15] The advantage of the present dataset rests in the variation at both the source *and* host country levels, which permits the estimation of trade-diversion effects. This was not possible in empirical studies that focused exclusively on a single host country.

The OECD data represents the first attempt to create a coherent dataset covering several host countries. The data have been drawn from population registers, residence or work permits, surveys and censuses taking place usually every 5 or 10 years. Due to the different timing of censuses, the reference year varies between 1999 and 2002, depending on the specific country. Some OECD countries had to be dropped due to large proportions of foreign-born populations with the country of origin unknown.[16] For host countries that were left in the sample, the values of unknown foreign-born populations did not exceed 2%. These unknown populations have been distributed using country-of-origin shares in the total number of foreign born in a concerned host country. The new entities on territories of the former Soviet Union and Yugoslavia have not been included due to differences in aggregation across host countries.[17]

[14] For more details see Donald and Lang (2007), p. 224-225.

[15] Other studies on trade and migration using the OECD migration data include working papers by Dolman(2007), and Felbemayr and Toubal (2008).

[16] These include Australia (16.2% unknown); the Czech Republic (28.2%); Mexico (41.9%); New Zealand (16.1%); Poland (41.1%); the Slovak Republic (9.3%); and Switzerland (14.7%). The borderline cases, Finland (3.8%) and Denmark (6.7%) were left in the sample.

[17] Turkey, the last OECD member in the sample, is in many respects closer to a typical developing country and its membership in the OECD owes more to strategic considerations rather than to the level of economic development. Nonetheless, despite being left out from the main regressions, the results with Turkey as a host country remain both quantitatively and statistically similar to the main regression results listed in Table 2. Results including Turkey can be provided upon request.

The figures for Germany were listed only by broad source regions instead of countries. For the Netherlands, the data included only the number of all who are foreign born instead of those over 15 years of age. I replaced the data for Germany with figures from the Federal Statistical Office of Germany and, since the available data for both Germany and the Netherlands covered total foreign-born population only, I adjusted them by the shares of immigrants over 15 years of age in the total foreign-born population by source country as recorded for comparatively open Belgium. As part of the sensitivity analysis in Section 6, I drop the two host countries and run all regressions to check for the robustness of the results.

4.2 Trade and remaining data

The data on bilateral exports and imports have been obtained from the Direction of Trade Statistics compiled by the International Monetary Fund. Trade volumes of especially smaller developing countries can vary substantially from year to year. For that reason five-year averages of real trade volumes over 1999-2003 have been chosen instead of using the data for a single year only. The five-year averages reduce an additional problem with zero observed exports and imports.[18] Finally, since the focus of the present study is immigrant networks and the home links of overseas Chinese communities quite likely cover both China and Hong Kong, the two entities are treated as a single country.

The remaining variables, common language and a measure of circle distance between capital cities were retrieved from Jon Haveman's web page[19] and added manually if values were missing. A dummy for the common colonial past was constructed from the histories of each colonial power detailed in *Wikipedia*. The dummy equals one if the country in question was either a colony or protectorate after 1945. As a measure of institutional quality, I use the five-year averages for countries i and j of the restricted Index of Economic

[18] While 23 out of the total 1,684 sample observations on exports from host countries (i.e. roughly 1.4 percent) reported zero trade in at least one year over the 1999-2003 period, none of them did so for the whole five-year period. For imports to host countries the figures equalled 57 (i.e., 3.4 percent) and 18 respectively. The tentative random-effect tobit estimates using the `xttobit` command in Stata produced coefficient estimates that were qualitatively and quantitatively very similar to results in Table 2. These can be provided upon request. The export figures are reported f.o.b., the import volumes are c.i.f.

[19] Jon Haveman's web page can be found at http://www.macalester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/TradeData.html#Gravity.

Table 1: Summary statistics, n=1,684.

Variable	Mean	Standard deviation	Minimum	Maximum
Exports_{ij}	353.49	2,156.04	0.001	62,824.19
Imports_{ij}	324.23	1,780.94	0	48,734.65
Host GDP^*_j	1,164,183	2,101,759	77,757.52	9,012,508
Source GDP^*_i	46,081.42	128,490.4	575.76	1,027,513
Host GDP/capita_j	21,786.83	7,720.79	9,306.51	36,720.11
Source GDP/capita_i	3,034.50	4,972.48	100,78	29,185.42
$\text{Immigrant stock}_{ij}$	12,365.68	65,698.51	0	8,359,180
$\text{Trade creation}_{ij}$	0.44	1.55	0	12.59
Host diversion_j	0.02	0.02	0.002	0.07
$\text{Source diversion}_i$	0.03	0.06	0.01	0.33
Distance_{ij}	7,300.51	3,487.39	375	19,594
$\text{Export share host}_j$	0.42	0.22	0.11	0.98
$\text{Export share source}_i$	0.19	0.14	0.03	0.76
$\text{Institutional quality host}_j$	73.19	6.39	58.53	81.01
$\text{Institutional quality source}_i$	44.58	13.98	15.09	78.5
$\text{Shared colonial past}_{ij}/\text{GDP}_i$	0.01	0.02	0	0.15
$\text{Common language}_{ij}/\text{GDP}_i$	0.01	0.03	0	0.16

*in millions of 1998 U.S. dollars

Freedom produced by the Heritage Foundation. The Index of Economic Freedom over 1999-2003 compiles evaluations of nine areas essential for functioning market environment. The restricted version includes only those areas that most closely relate to institutional quality in trade context – corruption, non-tariff trade barriers, rule of law, and regulatory burden – and drops inflation, fiscal burden, restrictions on banks, labor regulation, and government intervention. Finally, figures on population, GDP, GDP per capita, and export shares in hosts' GDP were collected from the World Development Indicators published by the World Bank. To avoid the potential endogeneity problem of the GDP variables, I use GDP and GDP per capita figures from 1998 as proxies. The main sample consists of 19 host countries and 90 source countries, generating an unbalanced panel of 1,684 observations. Table 1 presents the summary statistics for key variables.[20]

[20] The OECD data include immigrants' numbers by education levels. In order to exploit this potentially useful information, I constructed measures of relative shares of college-educated migrants and used them during the empirical model estimation. Given that the estimation results proved to be insignificant and sometimes with a negative sign, I do not present them in the final text. The output can be provided upon request.

5 Empirical results

The estimated coefficients for the trade creation and diversion terms are reported in Table 2.[21] The first columns for both exports and imports display the estimates from the benchmark OLS regression with regional dummies for host and source countries and clustering by host country. In the following columns, I present the results of Donald and Lang's (2007) 2-step procedure, where the trade creation estimates have been obtained in the 1st stage. Columns (2) and (5) contain the 2nd stage estimates of the source trade diversion for exports and imports. Columns (3) and (6) report the estimated coefficients of the host country trade diversion term.

Table 2: Main regression results, dependent variables real exports and imports 1999-2003.

	Real exports 1999-2003		
	(1) OLS regional dummies i and j	(2) 2-step estimates for source i	(3) 2-step estimates for host j
Trade creation$_{ij}$	0.085***	0.056***	0.056***
	(0.021)	(0.017)	(0.017)
Trade creation$_{ij}^2$	-0.184***	-0.001	-0.001
	(0.056)	(0.001)	(0.001)
Source diversion$_i$	-0.117	-2.486***	-
	(0.476)	(0.939)	
Host diversion$_j$	-9.099	-	-6.911
	(6.036)		(4.105)
R^2	0.852	0.670	0.595
N	1,577	1,684	1,684
	Real imports 1999-2003		
	(4) OLS regional dummies i and j	(5) 2-step estimates for source i	(6) 2-step estimates for host j
Trade creation$_{ij}$	0.073***	0.044**	0.044**
	(0.017)	(0.018)	(0.018)
Trade creation$_{ij}^2$	-0.223***	-0.001	-0.001
	(0.06)	(0.001)	(0.001)
Source diversion$_i$	-1.475***	-1.654*	-
	(0.5)	(0.965)	
Host diversion$_j$	-10.408**	-	-4.929
	(3.999)		(11.235)
R^2	0.856	0.496	0.600
N	1,577	1,684	1,684

Notes: The OLS with regional dummies account for clustering by host countries.
***,**,* - Significant at 1%, 5%, and 10% respectively. Standard errors in parentheses.

[21] For a complete list of all explanatory variables and estimation results see Table A.2 in the Appendix. For Liang and Zeger's (1986) OLS estimation with clustering, Equation 8 has been supplied with regional dummies to control for possible correlation of explanatory variables with unobserved region characteristics. The five regional dummies for host countries correspond to North America, East Asia, Northern Europe, Central Europe, and Southern Europe, the UK and Ireland representing the benchmark economies. For source countries the regions are Northern Africa and Arab states, Subsaharan Africa, South Asia and South-East Asia, with Latin American countries being the baseline economies.

Table 3: Examples of trade creation in response to a 10 percent boost in immigrant stock for different host and source countries.

Host j	Source i	GDP_i's % share in world GDP	Immigrant population m_{ij}	Exports creation in %	Imports creation in %
Canada	Bangladesh	0.14	19,515	0.24	0.19
	Tanzania	0.03	19,525	1.19	0.93
France	China	3.68	32,913	0.01	0.01
	Cameroon	0.03	33,125	0.95	0.75
the Netherlands	Pakistan	0.24	10,052	0.15	0.11
	Ghana	0.16	10,311	2.2	1.73
the UK	South Africa	0.44	124,658	0.27	0.21
	Kenya	0.044	125,491	2.70	2.12
the USA	Panama	0.039	132,975	0.68	0.53
	Cambodia	0.11	133,240	2.46	1.93

5.1 Trade creation

Regardless of specification and direction of trade, the estimated coefficients on trade creation are consistently positive, relatively stable, and significantly different from zero at least at the 5 percent significance level. The marginal trade creation impact of a 10-percent increase in immigrant stock m_{ij} depends on the level of m_{ij}, population of host j, and the output GDP_i of source country i (see Equation 8). This dependance differs from the studies using the natural logarithm of immigrant stock, where the marginal impact is fully described by the estimated regression coefficient. Table 3 provides examples of the implied export and import creation resulting from a 10 percent boost of immigrant stock for country pairs ij that have different levels of GDP_i, but are otherwise comparable in terms of both m_{ij} and N_j. The estimates suggest that for source countries with smaller GDP_i levels, a given number of immigrants connect to a relatively larger part of the source economy. Due to the gravity relationship linking output with trade, these connections then translate into relatively higher shares in trade between host j and source i.

The implied marginal trade creation effects lie within the interval $\langle 0, 1\rangle$ in more than 91,5 percent of country pairs in the sample and generally fall short of marginal effects re-

ported by studies using the natural logarithm of immigrant stock.[22] Apart from the measurement error of the trade creation term discussed in Section 3, lower marginal effects can be partly explained by the cross-sectional nature of the sample and low immigrant levels m_{ij} in a number of host countries. Focusing on the trade effect of a 10-percent increase in immigrant stock m_{ij} and holding other things constant, country pairs ij with smaller immigrant populations generate lower marginal trade effects as compared to observations with more numerous immigrant stocks. The smaller marginal impacts (as compared to earlier empirical studies) nonetheless apply to all host economies and pairs ij, regardless of immigrant population size m_{ij}.

Another potential explanation relates to the role of source country GDP_i. Immigrants from source countries with lower levels of GDP_i are predicted to trade relatively more than their counterparts from larger source economies. This is intuitive if trade is proportional to the GDPs of trading parties (as in the model from Section 3) and the immigrant trading technology has constant returns to scale, since then a given number of better-informed immigrants will generate a lower fraction of the overall trade volume.

The lower magnitude of the trade-creation effects can be also explained by the relative productivity of joint venture matches as captured by the ratio $\frac{b}{a-s}$ in Equation 6. While immigrants are more likely to understand source i's environment and business practices as compared to native agents from host j, they could lack the knowledge necessary for exports of more sophisticated and value added products. For example, Turkish traders in Germany might specialize in the trading of used cars or ethnic goods instead of power engines. Relatively less productive matching (as compared to matches initiated by host j's natives) might be rational especially if the immigrants' outside options in host j are not sufficiently profitable. The outside options of immigrants might be thought of as a function of profficiency in host-country language, legal status, and/or experience with the host's labor market. Given this assumption, they are likely to be lower than the opportunities of

[22] For example, a static version of the model by Girma and Yu (2002) produces a 1.6 percent increase in UK exports and a 1 percent rise in UK imports from non-Commonwealth countries. Head and Ries (1998) find a 1-1.3 percent boost for Canadian bilateral exports and 3.1-3.9 percent for imports. The study on U.S. exports by Herander and Saavedra (2005) reports 1.6 percent.

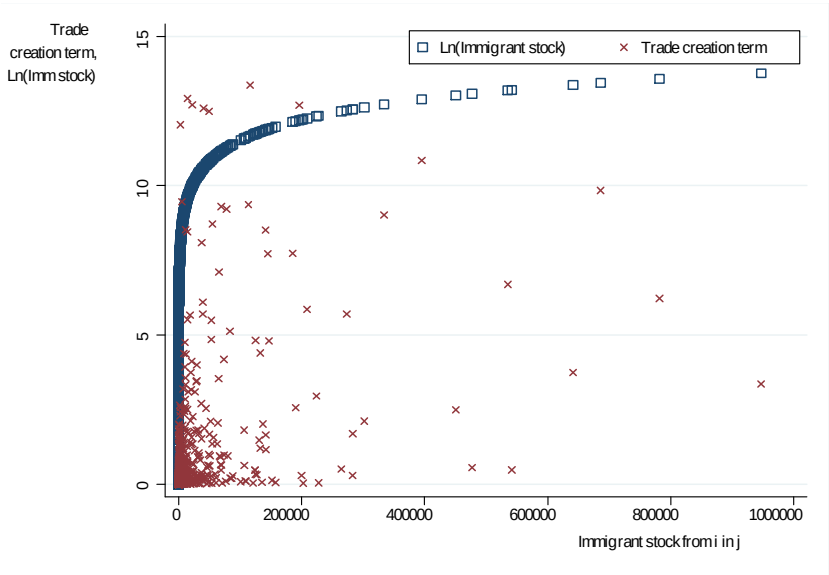
native agents.[23] Sectors that have some bearing on the immigrants' source country thus could provide one of few opportunities to employ immigrant human capital gainfully, even though the ultimate contribution to trade might be relatively lower than that of native agents.

Finally, immigrant networks could operate across a larger number of countries. In such a case, the matching mechanism in the model from Section 3 might be too restrictive and low trade creation estimates would be capturing only a fraction of the total effect.

5.1.1 Trade creation term vs. natural logarithm of immigrant stock

In this section, I focus on the relative performance of the trade creation term and the commonly employed level measures such as the natural logarithm of immigrant stock. Figure 2 illustrates the sample relationship between the absolute immigrant stock m_{ij}, its

Figure 2: Immigrant stock, its natural logarithm, and the trade creation term.



natural logarithm $\ln(m_{ij})$, and the trade creation variable derived in Section 3.

[23] The empirical study on Izraeli labor market by Friedberg (2000) found that immigrants' education obtained abroad is significantly less rewarded than education received locally. Similarly, Chiswick and Miller (1995) focus on the impact of language profficiency on immigrants' earnings in Australia and three other countries (the USA, Canada, and Israel), and show that higher fluency in host's language significantly increases immigrants' earnings.

The figure indicates that the trade creation term is only weakly related to the natural logarithm of immigrant stock.[24] Table 4 reports the estimates from regressions with host j and source i fixed effects and clustering by host country. The regressions employ both specifications of the immigrant variable, first separately and then simultaneously.

The coefficient estimates from the specification with the natural logarithm in columns (2) and (5) resemble results from the previous studies. For the present dataset, a 10 percent increase in the immigrant stock leads on average to a 1.06 percent boost of exports from and a 1.13 percent increase in imports to the host country. The natural logarithm specification of the immigrant variable, however, suffers from the ignorance of the trade partner's economic size and immigrants' share in host j's population.

Figure 3 summarizes the differences between the two specifications in predicted bilateral export increases following the 10-percent rise of immigrant population m_{ij} in host j. The horizontal line indicates the marginal effect obtained from the natural logarithm specification in Column 2 – i.e., the value of the coefficient on $Ln(Imm\ stock)_{ij}$, multiplied by 10.

Table 4: Trade creation term vs natural logarithm of immigrant stock, fixed effect estimates.

	Real exports 1999-2003			Real imports 1999-2003		
	(1)	(2)	(3)	(4)	(5)	(6)
Trade creation$_{ij}$	0.056***	-	0.02*	0.044**	-	0.025*
	(0.017)		(0.01)	(0.018)		(0.013)
Trade creation$^2_{ij}$	-0.001	-	-0.035	-0.001	-	-0.049
	(0.001)		(0.038)	(0.001)		(0.047)
Ln(Imm stock)$_{ij}$	-	0.106***	0.092***	-	0.113***	0.095***
		(0.02)	(0.021)		(0.024)	(0.025)
R^2	0.910	0.911	0.912	0.906	0.906	0.908
N	1,684	1,684	1,684	1,684	1,684	1,684

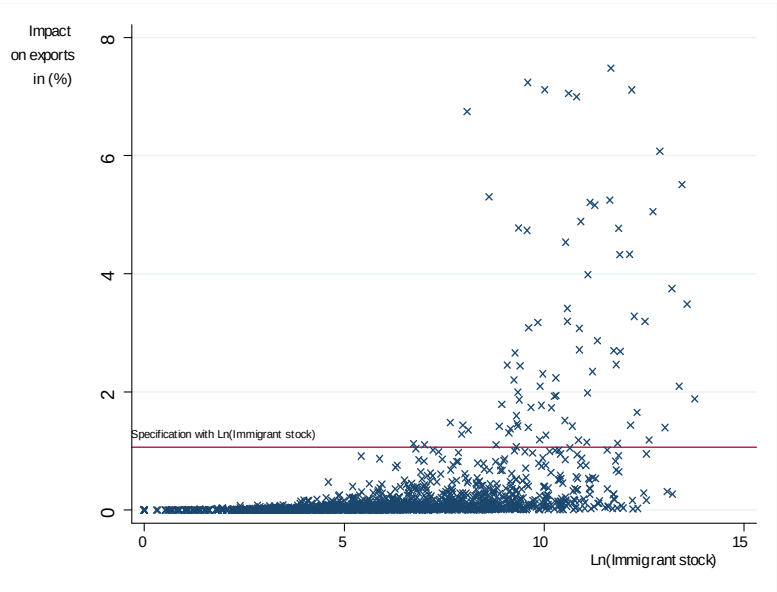
Note: All estimates account for clustering by host countries. Standard errors in parentheses.
***,**,* - Significant at 1%, 5%, and 10% respectively.

The implied effect from the natural logarithm specification thus remains the same regardless of the actual size of the immigrant community m_{ij}. In the case of the calculated marginal impacts obtained from the specification derived in Section 3 using results listed in Column (1) in Table 4, the trade increases center mostly around larger immigrant

[24] The correlation coefficient between the trade creation term and the natural logarithm of immigrant stock equals 0.126.

populations, leaving smaller communities m_{ij} without any notable effect on trade.

Figure 3: Marginal trade creation effects of a 10-percent increase in m_{ij}, matching model predictions and real exports 1999-2003.



Finally, Columns (3) and (6) in Table 4 report the results from the estimation including simultaneously the trade creation term and the natural logarithm of immigrant stock.[25] One can observe that despite a drop in the levels of the trade creation term, the combination of relative and absolute measures preserve the statistical significance of both. The levels and joint significance of the relative and absolute terms suggest that despite the significance of the trade creation term, and its account for the trade partner's output GDP_i and the relative size of the immigrant population $\frac{m_{ij}}{N_j}$, the model from Section 3 captures only part of the trade-immigration story. The next sub-section focuses on the trade diversion estimates.

[25] The 2nd stage estimates of host and source trade diversion coefficients did not change substantially and can be provided to the interested reader.

5.2 Trade diversion

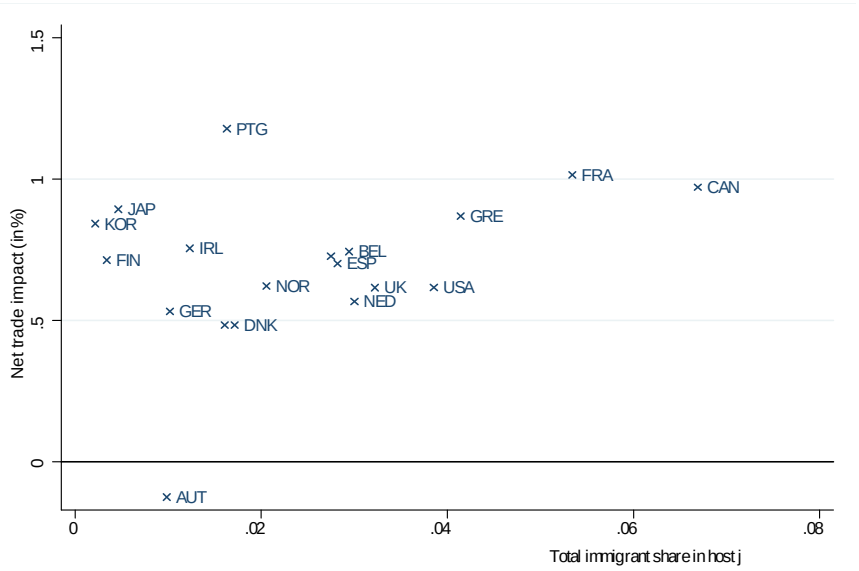
Regardless of specification and direction of trade, the host and source trade diversion terms in 2 have expected signs. Focusing on the estimates obtained through Donald and Lang's (2007) 2-step procedure, the source trade diversion terms differ from zero at least at the 10-percent probability level, and all trade diversion coefficients are negative and not statistically different from minus one as predicted by the model from Section 3. A one-percentage-point increase in the size of the total immigrant community $\sum_{j=1}^{J} m_{ij}$ relative to the source country i's population would result in a decrease in its *total* exports by roughly 2.5 percent and its *total* imports by 1.7 percent on average. The host diversion estimate is statistically not different from zero.

Figures 4 and 5 present the net overall effect on trade of host and source countries, using the coefficient estimates from the 2-step procedure by Donald and Lang (2007) that include the natural logarithm of immigrant stock and trade creation and diversion measures derived in Section 3.[26] Each figure provides an answer to one of two simple questions. 1) Do immigrant communities located in a given host j facilitate aggregate trade between host j and immigrants' countries of origin? 2) Do source countries with larger shares of population located in advanced OECD economies on aggregate benefit from immigrant-driven trade links?

To answer the first question, I consider the implied marginal impact on the sum of exports and imports of a balanced 10-percent increase across the total immigrant population $\sum_{i=1}^{I} m_{ij}$ for a given host j, holding total population constant. For the second question, I employ the same proportional increase of a given source i's natives located in OECD countries $\sum_{j=1}^{J} m_{ij}$, again fixing source i's population N_i. Figure 4 presents the predicted impact on the sum of exports and imports for 19 OECD host countries in the sample. Present results are consistent with the positive role of immigrant links found by previous studies, with Austria being the only OECD country with a negative predicted impact of

[26] I did not include the host trade diversion term in the computations of net trade effects, given that it was not statistically different from zero. Net trade effects on exports and imports for individual host and source countries can be found in Table A.1 in the Appendix.

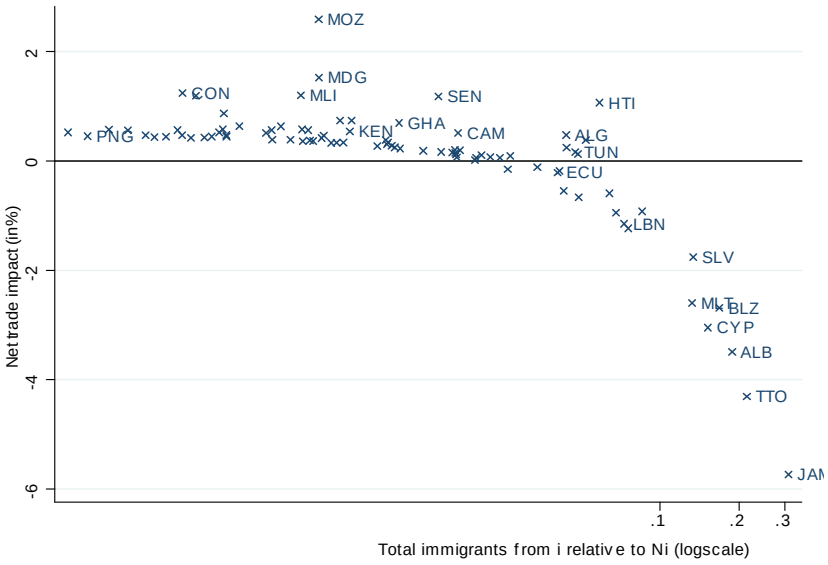
Figure 4: Host j's aggregate trade with source countries, and a 10% increase in total immigrant stock.



immigrants on trade. The positive role of immigrant links related to information provision, informal contract enforcement, and preferences for source-country products thus seem to dominate the potential losses due to the associated shifts in trade of the host countries.

A similar conclusion holds for source countries and their populations located in OECD host countries. Seventy-one out of 90 source economies show a positive marginal impact of immigrant links on the economies' total trade with OECD hosts. The marginal impact on trade on average declines with rising shares of source i's population located in OECD host countries. Since the trade creation term does not change substantially with the rising shares of source i's total overseas population, the trade diversion channel gradually gains in importance. The net effect of immigrants on trade might even turn negative in case the productivity of matches between immigrants and natives from source i fall short of productive matches forgone by host j's native agents. The next section focuses on the robustness of the estimated results.

Figure 5: Source i's aggregate trade with OECD countries, and a 10% increase in total immigrant stock.



6 Sensitivity analysis

6.1 Relative size of source markets, large immigrant stocks, and adjustments in migration data

The model and econometric specification in Section 3 assume the trade creation coefficient to be identical across all observations, and the only country-pair ij variation in estimated trade creation effects to be driven by differences across immigrant stocks m_{ij}, host populations N_j, and/or source countries' GDP_i. With a larger immigrant share $\frac{m_{ij}}{N_j}$ relatively to source-country GDP_i, immigrant networks might face a decreasing number of profitable trading opportunities or tougher competition between individual network participants, which translate into lower profit margins. The trade creation coefficient would then likely vary across different country groups.

To evaluate this hypothesis, I construct two additional variables *Trade creation large*$_{ij}$ and *Ln(Imm stock large*$_{ij}$*)*, which are equal to the values of *Trade creation*$_{ij}$ and *Ln(Imm stock*$_{ij}$*)* in case *Trade creation*$_{ij} > 1$, and zero otherwise. These variables should capture

Table 5: Estimates distinguishing source countries with relatively large immigrant communities, i.e. *Trade creation*$_{ij} > 1$.

	Real exports 1999-2003			Real imports 1999-2003		
	(1)	(2)	(3)	(4)	(5)	(6)
Trade creation$_{ij}$	0.607***	-	0.409***	0.522***	-	0.391**
	(0.085)		(0.133)	(0.101)		(0.141)
Trade creation large$_{ij}$	-0.545***	-	-0.427***	-0.473***	-	-0.435***
	(0.08)		(0.129)	(0.09)		(0.136)
Ln(Imm stock)$_{ij}$	-	0.128***	0.096***	-	0.117***	0.089***
		(0.022)	(0.026)		(0.026)	(0.03)
Ln(Imm stock large)$_{ij}$	-	0.002	0.028**	-	0.008	0.04***
		(0.007)	(0.012)		(0.008)	(0.012)
R^2	0.911	0.913	0.914	0.900	0.903	0.904
N	1,684	1,684	1,684	1,684	1,684	1,684

Note: All estimates account for clustering by host countries. Standard errors in parentheses. ***,**,* - Significant at 1%, 5%, and 10% respectively.

the trade impact of immigration for country pairs with large immigrant communities m_{ij} (in terms of host j's population) relatively to market size in country of origin i.[27] I then run the fixed effects regression allowing for clustering by host country and compare the obtained estimates with previous results. Table 5 presents the regression output.

The estimates show that given the use of the natural logarithm of immigrant stock, immigrants' contribution to trade between countries with relatively large immigrant communities in host j combined with small source-country i markets is no different from others. The situation becomes radically different once trade creation measure from Section 3 is employed. The coefficient estimates suggest that previous results from Table 2 in fact averaged the effects across country pairs with rather heterogeneous immigrant-trade links. The estimates maintain relatively high levels even after the simultaneous inclusion of both proxies for immigrant networks. A relatively small market size in country of origin i thus might prevent the full realization of benefits from immigrant-driven trade due to, for example, more intensive competition among traders and resulting lower markups. This is not to say that concerned country pairs do not benefit from immigrants at all. The signs and statistical significance of the natural logarithm of immigrant stock in fact indicate that the absolute size of immigrant community matters even more for source countries with

[27] The 142 out of 1,684 observations having trade-creation values above one consist mostly of trade partners with a former colonial relationship (38 out of 53 colonial pairs in the sample), or poorer/small economies with disproportionately large overseas diasporas.

relatively small markets.[28] This result is consistent with the study by White (2007b), given that these source countries have relatively lower GDP per capita levels with respect to the rest of the sample.

Apart from the heterogeneity across the trade creation dimension, the estimated outcomes might be possibly driven by a handful of source countries with large immigrant populations. To account for this possibility, for each host j, I drop five source countries (out of 90 non-OECD states) with the highest share in the overall immigrant stock. The levels and the statistical significance of the output, however, remain the same and can be provided upon request.

The discussion of the data on foreign-born persons in Section 4.1 mentioned the adjustments made to allow the inclusion of two key host countries, Germany and the Netherlands, into the sample. I run the whole estimation again and drop both host countries. Again, the results do not change substantially, and the coefficients of interest remain highly significant.

6.2 Endogeneity of immigrant variables

The potential endogeneity of trade creation and diversion terms might cast some doubt on the presented results. Over time, trade partners could learn about the living conditions in the other country and might pass the information further to potential migrants. Growing bilateral trade might likewise provide employment opportunities within the immigrant communities engaged in trading and thus reduce the *ex ante* uncertainty of agents considering migration.

While similar reasoning seems to be in line with the findings of the literature on international migration,[29] previous studies on immigrant networks have avoided the endogeneity issue. Indeed, finding a suitable instrument for the trade creation variable proves to be

[28] Figure A.1 and Figure A.2 in the Appendix show the net trade effects for a balanced 10-percent rise of immigrant stock, using the coefficient estimates from Columns (3) and (6) in Table 5 and the corresponding 2-stage estimates. Figure A.3 presents trade creation predictions for a 10-percent increase in m_{ij} generated by the matching model from Section 3.

[29] Focusing on the key pull and push factors shaping international migration decisions, Mayda (2005) finds a statistically significant positive effect of bilateral trade.

a daunting task. An exception is the study of Javorcik, Ozden, Spatareanu and Neagu's (2006) on migrant networks' links and foreign direct investment. The authors use the natural logarithm of population density and the share of passport costs in real GDP per capita in the source country from McKenzie (2005), both identified as significant push factors for migration. For the present purposes, however, the correlations between the stock of immigrants, population density in the source country, and passport costs seem to be negligible and in the former case even with the opposite sign. In the 2SLS regressions on exports and imports with the logarithms of both IVs and the natural logarithm of immigrant stock as the instrumented variable, the Shea partial R-squared failed to pass 0.01 for any combination of the instruments, and the joint F-tests in the first stage did not prove to be significant. The weakness of the available instruments thus precludes the quantification of the degree of endogeneity, at least in terms of the trade creation term.[30]

Moving to the trade diversion terms, any significant endogeneity problem seems to be of minor relevance. The trade diversion variables relate the *total* immigrant shares in host and source population to *bilateral* trade. If bilateral trade between countries i and j promotes international migration between the two yet not between the host or source country and other economies, its contribution to the total immigration shares would be most likely negligible.[31] Moreover, the mutual relationship between the immigration shares and bilateral trade should be positive, whereas the trade diversion terms establish a *negative* link. Hence, if anything, the endogeneity would underestimate the impact of trade diversion by immigrant networks.

[30] The instruments work better for subsets of countries from South-East Asia and Subsaharan Africa. Using the natural logarithm of immigrant stock, for the Subsaharan region the exogeneity of dependent variables could not be rejected, while in the case of South-Asian economies the exogeneity hypothesis has been rejected at least at the 2-percent level. The natural logarithm of immigrant stock remained nonetheless highly significant and stable in both trade directions.

[31] The shares in the host population for the largest source country i do not exceed 2.1 percent.

7 Conclusion

The study complements research on the links between immigrant networks and international trade. The trade creation measure derived within the matching framework points to the importance of the relative size of a given source country economy and immigrant network. I estimate the differential impact of immigrant links based on the GDP of their respective country of origin and find that the immigrant communities from relatively larger economies facilitate trade less than what is implied by existing studies.

While previous work focused largely on trade creation by immigrant networks, I also derive trade diversion measures capturing negative spillovers to host and source countries' total exports. While immigrant networks can mitigate some informal barriers to trade (e.g., the lack of information on foreign markets or ineffective contract enforcement institutions), the same networks' advantages coupled with the pervasive presence of informal trade barriers might lead to shifts in trade patterns previously known e.g., in the context of customs unions. By channeling trade to the immigrants' country of origin, potentially more profitable matches in other countries become lost. Using a dataset of 19 OECD countries, I find some empirical support for this hypothesis.

Apart from being statistically significant, the results are robust to the inclusion of commonly used level measures of immigrant stock. Nonetheless, more work needs to be done in the search for valid instruments that could better capture potential endogeneity concerns relating to the immigrant network variables. Future extensions that allow for heterogeneity in matching as well as country productivity could furthermore permit more precise estimates of both trade creation and diversion terms.

References

Anderson, J. E., "A Theoretical Foundation for the Gravity Equation", *American Economic Review*, 69, 106-116, 1979.

Bandyopadhyay, S., Coughlin C.C., and Wall H.J., "Ethnic Networks and US Exports", *Review of International Economics,* 16(1), 199-213, 2008.

Bergstrand, J. H., "The Hecksher-Ohlin-Samuelson Model, the Linder Hypothesis and the Determinants of Intra-industry Trade", *Economic Journal*, 100, 1216-1229, 1990.

Blanes, J. V., "Does Immigration Help to Explain Intra-Industry Trade? Evidence for Spain", *Review of World Economics*, 141(2), 244-270, 2005.

Casella, A. and Rauch, J. E., "Overcoming Informational Barriers to International Resource Allocation: Prices and Ties", *Economic Journal*, 113, 21-42, 2003.

Chiswick, R. B. and Miller, W. P., "The Endogeneity Between Language and Earnings: International Analyses", *Journal of Labor Economics*, 13(2), 247-287, 1995.

Co, C.Y., Euzent P., and Martin, T., "The Export Effect of Immigration into the USA", *Applied Economics,* 36(6), 573-583, 2004.

Combes, P. P., Lafourcade M., and Mayer, T., "The Trade Creating Effects of Business and Social Networks: Evidence from France", *Journal of International Economics*, 66, 1-29, 2005.

Deardorf, A. V., "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?", in J.A. Frankel (Ed.), *The Regionalization of the World Economy*, Chicago: University of Chicago Press, 1998.

Dolman, B., "Patterns of Migration, Trade and Foreign Direct Investment across OECD Countries", *DEGIT Conference Papers* c012_030, DEGIT, Dynamics, Economic Growth, and International Trade, 2007.

Donald, G. S. and Lang, K., "Inference with Difference-in-Differences and Other Panel Data", *Review of Economics and Statistics*, 89(2), 221-233, 2007.

Dunlevy, J. A, "The Influence of Corruption and Language on the Protrade Effect of Immigrants: Evidence from the American States", *Review of Economics and Statistics,* 88(1), 182–186, 2006.

Felbermayr, G. J. and Toubal, F., "Revisiting the Trade-Migration Nexus: Evidence from New OECD Data", available at SSRN: http://ssrn.com/abstract=1144063, 2008.

Friedberg, M. R., "You Can't Take It with You: Immigrant Assimilation and the Portability of Human Capital", *Journal of Labor Economics*, 18(2), 221-251, 2000.

Gereffi, G., "International Trade and Industrial Upgrading in the Apparel Commodity Chain", *Journal of International Economics*, 48, 37-70, 1999.

Girma, S. and Yu, Z., "The Link Between Immigration and Trade: Evidence from the U.K.", *Weltwirtschaftliches Archiv*, 138, 115-130, 2002.

Gould, D. M., "Immigrant Links to the Home Country: Empirical Implications for U.S. Bilateral Trade Flows", *Review of Economics and Statistics*, 76, 302-316, 1993.

Greif, A., "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition", *American Economic Review*, 83, 525-548, 1994.

Head, K. and Ries, J., "Immigration and Trade Creation: Econometric Evidence from Canada", *Canadian Journal of Economics*, 31, 47-62, 1998.

Helliwell, J. F., "National Borders, Trade and Migration", *Pacific Economic Review*, 3(3), 165-185, 1997.

Helpman, E. and Krugman, P., *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition and the International Economy*, Cambridge, MA: MIT Press, 1985.

Herander, G. M. and Saavedra, A. L., "Exports and the Structure of Immigrant-Based Networks: The Role of Geographic Proximity", *The Review of Economics and Statistics*, 87, 323-335, 2005.

Javorcik, B. S., Ozden, O., Spatareanu, M. and Neagu, C., "Migrant Networks and Foreign Direct Investment", manuscript, 2006.

Konečný, T., "Can Immigrants Hurt Trade?", *CERGE-EI Working Paper*, 329, 2007.

Liang, Y. K. and Zeger, L. S., "Longitudinal Data Analysis Using Generalized Linear Models", *Biometrika*, 73, 13-22, 1986.

Mayda, A. M., "International Migration: A Panel Data Analysis of Economic and Non-Economic Determinants", *IZA Discussion Paper 1590*, 2005.

McCallum, J., "National Borders Matter: Canada-U.S. Regional Trade Patterns", *American Economic Review*, 85, 615-623, 1995.

McKenzie, D. J., "Paper Walls are Easier to Tear Down: Passport Costs and Legal Barriers to Emigration", *World Bank Policy Research Working Paper* 3783, 2005.

Moulton, R. B., "Random Group Effects and the Precision of Regression Estimates", *Journal of Econometrics*, 32, 385-397, 1986

OECD, *Trends in international migration,* OECD Publishing, 2004.

Pedersen, P.J., Pytliková, N. and Smith, N., "Selection or Network Effects? Migration Flows into 27 OECD Countries, 1990-2000", *IZA Discussion Paper*, 2004.

Peng, M. W., *Behind the Success and Failure of U.S. Export Intermediaries*, Westport, CT: Quorum Books, 1998.

Piperakis, A.S., Milner, C. and Wright, P.W., "Immigration, Trade Costs and Trade: Gravity Evidence from Greece", *Journal of Economic Integration*, 18, 750-762, 2003.

Rauch, J. E. and Trindade, V., "Ethnic Chinese Networks in International Trade", *Review of Economics and Statistics*, 84, 116-130, 2002.

Rauch, J. E. and Trindade, V., "Information, International Substituability, and Globalization", *American Economic Review*, 93, 775-791, 2003.

Rauch, J. E. and Watson, J., "Starting Small in an Unfamiliar Environment", *International Journal of Industrial Organization*, 21, 1021-1042, 2003.

Tadesse, B. and White, R., "Does Cultural Distance Hinder Trade? A Comparative Study of Nine OECD Member Nations", forthcoming in *Open Economies Review*, 2007.

Tadesse, B. and White, R., "Cultural distance as a determinant of bilateral trade flows: Do immigrants counter the effect of cultural differences?", *Applied Economics Letters*, 1, 1-6, 2008a.

Tadesse, B. and White, R., "Do Immigrants Counter the Effect of Cultural Distance on Trade? Evidence from US State-level Exports", Journal of Socio-Economics 37(6), 2304-2318, 2008b.

Tadesse, B. and White, R., "Do Immigrants Enhance International Trade in Services? The Case of US Tourism Services Exports", Working paper, 2009.

Wagner, D., Head, K. and Ries, J., "Immigration and Trade of Provinces", *Scottish Journal of Political Economy*, 49(5), 507-525, 2002.

White, R., "An Examination of the Danish Immigrant-Trade Link", *International Migration*, 45(5), 61-86, 2007a.

White, R., "Immigrant-Trade Links, Transplanted Home-Bias and Network Effects", *Applied Economics*, 39(7), 839-852, 2007b.

White, R. and Tadesse, B., "Immigration Policy, Cultural Pluralism and Trade: Evidence from the White Australia Policy", *Pacific Economic Review*, 12(4), 489-509, 2007.

White, R. and Tadesse, B., "Cultural Distance and the US Immigrant-Trade Link", *The World Economy*, 31(8), 1078-1096, 2008a.

White, R. and Tadesse, B., "Immigrants, Cultural Distance and US State-Level Exports of Cultural Products", *North American Journal of Economics and Finance* 19(3), 331-348, 2008b.

8 Appendix

Table A.1: Net trade effect of a 10-percent increase in total immigrant stock.

OECD host countries	Net trade effect (in %)	Source countries	Net trade effect (in %)	Source countries	Net trade effect (in %)
Austria	-0,12	Chad	0,57	Mauritius	-0,6
Belgium	0,74	Chile	0,18	Morocco	0,38
Canada	0,97	China	0,43	Mozambique	2,58
Denmark	0,48	Colombia	0,15	Nepal	0,47
Finland	0,71	Congo	1,24	Nicaragua	0,24
France	1,01	Costa Rica	0,06	Niger	0,52
Germany	0,53	CoteD'Ivoire	0,57	Nigeria	0,47
Greece	0,87	Cyprus	-3,05	Oman	0,43
Ireland	0,76	Dem.Rep.Congo	-0,12	Pakistan	0,42
Italy	0,48	Dominican Rep.	-0,93	Panama	-0,67
Japan	0,89	Ecuador	-0,22	Papua N.Guinea	0,45
Korea	0,84	Egypt	0,37	Paraguay	0,39
Netherlands	0,57	El Salvador	-1,76	Peru	0,15
Norway	0,62	Eq.Guinea	0,09	Philippines	0,06
Portugal	1,18	Ethiopia	0,57	Qatar	0,33
Spain	0,7	Fiji	-0,95	Romania	0,07
Sweden	0,73	Gabon	0,38	Rwanda	0,87
UK	0,62	Ghana	0,69	Saudi Arabia	0,42
USA	0,62	Guatemala	-0,19	Senegal	1,17
Source countries	Net trade effect (in %)	Guinea	0,51	Seychelles	-1,24
		Haiti	1,06	South Africa	0,33
Albania	-3,49	Honduras	-0,55	Sri Lanka	0,2
Algeria	0,47	Indonesia	0,44	Sudan	0,47
Angola	0,12	Iran	0,28	Syria	0,31
Argentina	0,27	Israel	-0,16	Tanzania	0,52
Bahrain	0,22	Jamaica	-5,73	Thailand	0,36
Bangladesh	0,45	Jordan	0,16	Togo	0,73
Barbados	-6,99	Kenya	0,53	Trinidad and Tbg	-4,31
Belize	-2,69	Kuwait	0,07	Tunisia	0,16
Benin	0,63	Lao P.D.R	0,13	Uganda	0,63
Bolivia	0,34	Lebanon	-1,15	UAE	0,36
Brazil	0,39	Madagascar	1,52	Uruguay	0,02
Bulgaria	0,11	Malawi	0,56	Venezuela	0,24
Burkina Faso	0,56	Malaysia	0,33	Vietnam	0,19
Burundi	1,19	Mali	1,19	Yemen	0,44
Cambodia	0,51	Malta	-2,6	Zambia	0,56
Cameroon	0,56	Mauritania	0,74	Zimbabwe	0,46

Table A.2: Regression results, dependent variable real exports and imports 1999-2003.

	Real exports 1999-2003			Real imports 1999-2003		
	(1) OLS regional dummies i and j	(2) 2-step estimates for source i	(3) 2-step estimates for host j	(4) OLS regional dummies i and j	(5) 2-step estimates for source i	(6) 2-step estimates for host j
$\ln(GDP_i GDP_j)$	1.105***	0.783***	0.783***	0.996***	0.754***	0.754***
	(0.03)	(0.020)	(0.020)	(0.031)	(0.021)	(0.021)
$\ln(GDPcap_i GDPcap_j)$	-0.047	-0.203***	-0.203***	-0.013	-0.183***	-0.183***
	(0.044)	(0.035)	(0.035)	(0.044)	(0.036)	(0.036)
Source diversion$_i$	-0.117	-2.486***	-	0.073	-2.318**	-
	(0.476)	(0.939)		(0.432)	(0.973)	
Host diversion$_j$	-9.099	-	-6.911	-7.254	-	-4.929
	(6.036)		(4.105)	(5.690)		(11.235)
Trade creation$_{ij}$	0.085***	0.056***	0.056***	0.090***	0.044**	0.044**
	(0.021)	(0.017)	(0.017)	(0.024)	(0.018)	(0.018)
Trade creation$_{ij}^2$	-0.184***	-0.001	-0.001	-0.001*	-0.001	-0.001
	(0.056)	(0.001)	(0.001)	(0.0004)	(0.001)	(0.001)
Distance$_{ij}$	-0.669***	-1.080***	-1.080***	-0.779***	-1.153***	-1.153***
	(0.068)	(0.047)	(0.047)	(0.079)	(0.049)	(0.049)
Colony$_{ij}$	7.234**	5.133***	5.133***	7.211**	5.882***	5.882***
	(3.202)	(1.051)	(1.051)	(1.580)	(1.110)	(1.110)
Language$_{ij}$	4.565***	5.591***	5.591***	5.019***	5.081***	5.081***
	(1.455)	(0.725)	(0.725)	(1.742)	(0.766)	(0.766)
HF Index host$_j$	-0.024*	-	-0.015	-0.030	-	0.001
	(0.014)		(0.019)	(0.027)		(0.052)
HF Index source$_i$	0.012***	0.030***	-	0.013***	0.030***	-
	(0.002)	(0.004)		(0.002)	(0.004)	
Export share host$_j$	1.647***	-	2.773***	1.651***	-	0.601
	(0.351)		(0.233)	(0.295)		(0.688)
Export share$_i$	2.586***	1.786***	-	1.730***	1.246***	-
	(0.215)	(0.323)		(0.130)	(0.335)	
Constant	-10.016***	2.046	1.82	-11.505***	-2.320***	0.463
	(0.625)	(2.53)	(2.521)	(2.341)	(0.152)	(4.221)
R^2	0.852	0.670	0.595	0.856	0.496	0.600
N	1,577	1,684	1,684	1,577	1,684	1,684

Notes: The OLS with regional dummies account for clustering by host countries.
***,**,* - Significant at 1%, 5%, and 10% respectively. Standard errors in parentheses.

Figure A.1: A 10% increase in total immigrant stock and Host j's trade with source countries. Estimates accounting for relatively large immigrant communities (*Trade creation*$_{ij} > 1$, see Table 5).

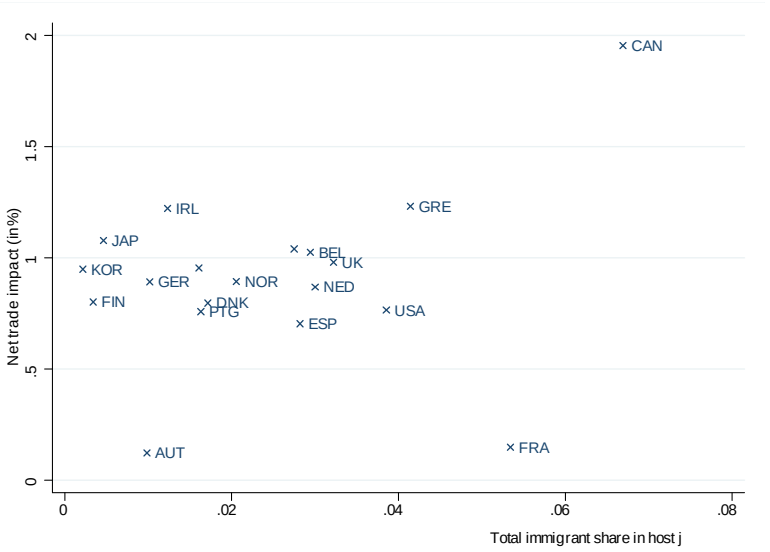


Figure A.2: A 10% increase in total immigrant stock and Source i's trade with OECD countries. Estimates accounting for relatively large immigrant communities (*Trade creation*$_{ij} > 1$, see Table 5).

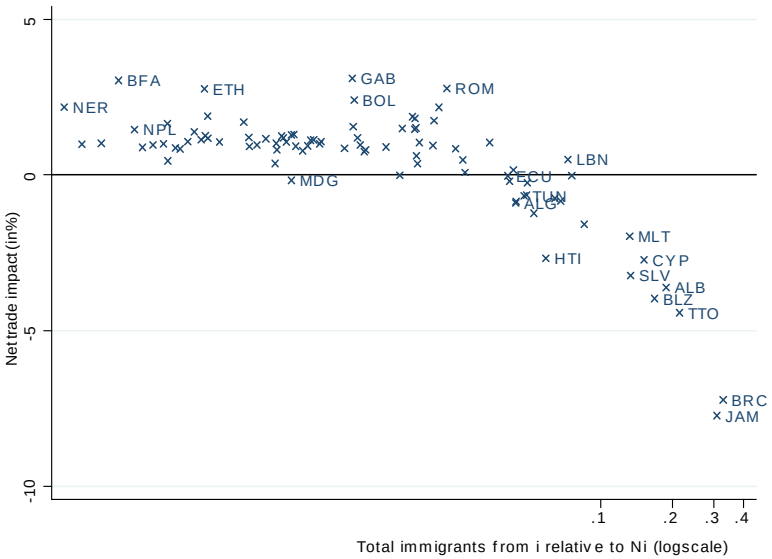


Figure A.3: Marginal trade creation effects of a 10-percent increase in m_{ij}, model predictions accounting for relatively large immigrant communities (*Trade creation*$_{ij} > 1$, see Table 5).

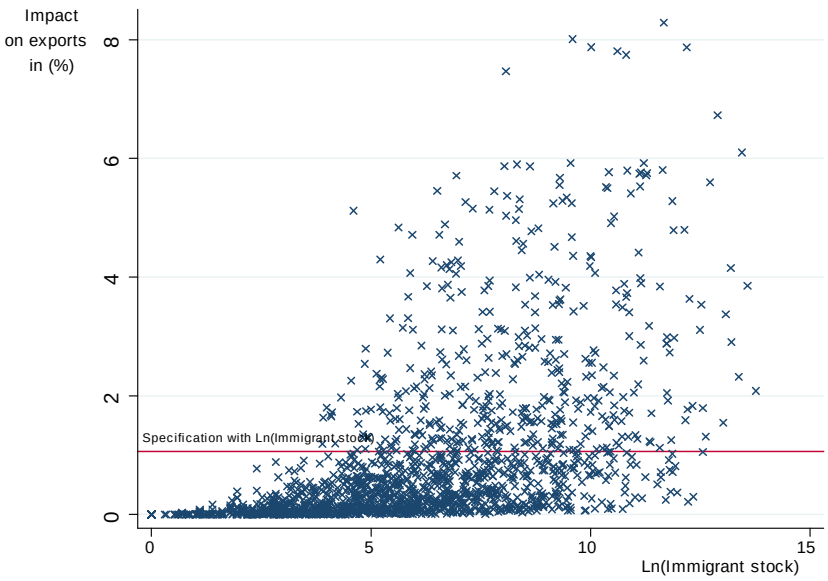


Table A.3: Correlation table.

	real exports$_{ij}$	real imports$_{ij}$	Host GDP$_j$	Source GDP$_i$	Host GDPcap$_j$	Source GDPcap$_i$	Imm. stock$_{ij}$	Trade creation$_{ij}$	Host diversion$_j$	Source diversion$_i$
exports$_{ij}$	1									
imports$_{ij}$	0,79	1								
Host GDP$_j$	0,54	0,51	1							
Source GDP$_i$	0,37	0,39	0,001	1						
Host GDPcap$_j$	0,19	0,06	0,37	0	1					
Source GDPcap$_i$	0,24	0,08	0	0,04	0	1				
Immigrant stock$_{ij}$	0,27	0,26	0,26	0,15	0,11	-0,05	1			
Trade creation$_{ij}$	0,10	0,11	0,14	-0,08	0,08	-0,1	0,39	1		
Host diversion$_j$	0,07	0,10	0,19	0,001	0,02	0,001	0,16	0,27	1	
Source diversion$_i$	-0,09	-0,14	0,01	-0,14	0,001	0,14	0,03	0,1	-0,01	1