CAN IMMIGRANTS HURT TRADE?

Tomáš Konečný

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

WORKING PAPER SERIES (ISSN 1211-3298)
Electronic Version

329

Working Paper Series **329**
(ISSN 1211-3298)

Can Immigrants Hurt Trade?

Tomáš Konečný

CERGE-EI
Prague, May 2007

ISBN 978-80-7343-128-0 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-117-3 (Národohospodářský ústav AV ČR, v.v.i.)

Can Immigrants Hurt Trade?

Tomáš Konečný[*]

CERGE-EI[†]

May, 2007

Abstract

This paper estimates the impact of immigrant network spillovers on international trade. Contrary to previous studies focusing mostly on the trade enhancing role of immigrant networks, the present framework allows for potential trade *diverting* effects. A simple matching model that incorporates *both* trade creation and diversion channels furthermore points at the importance of relative as opposed to absolute measures of immigrant networks. Using a new dataset of 19 OECD countries, I find that while immigrant networks indeed facilitate exports from host to source country, they simultaneously hurt trade with the host country's other trading partners. In addition, I find that the impact of information-related trade barriers might be negatively related to the economic size of the trading partner. In particular, the larger the trading partner, the smaller the trade benefits of lower information costs due to a shared common language.

Abstrakt

V článku odhaduji vliv vazeb imigrantů na domovskou zemi v mezinárodním obchodě. Narozdíl od předchozích studií, které zkoumaly pouze jejich potenciální přínos, se zaměřuji i na možné náklady ve formě odklonu obchodních toků. Součástí práce je jednoduchý model zachycující jak přínosy, tak i náklady imigrantských vazeb. Model dále poukazuje na důležitost relativních ukazatelů množství imigrantů, které byly v dosavadní literatuře opomíjeny. V empirické části používám nová data zahrnující 19 hostitelských zemí v rámci sdružení OECD. Výsledky naznačují, že ačkoliv imigranti rozvíjejí obchod s domovskou zemí, mají zároveň negativní vliv na obchod hostitelské země s ostatními obchodními partnery. V práci také ukazuji, že dopad informačních bariér na mezinárodní obchod je inverzní funkcí velikosti obchodního partnera.

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. 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 and Krešimir Žigić. The usual disclaimer applies.

[*] CERGE-EI, Politických vězňů 7, Prague, 111 21, Czech Republic. Email: tomas.konecny@cerge-ei.cz

[†] CERGE-EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic

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 the lack of information on available trading opportunities and imperfect contract enforcement. In the former case, the 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). In the second case, given the existence of insufficient contract enforcement institutions when trade parties originate from different jurisdictions, the potential renegation from the contract and the following loss 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 dispose of some sort of collective punishment mechanism that could substitute for inadequate enforcement institutions. In particular, numerous studies on informal barriers examine the impact on the 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 the host and source countries.

However, the combination of pervasive informal trade barriers and country-specific knowledge possessed by immigrants can also lead to the reduction of trade. Consider a German tractor producer who wishes to export its engines to either Vietnam or Thailand. Finding a reliable trade partner with the capacity to provide local distribution, marketing or maintenance services can be costly both in terms of money and time. Other things 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] While from the perspective of Germany the total exports do not change (or they even increase 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.

Gereffi's study on offshoring in the apparel industry provides yet another illustration (Gereffi, 1999). The author mentions 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. Gereffi argues that it is these networks that seem to shape many investment decisions.[2] In his example, the trade diversion would

[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. intermediaries' officers or managers are 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

take the form of unrealized offshoring projects in countries such as Bangladesh or Sri Lanka, i.e., in destinations with very low wages but lacking links to Chinese networks.

Since the previous studies did not address the potential trade-diverting mechanism driven by immigrant links, the present paper aims to fill the existing gap in the literature. It develops a simple framework for the study of network effects on trade and derives theory-based expressions for both trade creation and trade diversion by networks. Furthermore, the paper evaluates the empirical relevance of the trade-diversion channel using a new dataset on foreign-born populations located in 20 OECD member countries.

The following section reviews the spare theoretical literature on the trade-diversion effects of network ties. Section 3 covers the existing empirical research on the role of immigrant links in international trade and presents the model, while Section 4 discusses the data employed. The subsequent section covers econometric issues, results and sensitivity analysis. Section 6 concludes.

2 Theoretical literature

Several theoretical studies formalize the trade consequences of various aspects of asymmetric information. Casella and Rauch (2003) focus on the relationship between network ties and international prices. Rauch and Trindade (2003) study the impact of decreasing information costs on trade volumes, prices and welfare. Rauch and Watson (2003) analyze the mode of match formation in an environment with positive search costs and uncertainty relating to the trade partner's type.[3] Of these studies, Casella and Rauch (2003) is the only one that

ethnic Chinese networks reaches nearly 60%.

[3] A firm can either engage in a small joint project in order to learn more about its trade partner's quality, or it can place a large scale order right away and risk a higher probability of failure, or it can reject the matched partner and resume searching.

mentions potential trade-diversion effects driven by network ties.

The basic setup employed by Casella and Rauch (2003) is a 2 country x 1 good x 2 factor model in which the match quality represents the second factor of production besides the internationally immobile labor. The authors assume producers consisting of different types, each type summing to mass 1 located on a unit circle. Each producer can form a joint venture of the best possible quality with either a domestic partner or a foreign party. The quality of the joint venture is measured by the distance between any two producers and it directly impacts gains from a successful match.

A crucial ingredient in the model is network ties. While producers in the Home (labor-scarce) country possess complete information regarding potential matches in the domestic market, they do not know anything about their foreign counterparts and hence have to search randomly. Only a subgroup of domestic producers tied in a network benefit from full information on the available foreign matches. The model thus de facto introduces a matching technology that reduces existing frictions, yet this technology is reserved for a subgroup of producers only.

The equilibrium solution depends on the difference between the countries' labor endowment ratios.[4] In the case that the labor endowments happen to be sufficiently close, matches between domestic and foreign producers do not occur. The implied wage differential is simply too small to outweigh the uncertainty of matching with a foreign partner. As the labor endowment ratio and hence relative wages fall, tied home producers become active and with additional decreases a fraction of untied producers engage in further trade. This breakdown due to information ties leads to a partial insulation of one country from another, meaning

[4] Remember that the number of producers located in each country is set to one.

that the country's wages tend to be more sensitive to changes in the domestic rather than foreign labor supply and wage convergence becomes weaker. The introduction of ties also raises income and hence aggregate welfare relative to the benchmark case with pervasive uncertainty.

The positive welfare contribution of network ties in a two-country model could, however, become reduced or even reversed once a third country is added. Casella and Rauch (2003) argue that in the case when the strongest ties do not persist between the countries with the largest wage differentials in the absence of ties, trade creation by network ties will be mitigated by the previously nonexistent trade diversion channel. The strength of the trade diversion will depend on the size of the wage differentials between the two foreign countries. Compared to the baseline two-country solution, the closer the wages of a foreign country with network ties in comparison to the foreign country with the lowest relative wage, the smaller the trade diversion effect and the larger the trade creation effect (arising from transfer of labor to a cheaper country) would be. Large wage differentials between foreign countries will, on the contrary, magnify the trade diversion effect. Finally, if the wage differentials become wide enough, ties-induced trade diversion will be removed as producers will avoid the foreign country with network ties altogether and all international matches will be formed with the now substantially cheaper trading partner.

Unfortunately, Casella and Rauch (2003) do not present analytical expressions for the trade diversion channel. While for the three-country case such expressions could be derived manually, it remains unclear as to how the model would extend to a multicountry setting. In particular, the extension to three countries would not provide any information on the proper aggregation of wage differentials and ties to other trading partners (i.e. of the "third"

country). In order to avoid the complexities inherent in Casella and Rauch (2003), the simple multicountry framework developed in the next section allows for the trade diversion channel without any reference to relative wages.

3 Empirical studies and model

3.1 Previous empirical research on immigrant ties

The scarcity of theoretical studies on trade-diversion effects of immigrant networks is further supplemented by the absence of any empirical work in this area. The earlier pioneering studies focused on whether immigrant ties influence trade patterns at all and if so, their sole focus has been on trade creation. Nonetheless, the existing output still provides a useful consistency check, at least for the trade creation estimates discussed in Section 5.1.2.

Gould (1993) analyzes migration inflows in the U.S. using panel data from 1970 to 1986 and finds a significantly positive impact of immigrants on bilateral trade with their source countries. The implied long-run elasticities suggest a 10-percent increase in immigrant stock to increase U.S. exports to a concerned source country by 4.7 percent and U.S. imports by 8.3 percent. Gould (1993) explains the difference in estimates by the presence of both a taste effect and immigrant network effect in the imports equation, whereas for exports the taste effect should be absent. A similar exercise for the Canadian economy has been produced by Head and Ries (1998). The authors use two different measures of immigrant ties, 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 increase Canadian bilateral exports by 1.0-1.3 percent and imports by 3.1-3.9 percent. Rauch and Trindade (2003)'s study on ethnic Chinese networks delivers both an economically and

statistically significant role of Chinese minorities in bilateral trade promotion. Moreover, the authors conclude that while for Southeast Asian countries with larger population shares trade facilitation occurs mainly through the alleviation of information asymmetry (related mostly to trade partner characteristics), for other countries where ethnic Chinese form a rather small fraction of the population, the substitution for weak legal institutions becomes a key factor. Interesting research on the relevance of immigrant ties for the United Kingdom's trade has been conducted by Girma and Yu (2002). Girma and Yu (2002) work with the terms "universal" and "non-universal" immigrant ties, the former term being a label for personal contacts broadly consistent with Casella and Rauch (2003)'s matching frictions argument, the latter serving as a proxy for the general knowledge of source country's social institutions and market specifics. The operating hypothesis states that while universal links should matter equally for all countries, non-universal links should matter more in source countries with a more dissimilar institutional (legal, social, cultural) setup. Exploiting the different historical experience of Commonwealth and non-Commonwealth countries, the paper finds robust evidence that the impact of immigrant ties from Commonwealth countries on the UK's exports is insignificant. The results thus favor the "non-universal" against the "universal" dimension of immigrant ties.

More recent studies include Combes et al. (2005) and Herander and Saavedra (2005). The careful econometric study by Combes et al. (2005) analyzes the impact of internal migration in France on inter-regional trade. Migrants are found to double the volume of inter-regional trade flows in comparison to the situation without migrant networks. The estimated impact of business ties proves to be even stronger, magnifying trade flows by five times in some specifications. Furthermore, the authors quantify the share of informal

trade barriers that can be explained by migrant and business networks. According to their results, the negative impact of transport costs falls by as much as 60 percent and the trade-restricting border effect drops by more than three times. Finally, Herander and Saavedra (2005) explore the spatial dimension of immigrant networks. Focusing on trade-creation effects of immigrant networks operating within and between the U.S. states, 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 research by Herander and Saavedra (2005) is to my knowledge the only one to consider trade spillovers by immigrant networks. However, 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 of *another* ethnicity within a given country. The following subsection presents the estimation framework.

3.2 The empirical model

For the empirical evaluation of the trade diversion hypothesis I use a simple gravity framework that explicitly allows for matching in trade. The gravity relationship proportionally linking trade flows to the incomes of trading economies can be derived from a wide range

of international trade models.[5] The present model, however, does not build on any of them. Instead, it shows that the gravity might be consistent even with a very simple world economy with matching and no differences in productivity, endowments, or preferences.

Index a source country from which the immigrants come $i = 1, .., I$ and their host country $j = 1, .., J$. Assuming no differences in productivity, a country j's share in world output equals its population share in total population, i.e.

$$\frac{N_j}{N} = \frac{GDP_j}{GDP}, \tag{1}$$

where N_j stands for the total population in country j and N stands for the total world population. Further assume that the total N_j population in economy j consists of natives and immigrants from other countries i, where native agents are poorly informed with regard to available matches and immigrants from country i (denoted by m_{ij}) are fully informed. That is, assume native agents j could trade only with agents who were not matched by fully connected immigrants. Similarly, assume the immigrants possess full information on potential matches in their source country i, yet no information at all on matches relating to any other trading partners, so that in the end they would trade only with their source country i.

Figure 1 outlines the basic structure of the model with the two source (1 and 2) and two host countries (3 and 4). The picture shows that migrants from 1 residing in host countries 3 and 4 will match directly with an amount $\sum_{j=3,4} m_{1j}$ out of the total population N_1 located in the source country 1. In the host economy 4, on the other hand, $N_4 - \sum_{i=1,2} m_{i4}$ agents are poorly informed and thus face a random choice between other unmatched participants.

[5] Examples include Anderson (1979), Bergstrand (1990), Deardorff (1998) and Helpman and Krugman (1985).

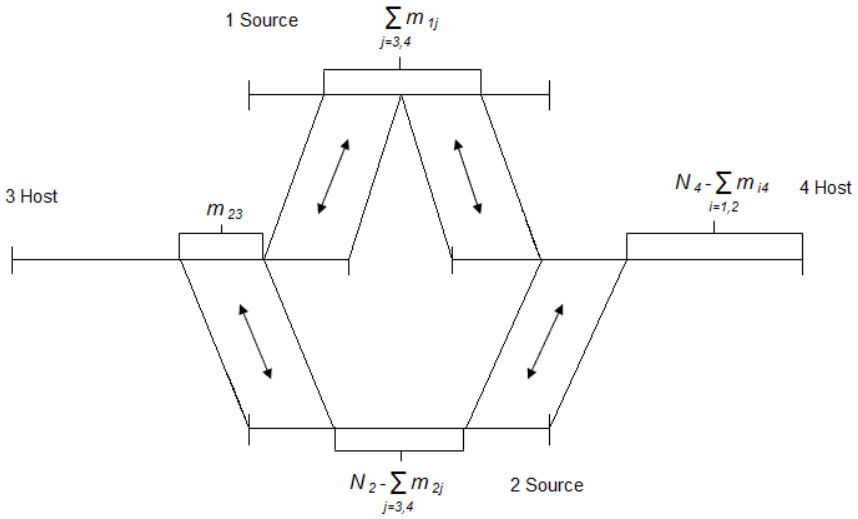

Figure 1: Model structure

Formally, the production generated by poorly informed agents in a host country j equals

$$GDP_{j.noIMM} = \left[\frac{\left(N_j - \sum_{i=1}^{I} m_{ij} \right)}{N} \right] GDP, \tag{2}$$

where $\sum_{i=1}^{I} m_{ij}$ represents the total number of immigrants in a host country j. Exports to a source country i by poorly connected agents in j will be proportional to $\left(N_i - \sum_{j=1}^{J} m_{ij} \right) / N$, the term $\sum_{j=1}^{J} m_{ij}$ accounting for the total number of immigrants from a source country i residing in all host countries $j = 1, .., J$ since these immigrants would directly benefit from network ties and would thus form matches only with residents in their source country. Thus, using (1), the exports to i generated by poorly connected agents in j will equal[6]

$$X_{ij}(poorly\ informed) = \frac{GDP_i GDP_j}{GDP} \left(1 - \frac{\sum_{j=1}^{J} m_{ij}}{N_i} \right) \left(1 - \frac{\sum_{i=1}^{I} m_{ij}}{N_j} \right) \tag{3}$$

and the exports to i generated by the immigrant group from i residing in j is

$$X_{ij}(immigrant\ network) = \frac{m_{ij}}{N} GDP = \frac{m_{ij}}{N_j} GDP_j,$$

[6] In principle, it might happen that the middle term in brackets and hence predicted exports can both turn negative. The situation corresponds to a hypothetical country with its overseas diaspora larger than the country's domestic population. In the present sample however, all observations are positive. I would like to thank Libor Dušek for this remark.

where use was made of (1) and m_{ij} equals the immigrant population from i residing in j. Summing the last two expressions, one obtains the relationship for total exports, which can be estimated accordingly:

$$X_{ij} = \frac{GDP_i GDP_j}{GDP} \left[\left(1 - \frac{\sum_{j=1}^{J} m_{ij}}{N_i} \right) \left(1 - \frac{\sum_{i=1}^{I} m_{ij}}{N_j} \right) + \frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP}} \right]. \tag{4}$$

Taking logarithms and approximating $\ln(1+x) \sim x$ for x small, one obtains

$$\ln X_{ij} = b_0 + b_1 \ln GDP_i GDP_j + b_2 \frac{\sum_{j=1}^{J} m_{ij}}{N_i} + b_3 \frac{\sum_{i=1}^{I} m_{ij}}{N_j} + b_4 \frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP}}. \tag{5}$$

The coefficients b_2 and b_3 indicate the size of trade diversion due to immigrant networks and are expected to be negative and both equal to minus one. Since b_2 refers to the population of a source country N_i and b_3 refers to the population of a host country N_j, I call the terms corresponding to these coefficients source country and host country trade diversion terms, respectively. The preceding discussion suggests that the host country diversion term should be of particular interest, since it is the explicit assumption that immigrant networks trade only with their source countries while disregarding other potentially profitable matches that drives the model. b_4 captures trade creation by immigrant network ties and is expected to be positive and equal to one.[7]

For the trade creation term not small, I employ a quadratic approximation with a negative expected sign. Intuition says that as $\frac{m_{ij}}{N_j}$ gets large, the migrants will tend to trade with each other instead of trading with their country of origin.

[7] While the immigrant ties introduced by the present matching model shift the geographical pattern of trade, they do 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 networks, the total trade effect would however turn positive despite the presence of trade diversion.

For estimation purposes, I use the following version of (5):

$$\ln X_{ij} = b_0 + b_1 \ln GDP_i GDP_j + b_2 \frac{\sum_{j=1}^{J} m_{ij}}{N_i} + b_3 \frac{\sum_{i=1}^{I} m_{ij}}{N_j} + b_4 \frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP}} +$$

$$+ b_5 \left(\frac{\frac{m_{ij}}{N_j}}{\frac{GDP_i}{GDP}} \right)^2 + a'z_{ij} + e_{ij}, \tag{6}$$

where a and z_{ij} are k x 1 vectors of other explanatory variables. These include distance, the product of GDP's per capita (both expressed in natural logarithms), membership in the EU, export share in the GDP of the host country as a proxy for export infrastructure, institutional quality measures and dummies for shared border, shared colonial past and common language.

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

Unlike other existing studies, however, I divide the colony and language dummy variables by the GDP of a source country i. The resulting continuous variables then imply a larger trade facilitating impact for smaller economies. The reasoning in general concords to Anderson and van Wincoop (2001), who focused on the "missing" trade between the U.S. states and Canadian provinces previously documented by McCallum (1995). 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 economy tends to be relatively more sensitive to trade barriers.

Finally, the formulation of the trade creation variable differs from the frequently used natural logarithm of immigrant stock.[8] Despite the lack of theoretical justification and zero predicted trade in the absence of immigrant networks, the natural logarithm formulation remains intuitively appealing and easy to interpret. On the other hand, the ratios derived within the present framework rely on an explicit model and emphasize relative rather than absolute measures of immigrant networks. In order to obtain an idea of the respective importance of ratios versus levels, I complement the formally derived relative measures with the natural logarithm of immigrant stock.

4 Data

4.1 Immigrants

One of the key reasons why the existing empirical studies on immigrant ties could not focus on trade diversion effects was their sole focus on a single host country (Rauch and Trindade, 2003 being the only exception). As a result, the trade diversion terms developed in the previous subsection simply could not be estimated, since they require variability at both the source *and* host country levels. I use the cross-country information on the stocks of foreign-born persons over 15 years of age for 20 OECD member countries. These data were obtained from the OECD Statistics Portal on Demography and Population. The main

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

possible reservation relates to the asserted comparability across countries. The principle sources for the above-mentioned database consist of population registers, residence or work permits, censuses and surveys, i.e. records which do not have migration statistics as a primary purpose. Furthermore, the frequency of censuses is usually 5 or 10 years; as a result, the reference year varies between 1999 and 2002, depending on the specific country. On the other hand, the OECD output represents the first attempt to create a coherent dataset covering more than a single host country.[9]

Some OECD countries had to be dropped due to large proportions of foreign-born population with the unknown country of origin.[10] For the remaining hosts, the small unknown populations have been distributed using country-of-origin shares in the total number of foreign born in a concerned host country. Furthermore, several source countries had to be dropped due to political changes that occured after the fall of the Soviet Bloc. Especially in the case of the Balkans and the descendants of the former Soviet Union, the reporting of the new entities differs among host countries.

Another reservation relates to the figures for Germany, which were listed only by broad source regions instead of countries, and for the Netherlands where the data included only the number of all foreign born instead of those over 15 years of age. For that reason 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-

[9] Further information on data sources and adjustments can be found on the OECD's web site at http://www.oecd.org/document/2/0,2340,en _2825_494553_38060354_1_1_1_1,00.html.

[10] These include Australia (16.2% unknown), Czech Republic (28.2%), Mexico (41.9%), New Zealand (16.1%), Poland (41.1%), Slovak Republic (9.3%), and Switzerland (14.7%). The borderline cases, Finland (3.8%) and Denmark (6.7%) were left in the sample. For other host countries the values of unknown foreign-born did not exceed 2%.

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 5, I drop the two host countries and run all regressions again to check for the robustness of results.

The main results presented in Section 5 cover 19 advanced industrial economies as immigrant hosts. The characteristics of the last OECD member in the sample, Turkey, are in many cases closer to a typical developing country and its membership in the OECD owes more to strategic considerations rather than the level of economic development. Nonetheless, despite being left out from the main regressions, the regression results with Turkey as a host country are shown in Appendix A (see Table A2).[11]

The last remark on immigrant data concerns 31 pairs (1.5 percent out of 2,071 observations) consisting either of very poor former colonies supplying huge immigrant populations and their centers (e.g. Suriname and the Netherlands, Congo and Belgium, Angola and Portugal), or small island economies with a disproportionate representation in the host countries (Fiji and Canada, Haiti and the U.S.A). These pairs were dropped due to extremely high trade creation values with a mean equal to 35.01 as compared to 0.26 for the rest of the sample. The sensitivity of the estimates to the inclusion of the outlying observations will be discussed in more detail in Section 5.3.2.

4.2 Trade and remaining data

[11] The estimates remain both quantitatively and statistically similar to those listed in Table 2.

Table 1: Summary statistics

Variable	Sample	Mean	Standard deviation	Minimum	Maximum
Exports	2,040	1,826	9,128	1	229,869
Host GDP*	2,040	1,169,211	2,113,976	77,758	9,012,500
Source GDP*	2,040	247,580	971,716	683	9,012,500
Host GDP/capita	2,040	21,780	7,743	9,307	36,720
Source GDP/capita	2,040	6,790	9,299	114	36,720
Immigrant stock	2,040	22,551	202,172	0	8,359,180
Trade creation	2,040	0.264	0.720	0	6.698
Host diversion	2,040	0.0601	0.0334	0.0030	0.1416
Source diversion	2,040	0.0420	0.0755	0.0006	0.4388
Distance	2,040	6,761	3,901	174	19,868
Export share	2,040	0.42	0.22	0.11	0.98
Institutions	2,040	22.74	18.35	-22.76	65.39
Colonial past	2,040	0.0025	0.0170	0	0.1532
Language	2,040	0.0058	0.0246	0	0.1532
Common border	2,040	0.0172	0.1298	0	1

*in millions of 1998 U.S. dollars

The data on bilateral exports have been obtained from the Direction of Trade Statistics compiled by the International Monetary Fund.[12] Since the precision of import figures tends to be generally higher than the precision of export figures, I use reported imports from host countries. Exports to especially smaller developing countries can vary substantially from year to year. For that reason a five-year average of real exports over 1999-2003 has been chosen, instead of using the data for a single year only. The averaging reduces an additional problem with zero observed exports. While 23 out of a total 2,160 observations (i.e. roughly 1 percent) reported no trade in at least one year over the 1999-2003 period, none of them did so for the whole five year period. Finally, since the focus of the present study are immigrant networks and the home links of overseas Chinese communities quite likely cover both China and Hong Kong, the two are treated as a single country.

The remaining variables, common language and a measure of circle distance between

[12] I employ exports rather than imports into host countries, since in the latter case the trade effect of the immigrant network tends to be further reinforced by trade driven by immigrants' preferences for source country products. The resulting estimates would then capture several channels that would be hard to disentangle. However, the estimates using import data will be provided soon, in order to obtain yet another consistency check of the main results.

capital cities were retrieved from Jon Haveman's web page[13] and added manually if values were missing. A dummy for a common colonial past used information from *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 difference between the five-year averages of the restricted Index of Economic Freedom produced by the Heritage Foundation. The Index of Economic Freedom over 1999-2003 compiles evaluations of nine areas critical for a functioning market environment. The restricted version however includes only those areas that most closely relate to institutional quality – 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 have been collected from the World Development Indicators published by the World Bank. To avoid the endogeneity problem, I use GDP and GDP per capita figures from 1998 as proxies. The main sample contains 19 host countries and 109 source countries, generating in total 2,040 observations (i.e., 19 x 109 - 31). Table (1) presents the summary statistics for key variables.

5 Empirical results

5.1 Trade diversion and creation terms

5.1.1 Trade diversion

The trade diversion expressions developed in Section 3.2. are host- and source-country specific and hence might interfere with the panel structure of the sample. First, any correlation

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

Table 2: Regression results, dependent variable Real exports 1999-2003.

Variable	(1) OLS	(2) OLS clustered	(3) Host FE	(4) Host RE	(5) Source FE clustered	(6) Source FE clust. + regions	(7) Wooldridge (2003)
Host diversion	-7.853***	-7.853***		-7.185***	-8.658***	-7.191***	-6.899**
	(0.601)	(2.457)		(0.251)	(2.138)	(2.807)	(2.920)
Source diversion	1.004***	1.004**	0.099	0.195			
	(0.270)	(0.419)	(0.254)	(0.254)			
Trade creation	0.448***	0.448***	0.417***	0.420***	0.419***	0.453***	0.360***
	(0.072)	(0.155)	(0.067)	(0.067)	(0.017)	(0.119)	(0.115)
Trade creation sq	-0.047***	-0.047*	-0.046***	-0.046***	-0.047**	-0.051**	-0.041**
	(0.014)	(0.024)	(0.014)	(0.014)	(0.022)	(0.020)	(0.019)
Colony	3.951***	3.951	2.466**	2.614**	3.115	2.278	3.014
	(1.263)	(4.038)	(1.241)	(1.240)	(3.805)	(3.488)	(2.098)
Language	3.812***	3.812**	4.268***	4.262***	3.188*	3.170*	4.351**
	(0.857)	(1.645)	(0.807)	(0.810)	(1.719)	(1.746)	(1.616)
#	2,040	2,040	2,040	2,040	2,040	2,040	2,040
R-squared	0.901	0.902	0.900	0.900	0.932	0.933	0.676

***, **, * - significant at 1%, 5% and 10%, respectively

with the respective unobservable country characteristics might introduce bias into the estimates. If, on the other hand, country dummies are included, the estimation of trade diversion effects becomes impossible. Another concern relates to the possible correlation between the residuals within the host country, which might lead to the underestimation of true standard errors. To address these issues, I proceed gradually and compare the coefficient estimates and standard errors coming from the baseline OLS specification in column (1) in Table 2 with those accounting for clustering as well as country and region effects. The benchmark OLS estimates have all the expected signs and economically reasonable values with the only exception of the source country trade diversion term. Column (2) repeats the OLS exercise with standard errors adjusted for clustering by host country. The almost fourfold increase of standard errors for the host trade diversion term in column (2) points at strong within-group correlation among residuals, yet keeps the basic result intact – larger immigrant shares in the host country population tend to reduce trade. The fourth column replaces fixed effects with random effects for the host country.[14] While these results are purely indicative, note that the trade diversion term lies within one standard error to the baseline OLS level and that other variables remain practically indistinguishable from the specification in column (3) and reasonably close to their corresponding OLS outcomes. Thus the OLS still seems to be a relatively good proxy.

The benchmark specification has until now ignored varying trade policies among the

[14] The validity of random effects hinges on the assumption of zero correlation between the regressors and residuals. Unfortunately, the Hausman test cannot resolve this issue as the calculated matrix $V(b) - V(B)$ is not positive definite. This could arise in finite samples due to sampling variance or the failure to meet the assumptions of the random effects model (i.e. the independence of regressors and residuals). The correlation between the residuals and regressors in the fixed effects specification however equals 0.008, so it is the former case that seems to matter. I decided to keep the estimates in the table nonetheless, since the use of the random effects specification was motivated by the possibility to directly estimate the trade diversion term, rather than higher efficiency.

source countries. In the absence of reliable trade policy measures for the 109 source countries included in the sample, the inclusion of source-country dummies becomes a necessity despite the loss of the source country diversion term. Column (5) delivers an increase in the level of the GDP-product term (complete results are reported in Table A2 in Appendix A). The host diversion term, however, remains within one standard deviation from its previous estimates. The last two columns report estimates that try to capture the specifics of both host and source countries. In column (6) I introduce regional dummies for host OECD countries[15] as well as source country fixed effects without any qualitative jumps in results. Finally in the last column, I use the two-step methodology employed in Wooldridge (2003) with first-stage fixed effects for both host and source countries. In the first stage, I estimate pair-specific variables, adjusting for clustering by the host country (the R-squared reported in Table 2, however, refers to the second-stage regression). In the second stage, I use the obtained coefficients on host-country fixed effects and use them as a left-hand-side variable with host country-specific terms as regressors. The host country-specific terms include the natural logartihms of real GDP and GDP per capita, host trade diversion term, share of exports in host's GDP, Heritage Foundation measure of institutional quality, and region dummies. Since the regressand is an estimate, the regression is weighted by the fixed-effects' standard errors. Both the first and second stage allow for clustering by host country. The results again seem to be in line with the preceding outcomes.

Summarizing the estimation results in Table 2, the host country trade diversion term is consistently negative, stable regardless of specification, and significant at least at the 5

[15] There are three regional dummies - the European OECD member countries, North America (the U.S.A. and Canada) and East Asia (Japan and South Korea).

percent level. According to the estimates, an increase in the share of immigrants (regardless of country of origin) in country j's population by one percentage point would result in a decrease in its *total* exports by roughly 6.9 percent.[16]

One should remember two important caveats when interpreting the scale of these estimates. First of all, as the average size of immigrant communities in the sample approaches 23,000, the considered one percentage point share increase would represent a huge population influx. This point becomes even more pronounced given the fact that in recent years most of the newcomers in the OECD countries originate from poorer countries with a smaller immigrant base vis-a-vis established migrant populations (Pedersen et al. 2002). To account for this problem, Table 3 presents an alternative measure of the trade diversion effect – the response of *total* exports to a 10 percent increase of the mean immigrant population in a host country j. As one could easily check, the implied fall in total exports becomes much smaller.

Second and perhaps more importantly, the underlying matching model departs from the assumption of equally informed agents and instead differentiates along the information dimension (through the introduction of network ties). The obtained estimates thus quantify the consequences of the relaxation of the homogeneity assumption and implied trade losses due to ties and informal trade barriers, with the relevant benchmark economy hosting poorly informed agents only and totaling the *same* population size as our economy with a fraction of fully informed traders (immigrants). In other words, the model substitutes poorly informed (natives) and fully informed agents *one-by-one*, whereas in practice the immigrants would rather *add* to the existing populations. In the latter case no trade diversion would ever

[16] The discussion as to why the coefficient is larger than theoretically predicted is given in Section 5.1.3.

Table 3: Predicted trade diversion after a 10 percent increase in immigrant stock

Country	Trade diversion (in %)
Austria	0.036
Belgium	0.049
Canada	0.085
Denmark	0.029
Finland	0.008
France	0.031
Germany	0.024
Greece	0.024
Ireland	0.054
Italy	0.018
Japan	0.006
Netherlands	0.033
Norway	0.032
Portugal	0.015
Spain	0.025
South Korea	0.002
Sweden	0.047
UK	0.038
USA	0.058

occur and hence any immigration policy recommendations based on trade diversion might be rather misleading.[17]

The matching model does not provide an answer to the question of how trade diversion might be related to the structure of immigrant populations in host countries. More specifically, one might wonder whether total immigrant populations consisting of mostly large immigrant communities divert trade more than more fragmented, smaller networks. In the case of trade creation, Gould (1993) found that for the U.S. economy the largest trade creating effect tends to be concentrated at relatively low immigrant numbers. Gould estimated that 90 percent of the immigrant information effect, which corresponds to trade creation in

[17] Take a host country of, say, total 10 million, consisting of 7 million natives and 3 million foreign-born. In the present situation the departure point is a host country with 10 million natives and no foreign-born (that's what the model dictates and what is meant by the one-by-one substitution of foreign-born 3 million for native 3 million), or perhaps a host with no natives and only foreign-born. In any case, the departure point is not a host country of 7 million natives. If someone was to evaluate whether to let more immigrants into any concerned host country, he should better focus on a comparison with the latter (7 million) case and the present estimates should not be relevant.

the present setting, will be exhausted at roughly 12,016 immigrants. For trade diversion, one might expect that as immigrant communities grow, profit opportunities from country-specific ties decline and other (potentially neighboring) markets could provide a rewarding alternative. Similarly, larger immigrant populations might establish numerous links to the host country economy and ultimately become more akin to native (poorly informed) agents. Hence ex ante one would expect the largest trade diversion effect to relate to smaller, more fragmented immigrant populations.

To get a rough idea of the issue, I take Gould's formula and calculate implied immigrant levels when 95 percent of the trade creation effect will be exhausted. Admittedly, the U.S. parameters derived by Gould would most likely vary across other host economies, but unfortunately there are no similar estimates for these countries. The obtained 24,967 immigrants almost exactly corresponds to the mean of the present sample. For larger populations I will assume that these 24,967 immigrants account for all trade creation and diversion generated by ties, while the remaining fraction will be treated as poorly connected. That is, large networks will contribute to trade diversion (and of course trade creation) only up to 24,967 immigrants. Then I re-estimate equation 6. The coefficient on trade diversion doubles to -14.95 consistent with expectations and is significant at the 10 percent level. Thus given that Gould's estimates are reasonably valid across all 19 OECD hosts in the sample, there seems to be some, albeit weak, support for the hypothesis that relatively more fragmented immigrant populations generate a stronger trade diversion effect.

5.1.2 Trade creation

Table 4: Trade creation, source country GDP and a 10 percent boost in mean immigrant stock.

Host	Source country	% share in world GDP	Population in j	Tr. creation in %
Belgium	Vietnam	0.10	6,204	2.41
	Romania	0.13	6,244	1.90
France	Sri Lanka	0.05	22,274	2.78
	Egypt	0.32	21,069	0.48
Germany	Tunisia	0.06	23,558	1.87
	Japan	16.64	25,994	0.01
Japan	Vietnam	0.10	10,630	0.37
	Indonesia	0.56	13,858	0.09
UK	Tanzania	0.02	31,931	5.15
	Greece	0.37	33,626	0.67
USA	Ecuador	0.05	283,065	5.00
	Iran	0.34	283,060	1.26

As in the case of trade diversion, the selection of a 1-percentage-point increase for comparative statics would not be of much value and hence again a 10 percent rise in a particular immigrant population will be considered. This exercise should moreover help comparison with previous studies.

While most of the existing research on immigrantion and trade sticks to the log-log specification, implying equal export elasticites regardless of the source economy's size or the relative size of the immigrant network of concern, the matching model developed in Section 2 suggests that the exclusion of both factors might lead to misleading results. To illustrate this, Table 4 lists the implied export increases into source country i resulting from a 10 percent boost of selected immigrant populations not far from host country j's mean immigrant population.

In order to compare the model's predictions with the results in the literature, I compute

Table 5: Unweighted mean for each host of implied trade creations

Host	Tr. creation (in %)
Austria	0.58
Belgium	1.04
Canada	1.67
Denmark	0.76
Finland	0.21
France	0.86
Germany	0.63
Greece	0.41
Ireland	0.61
Italy	0.60
Japan	0.07
Korea	0.03
Netherlands	0.79
Norway	0.87
Portugal	0.23
Spain	0.65
Sweden	0.97
UK	1.03
USA	1.11

an unweighted mean of the predicted increases of exports from host j to a source country i resulting from a 10 percent increase in the immigrant network. In general, the estimates appear roughly in line with the studies mentioned in Section 3.1. According to Girma and Yu (2002), a static version of their model produces a 1.6 percent increase for exports into non-Commonwealth countries as opposed to 1.03 percent from the matching model. Furthermore, Head and Ries (1998) find 1-1.3 percent for Canadian bilateral exports compared to 1.67 percent in this study. Gould (1993)'s long-run elasticity estimates for the U.S. exceed the present 1.11 percent by 3.6 percent. On the other hand, the U.S. study by Herander and Saavedra (2005) reports 1.6 percent. To my knowledge, there do not exist similar predictions for other OECD countries.

5.1.3 Trade creation vs. diversion

The matching model developed in Section 3.2 predicts that trade creation should balance with trade diversion and dominate it in case the matching technology possessed by immigrant networks is relatively more productive. The present estimates, however, suggest the opposite story, as the predicted *bilateral* trade creation does not suffice to outweigh the trade diversion of *total* exports. That is, the estimates of trade creation fail to reach the expected one, the coefficient on trade diversion tends to be far above its theoretical level and the overall network effect on the host country's exports is negative.[18]

How do these outcomes conform to our matching model? Effectively, having a fraction of the host population tied to a particular source country i de facto reduces the size of the concerned host economy for other trade partners. Even if the whole population of this host economy originated from the source country i and there would be no residual market to export from for other trade partners (all realized matches go to the source country i), with more productive matching technology the total trade of a concerned host would still increase. Now assume that equation (1) does not hold and productivity varies across economies. The assumption of equal productivity in production implies that country incomes and import demands should be proportional to the population and identical for equally sized economies. Once this assumption is relaxed, it might well happen that if trade is diverted from a market with higher import demand to less productive economies with lower import demand, the benefits of the more productive matching technology might be dominated by lost markets with high import demand (supported by higher productivity in production) and the resulting

[18] Note that while the trade diversion results of a 10 percent increase in the mean foreign-born population (see Table 3) might appear negligible, this effect refers to *total* exports of a host country.

overall trade effect would be negative.

Besides matching merely with their source country counterparts, immigrant networks could also operate across a larger number of countries. In such a case, the assumed matching mechanism might be too restrictive and the bilateral trade creation estimate would capture only a fraction of the total effect.

Another possible explanation might be the very assumption of a more productive matching technology possessed by immigrant networks, since the technology itself might aggregate over several dimensions. Arguably, immigrants are more likely to understand the local environment and business practices, on the other hand, they could lack the knowledge necessary for exports of more sophisticated products. For example, Turkish traders in Germany might specialize in exports of used cars instead of nuclear power engines.

The relatively low trade creation estimate might also be a consequence of the polynomial approximation, since the employed quadratic approximation could be sensitive to observations located further from zero. Section 5.3.2 will discuss this issue in more detail.

A final remark relates to how the present model is connected to the existing literature, which favors the use of level network variables such as natural logarithms of immigrant stock. While the relative size of immigrant networks as well as of trading parties seem to fare well empirically, the absolute levels could still remain important. I complement the relative measures of trade creation and diversion with the level variables typical in the literature on immigration and trade. As already mentioned, the most commonly employed level specification is the natural logarithm of immigrant stock. Table 5 presents the results.

The first two columns of Table 5 report the values of each measure individually. Using the specification with source country and host region fixed effects, the natural logartihm

Table 6: Absolute vs. relative measures

Variable	(1)	(2)	(3)	(4)
Trade diversion	-7.191**		-8.488***	-8.491***
	(2.807)		(2.821)	(2.820)
Trade creation	0.453***		0.087**	0.061
	(0.119)		(0.037)	(0.105)
Trade creation sq	-0.051***			0.005
	(0.020)			(0.017)
Ln immigrant stock		0.162***	0.143***	0.145***
		(0.026)	(0.025)	(0.025)
#	2,040	2,040	2,040	2,040
F-test	12.28***	40.38***	5.33**	3.69**
Tr. creation term(s)=0				
R-squared	0.933	0.936	0.936	0.936

***, **, * - significant at 1%, 5% and 10%, respectively

estimate in column (2) resembles the results found by the previous studies and confirms that the OECD sample contains a similar kind of information. One can also observe that the combination of relative and absolute measures preserves the statistical significance of both.[19]

To conclude, the joint significance of the relative and absolute terms suggests that despite support in the data, the matching model captures only a part of the immigrant story.

5.1.4 Colony and language terms

Both colony and common language have the expected signs. Colonial ties lose statistical significance once source-country fixed effects and clustering are allowed for. On the other hand, the role of a common language especially in smaller source economies seems to be robust for all specifications, though in some cases only at the 10 percent level. The argument by Anderson and van Wincoop (2001) thus finds partial empirical backing – the impact of information-related trade barriers is negatively related to the GDP of a trading partner. In particular, the larger the trading partner, the smaller the trade benefits of lower information costs due to a common language.

[19] In the specification with a squared trade creation term (column 4), however, the empirical relevance of the trade creation terms had to be established through a joint F-test.

The estimates for the language variable range from 3.8 to 4.4, depending on the specification. Take the mean value of GDP among countries sharing a common language as an example. In such a case the estimated contribution of easier communication amounts to roughly a 39 percent increase in exports from a given host country.[20]

5.2 Sensitivity analysis

5.2.1 Poorly informed migrants from OECD countries

A potential argument against the results from Section 5.1 might concern the composition of the immigration stock. Large communities originating from other OECD member states might qualitatively differ from immigrants born in less developed countries. In particular, both the supply and demand for network formation within the OECD context might be insufficient – intensive trade between the OECD economies over the last couple of decades led to an extended experience with foreign markets, IT and communication networks within the OECD are thicker, established business groups might provide alternative information channels, etc. On the supply side, the OECD migrants could be quite similar to the residents of the other OECD host countries and need not maintain any specific ties to their source country. Even if they did so, given the outside options, trade activities might be much less profitable for the OECD migrants relative to those coming from less developed countries. The assumption of active immigrant networks between the OECD economies thus might blur the true impact of ties with less developed countries.

I modify the specification of trade diversion and creation terms and treat immigrants from the OECD member states as poorly informed (i.e., natives). Table 7 shows the estimation

[20] The estimate used for calculation was 4.1, the midpoint of the above-mentioned interval.

Table 7: Regression results, dependent variable real exports 1999-2003

Variable	(1) Baseline OLS clustered	(2) Source FE clustered	(3) Source FE clust.+regions	(4) Wooldridge (2003)
Host diversion	-12.652***	-13.207***	-9.393***	-10.891**
	(4.385)	(4.118)	(4.760)	(4.504)
Source diversion	1.056**			
	(0.413)			
Trade creation	0.396**	0.374***	0.432***	0.344***
	(0.159)	(0.144)	(0.129)	(0.103)
Trade creation sq	-0.040	-0.041*	-0.048**	-0.036**
	(0.025)	(0.023)	(0.021)	(0.017)
#	2,040	2,040	2,040	2,040
R-squared	0.900	0.930	0.932	0.587

***, **, * - significant at 1%, 5% and 10%, respectively

Table 8: Trade creation* with poorly informed OECD migrants

Host	Tr. creation OECD migrants poorly informed (in %)	Change from original case
Austria	0.60	+0.02
Belgium	1.08	+0.04
Canada	1.86	+0.19
Denmark	0.80	+0.04
Finland	0.21	0
France	0.92	+0.06
Germany	0.64	+0.01
Greece	0.42	+0.01
Ireland	0.64	+0.03
Italy	0.68	+0.08
Japan	0.07	0
Korea	0.04	+0.01
Netherlands	0.84	+0.05
Norway	0.93	+0.06
Portugal	0.26	+0.03
Spain	0.74	+0.12
Sweden	1.04	+0.07
UK	1.13	+0.1
USA	1.25	+0.14

* Following a 10 percent increase in mean immigrant stock.

results. Because the signs and significance of other variables practically did not change, I report only the variables of interest (for full details see Table A3 in Appendix A). The trade creation as well as the source diversion term sticks to the original level. The host diversion term, however, increases substantially. An increase in the share of immigrants within country j's population by 1 percentage point would result in a decrease in its total exports by roughly 11 percent (using the estimate from column (4)). Hence not only are the results robust to the alternative specification, the trade diversion impact clear of the contribution by the OECD migrants becomes even more pronounced. Table 8 presents the calculated mean export response to a 10 percentage-point increase in immigrant stock. Similarly to the case of trade diversion with poorly informed OECD migrants, one can also observe stronger trade creation, though the increases are in most cases quite modest.

5.2.2 Functional forms, outliers and adjustments in migration data

The empirical version of the matching model relies on the quadratic approximation of the trade creation term for larger values. However, this could be one of the potential forces behind the lower-than-expected value of the trade creation variable, since the estimates might be sensitive to the inclusion of isolated observations located further from zero. A possible robustness check would then be to divide the trade creation data into several intervals and use a step function instead of the quadratic functional form. Note that since the distribution of the trade creation variable is skewed towards zero, the largest fraction of the data will fall into the lowest slots.

The resulting estimates illustrated in Figure 2 suggest that the quadratic function provides quite a reasonable approximation of the data. I draw step and quadratic functions for both the fully and poorly informed OECD migrants (see Section 5.3.1). While the individual intervals generate variable estimate levels, the general tendency towards a parabolic shape remains apparent.[21] Furthermore, the lower values of the step function indicate that the slope of the quadratic function might increase with the exclusion of more distant observations, thus bringing the trade creation coefficient closer to its predicted theoretical value.

While for the trade diversion term such an assumption might be justified,[22] the specification of the host trade creation term might potentially become too restrictive.

Another concern relates to outcomes being possibly driven by a handful of source countries, since the overall stock composition contains several large immigrant populations that

[21] Apart from the interval covering trade creation values between 1.2-1.4, all the steps proved to be statistically significant, usually at the 1-percent level.

[22] As an additional exercise, the imposition of higher order polynomials on the trade diversion term left the results practically unchanged both quantitatively and statistically.

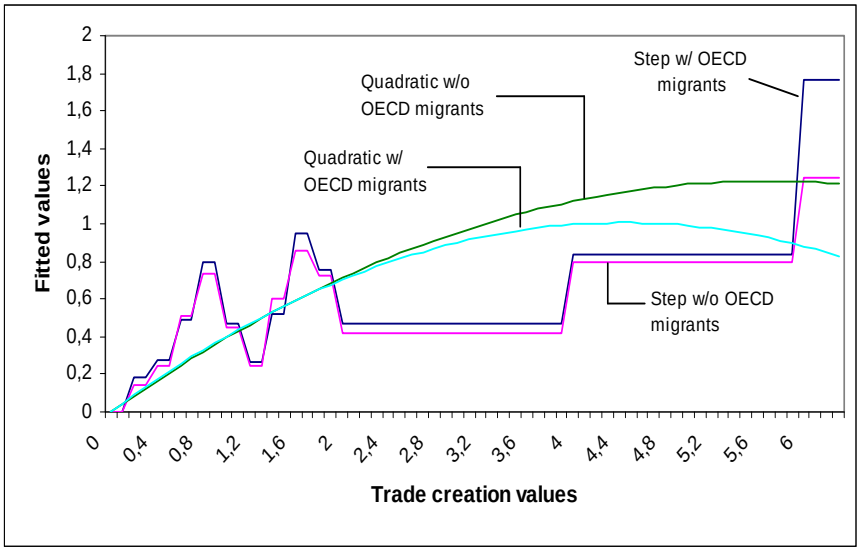

Figure 2: Step functions, quadratic approximation and trade creation

might dominate the rest of the data. Still treating the OECD member states' migrants as poorly connected, I drop the five countries with the highest share (out of 90 non-OECD states) of total immigrant stock.[23] Both the levels and the statistical significance of the output, however, remain the same.

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 remain highly significant.

5.2.3 Endogeneity

The ultimate reservation could regard the potential endogeneity of trade creation and diversion terms. Over time, trade partners could learn about the living conditions in the other

[23] The non-OECD source countries with the highest shares include Mexico (15.7%), China (4.3%), Turkey (3.7%), Philippines (3.5%) and Morocco (2.8%).

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 the undecided agents.

While similar reasoning seems to be in line with the findings of the literature on international migration,[24] previous studies on immigrant networks have avoided the endogeneity issue. Indeed, finding a suitable instrument for the trade creation variable proves to be a daunting task. A single exception is Javorcik et al. (2006), a study of migrant networks' ties 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.

The correlations of the two IVs and immigrant levels when all are expressed in natural logarithms are higher (0.14 and -0.21, respectively). Nonetheless, in the 2SLS regressions with the logartihms of both IVs and the natural logartihm of immigrant stock as the instrumented variable,[25] the Shea partial R-squared failed to pass 0.01 for any combination of the instruments and 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.

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

[25] In all regressions the trade diversion terms were dropped in order to allow the identification of the trade creation term.

Moving to the trade diversion term, it is hard to think of any significant endogeneity problem. The trade diversion variable relates the *total* immigrant share in the host population to *bilateral* trade. If bilateral trade between countries i and j promotes international migration between the two yet not between the host country j and other countries, its contribution to the total immigration share would be most likely negligible. Even the shares in the host population for the largest source country i do not exceed 3 percent.[26] Moreover, the mutual relationship between the immigration shares and bilateral trade should be positive, whereas the trade diversion term establishes a *negative* link. Hence, if anything, the endogeneity would then underestimate the impact of trade diversion by immigrant networks.

6 Concluding remarks

The study complements research on the links between immigrant networks and international trade. While previous work focused largely on trade creation by immigrant networks, I present a simple matching framework that allows for both trade creation and diversion. Immigrant networks can mitigate some informal barriers to trade, such as the lack of information on foreign markets or ineffective enforcement institutions. On the other hand, the same networks' advantages coupled with the pervasive presence of informal trade barriers might lead to trade losses similar to Viner's trade diversion by customs unions. By channeling trade exclusively to the immigrants' country of origin, potentially profitable matches in other countries become lost. Using a new dataset of 19 OECD countries, I find some empirical support for this hypothesis.

The trade creation and diversion measures derived within the matching framework point

[26] The only exception is the UK community in Ireland with 5.4 percent.

to the importance of the relative size of both the trading economies and immigrant networks. The benefits of a given number of immigrants will be larger for comparatively smaller source countries, since for smaller economies the implied decline in informal barriers spans a relatively larger part of bilateral trade. Consistent with these findings is the role of a common language in international trade, which was also found to be stronger for smaller economies. In the case of a mean-sized source country with a common language, such a stimulus boosts exports by as much as 39 percent.

However, the overall impact of immigrant ties becomes dominated by trade diversion. According to the estimates, an increase in the share of immigrants in a host country's population by one percentage point would result in a decrease in its *total* exports by roughly 6.9 percent. Moreover, under the assumption of poorly informed migrants from the OECD countries, the predicted fall rises to almost 11%.

Apart from being statistically significant, the results are robust to the inclusion of previously used level measures. However, further research should focus on a more flexible specification that could bring about more precise results on trade creation and diversion. In particular, future extensions should allow for heterogeneity in matching as well as production technologies.

Last but not least, future work should also focus on the search for valid instruments that could better capture potential endogeneity concerns.

References

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

Anderson, J. E. and van Wincoop, E., "Gravity with Gravitas: A Solution to the Border Puzzle", *NBER Working Paper 8079*, 2001

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

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

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

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

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

Pedersen, P.J., Pytlikova, 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

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

Wooldridge, J. M., "Cluster-Sample Methods in Applied Econometrics", *American Economic Review*, 93, 133-138, 2003

Appendix A:

Table A1: Countries in the Sample, n=109

Albania	Germany	Oman
Algeria	Ghana	Pakistan
Angola	Greece	Panama
Argentina	Guatemala	Papua New Guinea
Australia	Guinea	Paraguay
Austria	Guinea Bissau	Peru
Bahrain	Guyana	Philippines
Bangladesh	Haiti	Portugal
Barbados	Honduras	Qatar
Belgium	Indonesia	Romania
Belize	Iran	Rwanda
Benin	Ireland	Samoa
Bolivia	Israel	Saudi Arabia
Botswana	Italy	Senegal
Brazil	Jamaica	Seychelles
Bulgaria	Japan	Sierra Leone
Burkina Faso	Jordan	South Africa
Burundi	Kenya	Spain
Cambodia	Korea Rep.	Sri Lanka
Cameroon	Kuwait	Sudan
Chad	Lao P. Dem. Rep.	Suriname
Chile	Lebanon	Sweden
China	Liberia	Switzerland
Colombia	Libya	Syria
Congo	Madagascar	Tanzania
Costa Rica	Malawi	Thailand
Cote D'Ivoire	Malaysia	Togo
Cyprus	Mali	Trinidad and Tbg
Dem.Rep.Congo	Malta	Tunisia
Denmark	Mauritania	Turkey
Djibuti	Mexico	Uganda
Dominican Rep.	Morocco	UK
Ecuador	Mozambique	United Arab Em.
Egypt	Myanmar	Uruguay
El Salvador	Namibia	USA
Eq.Guinea	Netherlands	Venezuela
Ethiopia	New Zealand	Vietnam
Fiji	Nicaragua	Yemen
Finland	Niger	Zambia
France	Nigeria	Zimbabwe
Gabon	Norway	

Table A2: Complete regression results, dependent variable Real exports, n=2040 (n=2149 in Column (8))

Variable	(1) OLS	(2) OLS clustered	(3) Host FE	(4) Host RE	(5) Source FE clustered	(6) Source FE clust. + regions	(7) Wooldridge (2003)	(8) Turkey included
lnGDP1998	0.918***	0.918***	0.828***	0.838***	1.136***	1.121***	0.831***	1.122***
	(0.011)	(0.023)	(0.012)	(0.012)	(0.074)	(0.063)	(0.021)	(0.061)
lnGDPcap1998	-0.001	-0.001	0.070***	0.063**	-0.271	-0.218	0.103*	-0.211
	(0.019)	(0.056)	(0.021)	(0.021)	(0.339)	(0.332)	(0.052)	(0.331)
Host diversion	-7.853***	-7.853***		-7.185***	-8.658***	-7.191***	-6.899**	-7.084**
	(0.602)	(2.457)		(0.251)	(2.138)	(2.807)	(2.920)	(2.829)
Source diversion	1.004***	1.004**	0.099	0.195				
	(0.270)	(0.419)	(0.254)	(0.254)				
Trade creation	0.448***	0.448***	0.418***	0.420***	0.419***	0.453***	0.360***	0.366**
	(0.072)	(0.155)	(0.067)	(0.067)	(0.017)	(0.119)	(0.115)	(0.145)
Trade creation sq	-0.047***	-0.047*	-0.046***	-0.046***	-0.047**	-0.051**	-0.041**	-0.037
	(0.015)	(0.024)	(0.014)	(0.014)	(0.022)	(0.020)	(0.019)	(0.024)
Distance	-0.752***	-0.752***	-0.795***	-0.790***	-0.841***	-0.810***	-1.017***	-0.839***
	(0.030)	(0.055)	(0.029)	(0.029)	(0.131)	(0.138)	(0.086)	(0.133)
Colony	3.951***	3.951	2.466**	2.614**	3.115	2.278	3.014	2.650
	(1.263)	(4.038)	(1.241)	(1.240)	(3.805)	(3.488)	(2.098)	(3.450)
Language	3.812***	3.812**	4.268***	4.262***	3.188*	3.170*	4.351**	2.678
	(0.857)	(1.645)	(0.807)	(0.810)	(1.719)	(1.746)	(1.616)	(1.768)
Border	0.290*	0.290*	0.094	0.111	0.225	0.147	0.020	0.123
	(0.155)	(0.166)	(0.141)	(0.142)	(0.228)	(0.219)	(0.188)	(0.199)
Institutions	-0.010***	-0.010**	-0.010***	-0.010***	-0.010	-0.008	-0.006	-0.007
	(0.001)	(0.005)	(0.002)	(0.002)	(0.020)	(0.020)	(0.004)	(0.019)
Exports/GDP	1.899***	1.899***		1.529***	2.718***	2.543***	2.500***	2.564***
	(0.102)	(0.338)		(0.295)	(0.375)	(0.436)	(0.486)	(0.438)
Source EU dummy	0.155*	0.155	0.248***	0.238***	0.400**	0.413**	0.133	0.220
	(0.083)	(0.100)	(0.076)	(0.077)	(0.174)	(0.176)	(0.180)	(0.251)
Host EU dummy						0.245	0.127	0.245
						(0.232)	(0.271)	(0.229)
NAFTA						1.076***	0.121	1.160***
						(0.269)	(0.114)	(0.263)
East Asia						0.307	0.284	0.296
						(0.403)	(0.426)	(0.405)
Turkey								-2.664***
								(0.626)
Constant	-10.627***	-10.627***	-9.043***	-9.416***	-10.374*	-11.540*	-9.038***	-11.400*
			(0.411)	(0.889)	(0.480)	(0.489)	(5.653)	(6.034)
R-squared	0.901	0.902	0.900	0.900	0.932	0.933	0.676	0.923

***, **, * - significant at 1%, 5% and 10% level, respectively

Table A3: Complete regression results, OECD immigrants poorly informed, n=2040

Variable	(1) OLS clustered	(2) Source FE clustered	(3) Source FE clust. + regions	(4) Wooldridge (2003)
lnGDP1998	0.912***	1.107***	1.085***	0.816***
	(0.022)	(0.078)	(0.063)	(0.025)
lnGDPcap1998	-0.022	-0.314	-0.225	0.020***
	(0.056)	(0.348)	(0.346)	(0.004)
Host diversion	-12.652***	-13.207***	-9.393***	-10.891**
	(4.385)	(4.118)	(4.760)	(4.504)
Source diversion	1.056**		-10.231**	
	(0.413)		(4.480)	
Trade creation	0.396**	0.374***	0.432***	0.344***
	(0.159)	(0.144)	(0.129)	(0.103)
Trade creation sq	-0.040	-0.041*	-0.048**	-0.036**
	(0.025)	(0.023)	(0.021)	(0.017)
Distance	-0.772***	-0.856***	-0.837***	-0.945***
	(0.055)	(0.126)	(0.136)	(0.077)
Colony	4.500	3.841	2.532	3.009
	(4.138)	(3.796)	(3.433)	(2.135)
Language	3.633**	3.020*	3.104*	3.564*
	(1.717)	(1.820)	(1.819)	(1.728)
Border	0.386*	0.339	0.268	0.161
	(0.199)	(0.247)	(0.244)	(0.195)
Institutions	-0.013***	-0.014	-0.008	0.189
	(0.004)	(0.020)	(0.019)	(0.157)
Exports/GDP	1.457***	2.169*	2.060***	1.775***
	(0.332)	(0.487)	(0.480)	(0.515)
Source EU dummy	0.206*	0.387**	0.402**	0.673
	(0.104)	(0.173)	(.174)	(0.507)
Host EU dummy			0.307	0.418*
			(0.210)	(0.238)
NAFTA			0.210	0.748
			(0.386)	(0.438)
East Asia			1.039***	0.089
			(0.307)	(0.081)
Constant	-9.785***	-8.609	-10.328	-10.466**
	(0.820)	(5.924)	(6.631)	(4.570)
R-squared	0.900	0.930	0.932	0.587

***, **, * - significant at 1%, 5% and 10% level, respectively

Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the following institutional grants:

- Center of Advanced Political Economy Research [Centrum pro pokročilá politicko-ekonomická studia], No. LC542, (2005-2009),
- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2010);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Editors: Directors of CERGE and EI
Managing editors: Deputy Directors for Research of CERGE and EI

ISSN 1211-3298
ISBN 978-80-7343-128-0 (Univerzita Karlova. Centrum pro ekonomický výzkum
a doktorské studium)
ISBN 978-80-7344-117-3 (Národohospodářský ústav AV ČR, v. v. i.)

CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
http://www.cerge-ei.cz