Institutions Markets Technologies

IMT

INSTITUTE FOR ADVANCED STUDIES LUCCA

#05 2014

ISSN 2279-6894
IMT LUCCA EIC WORKING PAPER SERIES 04
May 2014

RA **Economics and institutional change**

The Migration Network Effect on International Trade

Rodolfo Metulini
Paolo Sgrignoli
Stefano Schiavo
Massimo Riccaboni

ISSN 2279-6894
IMT LUCCA EIC WORKING PAPER SERIES #05/2014

Piazza San Ponziano 6, 55100 Lucca

Research Area
Economics and institutional change

The Migration Network Effect on International Trade

Rodolfo Metulini
IMT Institute for Advanced Studies Lucca

Paolo Sgrignoli
IMT Institute for Advanced Studies Lucca

Stefano Schiavo
Department of Economics and Management, University of Trento

Massimo Riccaboni
IMT Institute for Advanced Studies Lucca

The Migration Network Effect on International Trade

Rodolfo Metulini[1], Paolo Sgrignoli[1], Stefano Schiavo[2,3], and Massimo Riccaboni*[1,4]

[1]IMT Institute for Advanced Studies Lucca - Laboratory of Innovation Management and Economics

[2]University of Trento - School of International Studies and Department of Economics and Management

[3]Observatoire Français des Conjonctures Économiques - Département de Recherche Innovation et Concurrence

[4]KU Leuven - Department of Management, Stategy and Innovation

Abstract

This paper studies the relationship between migration and trade, with the aim of measuring both direct and indirect network effects. We analyze trade of differentiated and homogeneous goods using an econometric approach inspired by spatial econometrics, proposing a new way to define country neighbors based on the most intense links in the migration network. We find that migration significantly affects trade across categories both in direct and in indirect way. The indirect impact highlights a stronger competitive effect of third country migrants for homogeneous goods. We also confirm that the effect of migration channels is higher on differentiated goods.

Keywords: Trade; Migration; Gravity model; Spatial econometrics, Networks
JEL Codes: F14, F22, C21

1 Introduction

Since the mid Nineties a growing body of research has investigated the relation between human migration and international trade. Whereas the standard Heckscher-Ohlin model

*m.riccaboni@imtlucca.it

suggests that the movement of goods across borders can provide a substitute for the movement of production factors, the empirical bottom line of these more recent works is that the two actually complement each other. This appears to hold for different countries (the US, Canada, Spain, Italy and France, to name just a few, see respectively Gould 1994, Head & Ries 1998, Peri & Requena-Silvente 2010, Bratti et al. 2012, Briant et al. 2014) and has recently been confirmed by a meta-analysis covering 48 different studies (Genc et al. 2011). As it has often happened in the international trade literature, empirical findings have percolated to economic theory, with recent models being able to accommodate the complementarity between migration and trade (Felbermayr et al. 2012).

We contribute to this growing field of research with a novel methodological approach that combines network analysis and spatial econometric techniques. On the one hand, this allows us to assess both the direct and the indirect effect of migration on trade without focusing on a single ethnic community at a time, as customarily done in the existing literature. On the other hand, spatial econometrics allows us to effectively account for the interdependences among trade flows that would otherwise lead to inconsistent (or even biased) estimates.

Most of the empirical literature we refer to shares a common strategy, based on the estimation of a log-linear gravity model where bilateral trade flows are regressed over standard explanatory variables (economic mass and distance), the stock of migrants from specific partner countries and other controls aiming at capturing various types of trade costs (common language, colonial relationships and the like). The two main strands of research that have emerged investigate the direct relation between trade and migration (i.e. the impact of migration from A to B on import/export flows between the same countries), and the existence of indirect or "network" effects (migration from A to both B and C not only affects trade from A to B and from A to C, but also establishes a connection between B and C due to the presence of a community of expatriates with the same background in both countries). The core of the argument (see for instance the seminal contribution by Rauch & Trindade 2002) is that formal and informal links among co-ethnic migrants in other countries and at home facilitate trade by providing potential trading partners with easier access to valuable, i.e. qualified, information. The pro-trade effect thus stems from the reduction of the trade barriers and search costs associated with market transactions. Since these costs are likely to be larger for international trade due to distance, language and cultural differences, legal provisions and the like, ethnic networks end up being especially relevant in facilitating cross-border transactions.

Indeed, one of the central results in the literature is that the positive effect of migration on trade is larger for "differentiated goods", i.e. those items that are not homogeneous and are not traded in organized exchanges therefore rendering that knowledge about counterpart reputation particularly valuable (Rauch & Trindade 2002).[1] Similar results have been replicated by a number of subsequent works using a variety of datasets and techniques. Peri & Requena-Silvente (2010), for instance, analyze the Spanish case and find that doubling the number of immigrants from a given country increases export to the same destination by 10 percent. This effect is higher for firms selling differentiated products and for more distant countries (geographically or culturally). All of these elements are consistent with the notion that networks (in this case the presence of a large community of expatriates and their connections with co-nationals at home and abroad) lower the hurdle in terms of economic interactions, providing better access to information and trade opportunities and reducing the fixed costs associated with entry into a foreign market. Aleksynska & Peri (2013) focus on the share of migrants involved in business activities rather than the total migrant population, and find a significant effect, even after controlling for the overall bilateral stock of migrants. Using trade data on Italian provinces, Bratti et al. (2012) find that the presence of migrants boosts both import from and export to their home countries, with the former effect being much larger. In the literature, this difference is interpreted as signaling a second channel through which migration affects trade, namely a home-country bias in demand by ethnic communities. Briant et al. (2014) also use a fine geographical disaggregation based on French departments to investigate the effect of migration on trade in goods with different degrees of complexity, as well as across countries with various levels of institutional quality. Migration is more relevant for complex goods, regardless of the quality of institutions in the partner country, whereas it matters also for simple products only matter when the institutional quality of the source country is low. A similar substitution effect between migrants and institutions is found in Ehrhart et al. (2014), who focus on African countries.

In parallel to these developments in the trade-migration literature, the past decade has witnessed important advances in both the theoretical foundations of the gravity model, and its estimation methods (Anderson & Wincoop 2003, Deardorff 1998). The literature has suggested that special care has to be applied in the empirical analysis to account for the interdependencies between trade flows that are inherent to the estimation of a general equilibrium model. In fact, Anderson & Wincoop (2003) show that bilateral export does not only depend on bilateral trade costs, the size of the trading economies and other dyad-specific characteristics, but also on *multilateral trade resistance* (MTR) i.e. the overall set of trade barriers that exporter and importer countries face. Several ways to account for MTR have been proposed: these involve the use of country-specific

[1] Although subsequent work has shown that the actual magnitude of this pro-trade effect is smaller than originally estimated (see Felbermayr et al. 2010), its existence and its specific importance for differentiated goods is confirmed.

effects (Feenstra 2003), export- and import-specific dummies (Anderson & van Wincoop 2004), measures of geographic remoteness (Helliwell 1998), as well as more sophisticated methods (see Head & Mayer 2013, for an excellent survey). Behrens et al. (2012) tackle the issue borrowing from the spatial econometrics literature (see also LeSage & Pace 2008, for an earlier contribution along the same lines, with an application to bilateral migration flows): they suggest using a spatial autoregressive moving average specification as a proxy for MTR, which results in a consistent estimation of the gravity equation.[2]

We build on both the aforementioned streams of literature to estimate the effect of migration on trade using spatial econometrics to adequately account for interdependences in trade flows. In fact, the key innovation proposed in the paper rests on the fact that spatial autocorrelation matrix is based on topological rather than geographical distance. More precisely, we build a world-wide network of migration connecting countries, and use distance in the network to define proximity. Hence, we proxy MTR introducing the global migration network into the model, assuming that migration network filters out the heterogeneity on the relative trade costs faced by exporting and importing countries. Our tests confirm that controlling for the global migration network eliminates the spatial autocorrelation, thus supporting our intuition.

The rest of the paper is structured as follows. Our empirical strategy is laid down in Section 2, which illustrates the rationale for our approach, the model specification, and the data used. Section 3 discusses our main results, while some concluding remarks are elaborated in Section 4.

2 Empirical strategy

The combination of network analysis and spatial econometrics we propose in the paper is summarized in Figure 1. We assume that trade between i and j depends both on variables specific to the country-pair (e.g distance, stock of bilateral migrants), but is also affected by third-country effects. In particular, we focus our attention on the potential impact that migrants from third countries (say k) may have on bilateral trade between i and j. Let k labels neighbors of the origin country i in the migration network: this means that there is a significant number of people born in i and resident in k.[3] Migration from k to j represents the third-country (indirect) effect we take into account in the empirical analysis. In other words, we investigate whether migration from k to j affects export

[2] The need to account for spatial autocorrelation in trade flows had been already recognized in Porojan (2001), although that paper suffers from serious methodological limitations pointed out by Johnston et al. (2003).

[3] What represents a *significant* number of migrants is explained in Section 2.1 below.

from i to j, given the existence of a strong migration link from i to k. Similarly, we could let h be a migration neighbor of the destination country j. In this case migrants from i to h should represents the indirect channel affecting trade from i to j. However, we have no theoretical and empirical reason to model this second type of network dependence.

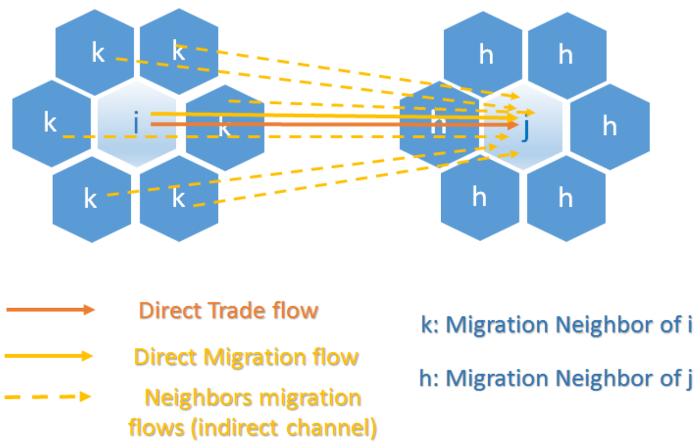

Figure 1: Exemplifying representation of the direct and indirect migration channels (origin-side)

2.1 Gravity models and spatial interaction

As mentioned above, the standard approach used in the empirical on migration and trade entails the estimation of a gravity model augmented with the stock of migrants. We follow a similar strategy and model bilateral trade in terms of per capita GDP to control for purchasing power and population to control for size. Following Baltagi et al. (2007) we construct pair-specific measures of both GDP and population rather than separately including information on both the origin and the destination countries, as this allows us to better interpret of our variables of interest. The control variables are defined as $GDPpc\_sum_{ij} = log(GDPpc_i + GDPpc_j)$ and $population\_sum_{ij} = log(population_i + population_j)$. Moreover, we also introduce similarities indexes defined as as $GDPpc\_sim_{ij} = (1 - (\frac{GDPpc_i}{GDPpc_i - GDPpc_j})^2 - (\frac{GDPpc_j}{GDPpc_i - GDPpc_j})^2)$ and $population\_sim_{ij} = (1 - (\frac{population_i}{population_i - population_j})^2 - (\frac{population_j}{population_i - population_j})^2)$. Last, the model includes the stock of migrants and a number of standard controls such as geographic contiguity (*contig*), common language (*comlang*), common currency (*comcur*), colonial ties (*colony*) and participation into regional trade agreements (*rta*).

Since the seminal contribution by Anderson & Wincoop (2003) recent empirical works recognize the importance of adequately account for MTR, i.e. to consider interdependencies among trade flows, that stem from the estimation of a model resulting from a general equilibrium framework. A number of alternative methods have been proposed to deal with this issue, most of which are very effectively summarized by Head & Mayer (2013). Here we concentrate on two: the first entails augmenting the gravity model with exporter- and importer-specific dummies; the second models MTR in a way similar to spatial autocorrelation. In particular, Behrens et al. (2012) suggest a spatial autoregressive moving average specification for the gravity model, which results in consistent estimates of the parameters. They argue that the baseline fixed effects specification does not fully succeed in capturing the MTR dependencies in the error structure, and indeed find that the residuals still display a significant amount of autocorrelation. Anselin & Arribas-Bel (2013) demonstrated by means of a series of simulation experiments that fixed effects correctly remove autocorrelation only is some specific cases. In the empirical analysis we use the Moran I test to check for the presence of autocorrelation in the standard gravity model, and the ability of our specification accounting for spatial contiguity in the migration matrix to adequately proxy for MTR, and therefore remove this autocorrelation in the residuals.

To model the spatial autoregressive component one generally uses an $n \times n$ weight matrix (W) that defines the set of neighbors: most frequently W is based on spatial contiguity, so that $[w_{ij}] = 1$ if i and j share geographical borders, and 0 otherwise.[4]

It was recently argued that the matrix can be both spatial or non-spatial. Accordingly, several proposals have been made in the literature, such as using the technological similarities or the transport costs instead of spatial metrics. One of the newest suggestions, however, is to analyze the effect of network-propagation, viewed both as an alternative and a complement to the spatial effect. LeSage & Pace (2011) discuss the possibility of jointly modeling spatial and non-spatial dependence through a double autoregressive component that make use of two different weight matrix specifications (Elhorst et al. 2012). In general, network theory and spatial econometrics are intimately connected. Leenders (2002) proposes using Spatial Autoregressive models employing an ad-hoc W matrix based on network relations (in terms of social influences and communication); Farber et al. (2010) analyze the relationship between the topology property of networks and the properties of spatial models, performing several simulation tests. Manski (1993) gives a seminal contribution, as it lays the foundation for analyzing the exogenous, endogenous and correlated effects that researchers encounter both in network and econometric theory. Lee et al. (2010), following Mansky's work, propose a specification for estimating network models in presence of exogenous, endogenous and correlated effects. Furthermore, the correct specification for the estimation of network models has become a popular object

[4] Other formulations are based on inverse distance.

of study as of late (Bramoullé et al. 2009, Chandrasekhar & Lewis 2011).

To control for autocorrelation we use the matrix describing the migration network, so that topological distance replaces the more usual spatial weight matrix. In order to identify the significant links, we use a stochastic benchmark based on the hypergeometric distribution, as recently done in Riccaboni et al. (2013). The procedure starts from the null hypothesis that treat all links are randomly assigned following an hypergeometric probability distribution. For each pair of countries, we can thus compute the probability that the observed link weight comes from the same distribution, which takes as parameters the out- and in-strength of the nodes, plus the total amount of migrants observed in the network. Hence the procedures takes into account the heterogeneity of countries with respect to the total number of migrants and allows us to retain only those links that represent a significant departure from the hypergeometric benchmark.[5] The specification of the W matrix then becomes:

$$W^M : \begin{cases} w^M_{i,j} = 1, & \text{if } i \text{ has a significant migration} \\ & \text{relationship with } j \\ w^M_{i,j} = 0, & \text{otherwise.} \end{cases}$$

where the specific kroeneker transformation is applied so that the set of neighbors for each country-pair includes neighbors of the exporter countries.[6]

2.2 Model specification and estimators

Using spatial econometrics, the measure of the spatial (network) association in the origin-destination trade flow specification can be based on two classes of models (LeSage & Pace 2008): *Spatial autoregressive models* (SAR) and *Spatial Durbin / Spatial error models* (SDM/SEM). The former consists in the inclusion of either a spatially lagged dependent variable or of a spatial autoregressive process in the residual term, motivated by significant spatial autocorrelation in the dependent variable. This model can be augmented with the inclusion of the spatial lagged residuals, and it is called *Spatial autoregressive error model* (SARAR). The latter can be motivated by a statistical nuisance and it is best described as a proxy for missing variables that follow a meaningful spatial pattern. The

[5]We set the cutoff at 1%.

[6]To account for this, the W^M_K has dimension $n^2 \times n^2$ and it is generally constructed as the Kronecker product of W^M with the identity matrix I (as proposed in LeSage & Pace 2008):

$$W^M_K = W^M \otimes I.$$

In a panel framework one needs to account for the time index so that the matrix has to be pre-multiplied by a diagonal matrix of dimension t: $W^M_{K,t} = I_t \otimes W^M_K$.

econometric representation of the models can be illustrated as:

$$\text{SAR} \qquad y = \rho \mathbf{W} y + \mathbf{X}\beta + \epsilon \tag{1}$$

$$\text{SARAR} \qquad y = \rho \mathbf{W} y + \mathbf{X}\beta + \lambda \mathbf{W}\epsilon + \epsilon \tag{2}$$

$$\text{SDM} \qquad y = \rho \mathbf{W} y + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + \epsilon \tag{3}$$

which becomes the SEM model in the event that included and excluded variables are not correlated (common factor tests can be performed, see LeSage & Pace 2008)

$$\text{SEM} \qquad y = \mathbf{X}\beta + \lambda \mathbf{W}\epsilon + \epsilon. \tag{4}$$

The Durbin model can also take into account both the spatially lagged dependent variable and the spatial autoregressive process in the residuals: this augmented version of the SDM model (also called *Manski*) that fully accounts for all possible spatial dependency takes the form:

$$\text{Manski} \qquad y = \rho \mathbf{W} y + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + \lambda \mathbf{W}\epsilon + \epsilon \tag{5}$$

where y is the dependent variable, $\mathbf{X}$ is the matrix of the explanatory variables and ϵ represents the residuals. $\mathbf{W}$ is the (spatial) weight matrix, while β, γ, λ and ρ are the coefficients to be estimated. However, Elhorst (2010) argues that the SDM is the only model that provides unbiased parameter estimates and correct standard errors, even if the true data-generation process is any of the other spatial regression models mentioned above.

In spatial models, the presence of intrinsic endogeneity due to the inclusion of a spatial lag of the dependent variable among the controls renders OLS estimation inconsistent. The standard alternative in this literature is the concentrated maximum likelihood (CML) estimator proposed by Anselin (1988) and revised by LeSage & Pace (2008).[7] Last, we are aware of the fact that a log-log model implies non-realistic assumptions about homoscedasticity in the residuals, and will explicitly test for this in the empirical analysis.

[7]Fitting a CML estimator on a log-log gravity model disregards the presence of zero trade flows, which represent around 20 percent of our sample. The standard literature has addressed it by considering trade flows as count processes and fitting Poisson or negative binomial models. However, to the best of our knowledge no extension of this approach exists that combines it with spatial autoregressive models. The alternative to fit a zero inflated Poisson model in which the spatial effect is captured by spatial-filtering eigenvectors (see Lionetti & Patuelli 2009) would prevent us from distinguishing between direct and indirect spatial effects.

2.3 Data

Data regarding migrants come from the World Bank's Global Bilateral Migration dataset (Özden et al. 2011): it is composed of matrices of bilateral migrant stocks spanning five decades from 1960 to 2000 (5 census rounds), and based primarily on the foreign-born definition of migrants. It is the first and only comprehensive picture of bilateral global migration over the second half of the 20th century, taking into account a total of 232 countries. The data reveal that the global migrant stock increased from 92 million in 1960 to 165 million in 2000. Quantitatively, migration between developing countries dominates, constituting half of all international migration in 2000, whereas flows from developing to developed countries represent the fastest growing component of international migration in both absolute and relative terms.

For international trade, we use the NBER-UN dataset described by Feenstra et al. (2005), disaggregated according to the Standardized International Trade Code at the four-digit level (SITC-4). For each country it provides the value (expressed in thousands of US dollars) exported to all other countries, for 775 product classes. In our analysis, we focus on the years 1970, 1980, 1990 and 2000.

Looking at the SITC product code of goods traded between each country pair allows us to apply Rauch & Trindade's (2002) classification and distinguish between homogeneous and differentiated goods. Trade in the latter type of products are more heavily influenced by the presence of migrant networks, as buyers and sellers need to look for relevant information that is not easily embedded in prices.

We only consider countries present in both datasets: this results in a final sample of 146 countries (nodes) that have active interactions in both trade and migration. All the other controls used in the regressions (e.g. contiguity, common language, etc.) have been retrieved from the CEPII dataset documented in (Mayer & Zignago 2011).

3 Results

We conduct a panel regression estimation using pooled data from the years 1970, 1980, 1990 and 2000.[8] We employ three different dependent variables: (*i*) total exports; (*ii*) export of differentiated goods; and (*iii*) export of homogeneous goods.

We start by estimating a baseline gravity model for total trade without migration using pooled OLS; results, presented in the first column of Table 1, are in line with the literature. In column 2 of the table we add the stock of migrants to the model, where

[8] A cross sectional analysis was also performed for the years 1970 and 2000 as a robustness check. Results are available upon request.

we note that the migration coefficient (0.129) is in line with the meta-analysis by Genc et al. (2011), who report coefficients that vary between 0.13–0.15. Moreover, we find that adding migration to the explanatory variables lowers the impact of distance. This is in good agreement with the literature (see for instance Felbermayr et al. 2012) and suggests that distance picks up a host of formal and informal informational barriers.

Table 1: Gravity results with OLS and FE models, with and without instrumenting migration for reverse causality

Non instrumented	base	total trade		diff. goods		homog. goods	
	ols	ols	fe	ols	fe	ols	fe
distance	-.858***	-.706***	-1.002***	-.669***	-.055***	-.728***	-1.011***
GDPpc_sum	1.746***	1.654***	-	1.732***	-	1.497***	-
GDPpc_sim	.933***	.888***	-	.851***	-	.868***	-
population_sum	1.622***	1.476***	-	1.446***	-	1.438***	-
population_sim	.774***	.703***	-	.617***	-	.742***	-
contig	.268***	.168***	.79**	.184***	.144***	.118***	.017
comlang	.188***	.082***	.129***	.118***	.244***	.108***	.093***
colony	.604***	.471***	.455***	.375***	.313***	.401***	.443***
comcur	.360***	.289***	.298***	.345***	.270***	.248***	.300***
rta	.187***	.148***	.005	.324***	.009	.074**	0.041
migration		.129***	.128***	.133***	.140***	.109***	.113***
R^2 adj	.639	.639	.752	.629	.820	.604	.716
obs	29784	24105	27217	20908	23467	22256	24813
Instrumented		total trade		diff. goods		homog. goods	
		ols	fe	ols	fe	ols	fe
distance		-.776***	-1.064***	-0.680***	-.075***	-0.805***	-1.086***
GDPpc_sum		1.896***	-	1.944***	-	1.687***	-
GDPpc_sim		.955***	-	.899***	-	.915***	-
population_sum		1.659***	-	1.594***	-	1.590***	-
population_sim		.783***	-	.676***	-	.815***	-
contig		.229***	.074*	.228***	.144***	.201***	.025
comlang		.149***	.114***	.152***	.239***	.172***	.071***
colony		.384***	.429***	.283***	.288***	.339***	.429***
comcur		.175**	.088	.201**	.067***	.206**	.128*
rta		.093***	-.035	.280***	-.029	.027	-.007
migration		.088***	.121***	.109***	.135***	.070***	.105***
R^2 adj		.636	.746	.608	.806	.589	.707
obs		17448	18551	15261	16124	16211	17039

A specification that includes origin- and destination-specific fixed effect has been widely applied in estimating the gravity equation for international trade, to accounts for MTR. Here we opt for importer and exporter time-varying fixed effects (FE) as suggested by the most recent literature (Felbermayr et al. 2012, Head & Mayer 2013) and find a migration coefficient of the same magnitude as before (0.129 with OLS, 0.128 with FE). Columns 4–7 of Table 1 report OLS and FE results for export of differentiated and homogeneous goods: the migration coefficient is higher in the former case, in line with expectations.

An important issue that has recently moved to center stage is potential endogeneity biases. Since the causal relationship between trade and migration can hold both ways, to disentangle the effect of migration on trade one needs to adopt an instrumental variable strategy. We follow the literature (Felbermayr et al. 2012, Briant et al. 2014) and use data from the previous decade ($migration_{t-1}$) as an instrument for contemporaneous migration. Results for an F-test on the validity of instruments and a Durbin-Wu-Hausmann test for endogeneity are reported in table 2: for all the three dependent variables they confirm the presence of endogeneity and the necessity to use instruments, as well as the validity of the IV strategy adopted. The migration coefficients using the IV model (columns 8–13 of Table 1) are lower than in the standard OLS, but the positive effect of migration on trade persists and remains larger in the case of trade in differentiated goods.

Table 2: Tests for migration endogeneity and instruments

	total trade	diff. goods	homog. goods
Correlation between $Trade_t$ and $Migration_t$	0.35	0.37	0.29
Correlation between $Trade_t$ and $Migration_{t-1}$	0.28	0.29	0.22
First stage test for the validity of the instrument	>37.75	>37.75	>37.75
Durbin-Wu-Hausman for the endogeneity in the model	14.16	4.70	12.77

The Moran I test on the residuals of the unconstrained gravity model confirms the presence of residual autocorrelation. Here, our unconstrained gravity model corresponds to the baseline OLS. As we can see in the columns 2-3 of Table 3, the OLS residuals still display some positive autocorrelation, measured with both the spatial weight matrix (column 2) and with our migration network matrix (column 3).[9] The autocorrelation is significant for all the classifications (all trade, differentiated and homogeneous goods). The FE model that incorporates origin- and destination-specific effects to account for the MTR does not properly capture all the residual autocorrelation: the Moran I tests (columns 4 and 5 of Table 3), still finds a significant (negative) autocorrelation. This motivates the use of the SDM/SEM model in the rest of the analysis, since we were able

[9]The *spatial weight matrix* is constructed using the k-nearest neighbors method. To make the network and spatial weight matrices comparable in terms of concentration, we choose $k = 15$, resulting in a spatial weight matrix having a mean number of 18.38 neighbors based on geographic proximity.

to empirically confirm the findings of Behrens et al. (2012) regarding the lack of the FE formulation to fully filter out all of the residual autocorrelation.

Table 3: Moran I test for autocorrelation on the residuals of the gravity model estimated by OLS (colums ii. and iii.), FE (columns iv. and v.) and SDM (columns vi. and vii.)

	OLS		FE		SDM	
matrix	15 near.neigh. contiguity	Migration network	15 near.neigh. contiguity	Migration network	15 near.neigh. contiguity	Migration network
total	0.078	0.077	-0.011	- 0.008	-0.000	0.001
z-score (p-val)	29.21(0.000)	28.01 (0.000)	-4.40 (0.000)	-3.49 (0.000)	-1.05 (0.144)	0.139 (0.444)
differentiated	0.087	0.081	-0.010	-0.009	0.001	-0.012
z-score (p-val)	28.14 (0.000)	25.44 (0.000)	-3.83 (0.000)	-3.43 (0.000)	1.152 (0.123)	-2.297 (0.011)
homogeneous	0.075	0.081	-0.012	-0.011	-0.001	0.001
z-score (p-val)	26.07 (0.000)	27.31 (0.000)	-4.87 (0.000)	-4.16 (0.000)	-1.209 (0.116)	0.299 (0.381)

Adding the spatial autoregressive components to the gravity model seems therefore fundamental in order to grasp the potential contributions of the network of migration, and to test whether this network structure can capture the residual autocorrelation stemming from MTR. In order to do so, we make use of the previously computed 146×146 matrices for the 2000 time period, representing the network of country to country migrations. We perform both the SAR/SARAR and the SDM/SEM models with the CML estimator using network matrices as weights.[10] To choose from different specifications of the model we perform a likelihood ratio test, starting from the most general case (SDM) as suggesed by Elhorst (2010).

The first three columns of Table 4 report the results obtained from the estimation of the following final equation:

$$T = \rho \boldsymbol{W_t^M} T + \sum_{k=1}^{K} \mathbf{X}_k \beta_k + \sum_{k=1}^{K} \boldsymbol{W_t^M} \mathbf{X}_k \gamma_k + \epsilon \tag{6}$$

where T is the dependent variable, ρ is the scalar coefficient of the lagged trade term to be estimated, β and γ are the $k \times 1$ vectors of coefficients to be estimated for, respectively, the explanatories and the lagged explanatories X_k, where the regressors k are the following: *distance*, *GDPpc_sum*, *population_sum*, *GDPpc_sim*, *population_sim*, *migration*, *contig*, *comcur*, *comlang*, *colony* and *rta*.[11] Finally, W_t^M is the $n^2 * t \times n^2 * t$ network

[10]We also compute a CML estimator, separately, using the spatial matrix based on geographic proximity. Results are available upon request. On this issue, LeSage & Pace (2011) discuss the conjoint use of two or more weight matrices in the same model (one spatial and the other non-spatial), but some pitfalls emerge. We may analyze this in future developments.

[11]All the data (except for the dummies) are in log10.

weight matrix relative to migration.

We have performed the common factor test for SDM versus SEM. Results point toward the SDM specification, which accounts for the lagged dependent variable, and lagged explanatory variables. Likelihood ratio tests for the choice between SAR/SARAR models and SDM were also performed, leading to favour the SDM. The SDM, infact, as confirmed in the literature (Elhorst 2010) is able to correct for the parameters mispecification due to autocorrelated omitted variables, even when the true model is not a SDM. However, in order to let our work comparable with Behrens et al. (2012), we also have estimated the SAR and the SARAR specifications. As we can see in table 4, the estimated ρ parameter for the lagged dependent variable is positive, while Behrens et al. (2012) found this parameter to be negative in the SARAR specification. We also found a negative ρ when performing SARAR model.[12]

The SDM model controls both for the dyad and for the migration network lagged explanatory variables, in order to allow changes in a given explanatory variable associated with a single country-pair to affect the pair itself, and to potentially reverberate across all other dyads indirectly. This rich set of information increases the difficulty of interpreting the regression results. For the sake of clarity, we therefore calculate the direct and indirect impacts as suggested by Pace & LeSage (2009) and discussed by LeSage & Thomas-Agnan (2014) for exogenous and endogenous flow models. We present the figures in Table 5.[13]

Comparing the first three columns of Table 4 with the upper panel of Table 5, we see that the direct effect of migration is in line with OLS and FE results displayed above (see Table 1).

Analyzing the total effects, we note a negative indirect coefficient for differentiated goods, which significantly lowers the total impact of migration on trade. One possible interpretation of this negative indirect effect is that migrants also bring knowledge, competences and business contacts that are particularly relevant for producing and exporting differentiated goods. As a result, migration from i to h may erode i's ability to export specific goods to other markets (e.g. to country j), making h a better competitor. So, we decided to estimate and report both the SDM regression results without (first three columns of table 4 and upper panel of table 5) and with (last three columns of table 4 and bottom panel of table 5) controlling for this phenomena. To control for this effect

[12] SAR and SARAR regression results are available upon request

[13] We compute these models in R with the spdep package. The models have been fitted using Monte Carlo simulations with 1000 replications using traces of powers of the network weight matrix, which considerably reduces computation time.

Table 4: Results from pooled panel SDM model with instrumented migration. Without (i) and (ii) with controls for import strength

	(i) **Baseline**			(ii) **Import strength**		
	total trade	diff. goods	homog. goods	total trade	diff. goods	homog. goods
distance	-0.784***	-0.689***	-0.810***	-0.785***	-0.681***	-0.816***
GDPpc_sum	1.924***	1.989***	1.701***	1.919***	2.229***	1.589***
GDPpc_sim	0.967***	0.939***	0.915***	0.964***	1.060***	0.867***
population_sum	1.662***	1.596***	1.588***	1.659***	1.786***	1.497***
population_sim	0.787***	0.679***	0.817***	0.785***	0.799***	0.759***
contig	0.217***	0.215***	0.188***	0.217***	0.180**	0.207***
comlang	0.140***	0.137***	0.166***	0.140***	0.145***	0.162***
colony	0.385***	0.279***	0.341***	0.384***	0.253***	0.356***
comcur	0.166***	0.182***	0.209***	0.167***	0.201***	0.202***
rta	0.115*	0.297***	0.045	0.113*	0.297***	0.046
migration	0.092***	0.115***	0.074***	0.092***	0.141***	0.063***
im_strength_net	-	-	-	0.003	-0.245***	0.125**
W.distance	0.177***	0.286***	0.110**	0.174***	0.251**	0.100***
W.GDPpc_sum	-0.222***	-0.275***	-0.159**	-0.234***	-0.173**	-0.197***
W.GDPpc_sim	-0.051	-0.195***	0.015*	-0.059**	-0.077*	-0.031
W.population_sum	0.008	-0.059	0.017	0.005	-0.057	0.035
W.population_sim	-0.060	-0.050	-0.071	-0.059	-0.063	-0.058
W.contig	0.026	0.111***	0.048	0.028	0.098***	-0.043
W.comlang	-0.006	0.137***	-0.002	-0.005	-0.006	-0.002
W.colony	0.008	0.043*	-0.068	0.009	0.024	-0.047
W.comcur	0.090***	0.239***	0.067**	0.093**	0.173***	0.103***
W.rta	-0.140***	-0.080*	-0.179**	-0.143***	0.052	-0.205***
W.migration	-0.023**	-0.014*	-0.044*	-0.026**	-0.004	-0.043**
W.im_strength_net	-	-	-	0.011	-0.011	0.003
ρ	0.035	0.051	0.051	0.033	0.042	0.039

we include total import by j net of imports from i among the controls:

$$im.strength.net_j = \sum_{k \neq i} T_{kj} - T_{ij} \quad (7)$$

Results that account for import strength appear in columns 4–6 of Table 4, and in the bottom part of Table 5. The additional control turns out highly significant and negative for differentiated goods, suggesting that export of such products from i to j is substituted by trade from other sources. Moreover, migration coefficients change considerably: accounting for import strength of the destination country, the total effect of migration for the differentiated goods is now significantly higher than for homogeneous goods (0.143 versus 0.021).

Table 5: Impacts from pooled panel SDM model with instrumented migration. Without (i) and with (ii) controlling for import strength

	Baseline specification								
	total trade			different. goods			homogen. goods		
	direct	indirect	total	direct	indirect	total	direct	indirect	total
distance	-0.782	0.154	-0.628	-0.685	0.261	-0.425	-0.810	0.071	-0.738
GDPpc_sum	1.923	-0.160	1.762	1.987	-0.180	1.807	1.700	-0.075	1.625
GDPpc_sim	0.967	-0.018	0.949	0.937	-0.153	0.783	0.915	0.064	0.979
population_sum	1.663	0.067	1.729	1.596	0.024	1.620	1.589	0.102	1.692
population_sim	0.786	-0.034	0.753	0.679	-0.015	0.663	0.817	-0.031	0.786
contig	0.217	0.035	0.252	0.217	0.127	0.344	0.189	0.060	0.248
comlang	0.140	-0.001	0.139	0.138	0.085	0.223	0.166	-0.035	0.130
colony	0.385	0.022	0.407	0.280	0.060	0.340	0.340	-0.053	0.287
comcur	0.167	0.098	0.265	0.185	0.259	0.444	0.210	0.081	0.291
rta	0.114	-0.139	-0.025	0.296	-0.068	0.228	0.046	-0.184	-0.138
migration	0.092	-0.021	0.071	0.115	-0.009	0.106	0.073	-0.042	0.031
	Controlling for import strength								
	total trade			different. goods			homogen. goods		
	direct	indirect	total	direct	indirect	total	direct	indirect	total
distance	-0.783	0.152	-0.631	-0.679	0.229	-0.449	-0.815	0.070	-0.745
GDPpc_sum	1.918	-0.174	1.743	2.228	-0.081	2.146	1.588	-0.140	1.448
GDPpc_sim	0.964	-0.028	0.936	1.059	-0.034	1.025	0.867	0.003	0.870
population_sum	1.660	0.062	1.722	1.787	0.020	1.806	1.498	0.096	1.593
population_sim	0.785	-0.034	0.751	0.798	-0.031	0.767	0.759	-0.030	0.729
contig	0.217	0.036	0.253	0.181	0.110	0.290	0.207	0.052	0.259
comlang	0.140	0.002	0.142	0.145	-0.000	0.144	0.162	0.004	0.158
colony	0.385	0.023	0.407	0.253	0.036	0.289	0.355	-0.035	0.321
comcur	0.168	0.101	0.268	0.203	0.187	0.390	0.203	0.114	0.317
rta	0.112	-0.143	-0.030	0.296	-0.041	0.255	0.044	-0.210	-0.165
migration	0.092	-0.023	0.068	0.141	0.002	0.143	0.062	-0.042	0.021

All in all, looking to the results controlling for import strength, we note that the GDP coefficients slightly decrease for the effect of the inclusion of the lagged GDP terms (W.GDPpc_sum), that highlights a negative indirect impact. The distance coefficient also decreases when we introduce lagged terms. In particular, distance matters more for trade of homogeneous goods compared with differentiated goods, to which corresponds a total impact of -0.449, signficantly smaller than the traditional distance coefficient for total trade, which vary from -0.7 to -1 in the literature. Interestingly, we find a negative indirect effect for the RTA dummy: this can be easily rationalized if we think that a trade agreement between a country's export partners is likely to have negative "indirect" effect on that country's ability to export because of trade diversion effects.

The impact of migration on trade is significantly higher for differentiated goods compared to homogeneous ones: the gap in the effect becomes even larger when we consider the total impact rather than only the direct one. In fact, we find a negative indirect impact of migration on total and homogeneous goods trade, while the counterpart for differentiated is close to zero. The negative indirect impact can be interpreted as a competition effect: having more third country migrants in the importer country reduces trade between the country pair. This is true for total trade and homogeneous goods trade, but not for differentiated goods and it is likely to depend on the fact that the latter are more difficult to substitute for, so that they suffer less competition from third countries.

We next control for the residual autocorrelation in the SDM model using the Moran I test. Looking to columns 6-7 of Table 3, we obtain encouraging results: using both the spatial matrix and the migration network matrix, the autocorrelation that was present both in the OLS and the FE residuals is no longer significant. This provides further support to our statement: the SDM model associated with a weight matrix based on network of migration successfully captures MTR. We also check for the normality assumption of the CML residuals in the selected model: they are normally distributed, confirming that the model is reasonably well-specified.

All in all, the controls for network interdependencies are always significant in our analysis. This means that the baseline gravity model does not account for network effects, which play a relevant role in shaping the world trade web. Furthermore, trade in differentiated goods is more strongly affected by migration, as predicted by Rauch.

4 Conclusions

Increased data availability both at national and international levels has triggered a host of research on the relationship between trade and migration. We contribute to this line of research by applying spatial econometric techniques that exploit topological distance on networks, rather than the usual geographical standard geographic space, in order to look at direct and indirect effects of migration on trade. In this way we can investigate the network effects suggested by Rauch's seminal papers from a global perspective, rather than focusing on a single ethnic network as done in the literature so far.

Our work also contributes the literature of spatial economics /econometrics that aims to control for the multilateral resistance terms in the constraint gravity equation for trade. We can draw several conclusions. First, accounting the multilateral resistance terms by means of a SDM specification using a migration network weight matrix, we filter out the residual autocorrelation. Furthermore, from a qualitative point of view we confirm the finding that migration has a larger impact on differentiated products, both at direct and global (network) level. Indeed, the negative effect that third-country migrants have on trade of homogeneous goods (testified by the negative indirect impact found in the estimation results) and that we rationalize as a competition effect, is no longer there when we focus on differentiated goods.

In future work we plan to accommodate multiple and different network effects in this setting. Moreover, we plan to investigate the changing role that different types of migrants (high-skilled, low-skilled) play in favoring the trade of specific commodities.

Acknowledgments

The authors acknowledge funding from the Italian Ministry of Education, University and Research (MIUR) through the National Research Program of Italy (PNR), the CRISIS Lab project, and the project "The global virtual-water network: social, economic, and environmental implications" (FIRB - RBFR12BA3Y). They would like to acknowledge helpful comments from Giorgio Fagiolo (S.Anna, Pisa) and Thomas Scherngell (AIT, Wien), and from partecipants to the SITIS 2013 conference in Kyoto. M.R. acknowledges funding from the Multiplex FP7 project (Foundational Research on MULTIlevel comPLEX networks and systems).

References

Aleksynska, M. & Peri, G. (2013), 'Isolating the network effect of immigrants on trade', *The World Economy* **37**(3), 434–455.

Anderson, J. E. & van Wincoop, E. (2004), 'Trade Costs', *Journal of Economic Literature* **42**(3), 691–751.

Anderson, J. E. & Wincoop, E. V. (2003), 'Gravity with Gravitas: A Solution to the Border Puzzle', *American Economic Review* **93**(1), 170–192.

Anselin, L. (1988), *Spatial econometrics: methods and models*, Vol. 4, Springer.

Anselin, L. & Arribas-Bel, D. (2013), 'Spatial fixed effects and spatial dependence in a single cross-section', *Papers in Regional Science* **92**(1), 3–17.

Baltagi, B. H., Egger, P. & Pfaffermayr, M. (2007), 'Estimating models of complex fdi: Are there third-country effects?', *Journal of Econometrics* **140**(1), 260–281.

Behrens, K., Ertur, C. & Koch, W. (2012), "Dual gravity': Using spatial econometrics to control for multilateral resistance', *Journal of Applied Econometrics* **27**(5), 773–794.

Bramoullé, Y., Djebbari, H. & Fortin, B. (2009), 'Identification of peer effects through social networks', *Journal of econometrics* **150**(1), 41–55.

Bratti, M., De Benedictis, L. & Santoni, G. (2012), On the pro-trade effects of immigrants, Discussion Papers 6628, Institute for the Study of Labor (IZA).

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

Chandrasekhar, A. G. & Lewis, R. (2011), Econometrics of sampled networks. mimeo.

Deardorff, A. (1998), Determinants of bilateral trade: does gravity work in a neoclassical world?, *in* 'The regionalization of the world economy', University of Chicago Press, pp. 7–32.

Ehrhart, H., Goff, M. L., Rocher, E. & Singh, R. J. (2014), Does migration foster exports? evidence from africa, Policy Research Working Paper 6739, World Bank.

Elhorst, J. P. (2010), 'Applied spatial econometrics: raising the bar', *Spatial Economic Analysis* **5**(1), 9–28.

Elhorst, J. P., Lacombe, D. J. & Piras, G. (2012), 'On model specification and parameter space definitions in higher order spatial econometric models', *Regional Science and Urban Economics* **42**(1), 211–220.

Farber, S., Páez, A. & Volz, E. (2010), Topology, dependency tests and estimation bias in network autoregressive models, *in* 'Progress in Spatial Analysis', Springer, pp. 29–57.

Feenstra, R. C. (2003), *Advanced international trade: theory and evidence*, Princeton University Press.

Feenstra, R. C., Lipsey, R. E., Deng, H., Ma, A. C. & Mo, H. (2005), World trade flows: 1962-2000, Working Paper 11040, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w11040*

Felbermayr, G., Grossmann, V. & Kohler, W. (2012), Migration, international trade and capital formation: Cause or effect?, Discussion Papers 6975, Institute for the Study of Labor (IZA).

Felbermayr, G. J., Jung, B. & Toubal, F. (2010), 'Ethnic networks, information, and international trade: Revisiting the evidence', *Annales d'Economie et de Statistique* **97/98**, 41–70.

Genc, M., Gheasi, M., Nijkamp, P. & Poot, J. (2011), The impact of immigration on international trade: a meta-analysis, Discussion Paper 6145, IZA - Institute for the Study of Labor.

Gould, D. M. (1994), 'Immigrant links to the home country: empirical implications for us bilateral trade flows', *The Review of Economics and Statistics* **76**(2), 302–316.

Head, K. & Mayer, T. (2013), Gravity equations: Workhorse, toolkit, and cookbook, Sciences Po Economics Discussion Papers 2013-02, Sciences Po Departement of Economics.

Head, K. & Ries, J. (1998), 'Immigration and trade creation: econometric evidence from Canada', *Canadian journal of economics* **31**(1), 47–62.

Helliwell, J. F. (1998), *How much do national borders matter?*, Brookings Institution Press, Washington, D.C.

Johnston, R., Hepple, L., Hoare, T., Jones, K. & Plummer, P. (2003), 'The mistreated model: Some technical comments on porojan's paper on 'trade flows and spatial effects'', *Open Economies Review* **14**(1), 11–14.

Lee, L.-F., Liu, X. & Lin, X. (2010), 'Specification and estimation of social interaction models with network structures', *The Econometrics Journal* **13**(2), 145–176.

Leenders, R. T. A. (2002), 'Modeling social influence through network autocorrelation: constructing the weight matrix', *Social Networks* **24**(1), 21–47.

LeSage, J. P. & Pace, R. K. (2008), 'Spatial econometric modeling of origin-destination flows', *Journal of Regional Science* **48**(5), 941–967.

LeSage, J. P. & Pace, R. K. (2011), 'Pitfalls in higher order model extensions of basic spatial regression methodology', *The Review of Regional Studies* **41**(1), 13–26.

LeSage, J. P. & Thomas-Agnan, C. (2014), 'Interpreting spatial econometric origin-destination flow models', *Journal of Regional Science* .

Lionetti, S. & Patuelli, R. (2009), Trading Cultural Goods in the Era of Digital Piracy, Quaderni della facoltà di Scienze economiche dell'Università di Lugano 0907, USI Universit della Svizzera italiana.

Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', *The review of economic studies* **60**(3), 531–542.

Mayer, T. & Zignago, S. (2011), Notes on CEPII's distances measures: The GeoDist database, Working paper 25, CEPII.

Özden, Ç., Parsons, C. R., Schiff, M. & Walmsley, T. L. (2011), 'Where on earth is everybody? The evolution of global bilateral migration 1960–2000', *The World Bank Economic Review* **25**(1), 12–56.

Pace, R. K. & LeSage, J. (2009), *Introduction to spatial econometrics*, Chapman and Hall/CRC, Boca Raton, FL.

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

Porojan, A. (2001), 'Trade flows and spatial effects: the gravity model revisited', *Open economies review* **12**(3), 265–280.

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

Riccaboni, M., Rossi, A. & Schiavo, S. (2013), 'Global networks of trade and bits', *Journal of Economic Interaction and Coordination,* **8**(1), 1–24.

Institutions
Markets
Technologies
IMT
INSTITUTE FOR ADVANCED STUDIES LUCCA