OESTERREICHISCHE NATIONALBANK
EUROSYSTEM

Working Paper 275

Immigration and Trade Creation:
Evidence from the Extensive and Intensive Margins

Ana Abeliansky, Inmaculada Martinez-Zarzoso, Anna Raggl

OeNB Working Paper Series
41 pages

Immigration and Trade Creation: Evidence from the Extensive and Intensive Margins

This paper examines how immigration to OECD countries shapes trade. Using data on migrants from about 180 origin countries (1995–2023) and structural gravity estimations, it shows that migration clearly boosts trade. The main effect appears on the extensive margin, meaning that migrants help countries start trading new products or open new trade relationships. In contrast, effects on the intensive margin — trading more of the goods countries already exchange — are weaker. Migrant networks reduce information and trust barriers, making it easier for firms to export new products.

Authors

Ana Abeliansky
Oesterreichische Nationalbank (OeNB), Vienna University of Economics and Business, ana.abeliansky@oenb.at

Inmaculada Martinez-Zarzoso
University Jaume I, University of Goettingen, imartin@gwdg.de

Anna Katharina Raggl
Oesterreichische Nationalbank (OeNB), anna.raggl@oenb.at

JEL classification
F10; F14; F22

Keywords
immigration; trade margins; extensive margin; gravity model; endogeneity; market entry; regularization programs; OECD

Robust pro-trade effects of migration

Focusing on trade between OECD countries and their trading partners, we find robust pro-trade effects of migration. Migrant networks matter most for expanding the extensive margin, enabling the exchange of new products and creating new trade relationships rather than intensifying pre-existing flows.

Strength of effect is heterogeneous

Initial migrants provide the largest trade-creation role, and that effect declines as diasporas grow. We also find higher pro-trade effects of migration if the origin or destination country is less economically developed.

Catalyst for economic integration

Migrant networks help to reduce information frictions and cultural barriers, enabling firms to enter new markets and diversify the range of traded products. Thus, immigration policies, including regularization programs, can indirectly foster trade by lowering the fixed costs of market entry.

Immigration and Trade Creation: Evidence from the Extensive and Intensive Margins

Ana Abeliansky[a,b], Inmaculada Martinez-Zarzoso[c,d], and Anna Katharina Raggl[a]

February 26, 2026

a) Oesterreichische Nationalbank (OeNB)
Otto-Wagner-Platz 3, 1090 Vienna, Austria

b) Vienna University of Economics and Business
Department of Economics
Welthandelsplatz 1, 1020 Vienna, Austria

c) University Jaume I
Institute of International Economics
Av Sos Baynat SN, 12006 Castellon, Spain

d) University of Goettingen
Department of Economics
Platz der Goettinger Sieben 3, 37073 Goettingen, Germany

Abstract

This paper reexamines the trade creation effects of migration using a large sample of 180 countries from 1995 to 2023. We estimate a structural gravity model of bilateral trade augmented with migration stocks, investigating whether immigration to OECD countries relates to the extensive and intensive margins of trade. To address potential endogeneity, we apply a two-step control function approach based on a first-stage gravity model of migration that exploits bilateral and time variation in migration regularization policies across OECD countries. Results obtained using PPML estimators show robust pro-trade effects of migration through the extensive margin, with non-linearities and heterogeneity across product types. The findings show that migration primarily fosters exports of new, differentiated products, while effects on the intensive margin are weaker for different types of goods. These results stress the role of migrant networks as facilitators of new trade relationships and highlight policy implications for migration and trade integration.

JEL classification: F10; F14; F22

Keywords: immigration; trade margins; extensive margin; gravity model; PPML; CFA; expectiles; OECD

Disclaimer: The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the OeNB or the Eurosystem.

Non-technical summary

This paper examines how immigration shapes international trade by analyzing a large dataset covering migration to Organisation for Economic Co-operation and Development (OECD) countries and their trade with 180 partner countries between 1995 and 2023. We are particularly interested in whether migrants mainly help countries start trading new products (extensive margin) or mainly deepen existing trade flows (intensive margin). The central argument is that migrants bring information, contacts and cultural knowledge that reduce trade costs, especially the fixed costs inherent in establishing new trade relationships. These mechanisms are particularly important for overcoming informational frictions that often prevent firms from entering foreign markets.

We estimate state-of-the-art gravity models of trade and combine them with a control function approach to address the possibility that trade and migration influence each other. We use variations in changes in migration regularization policies in OECD countries for identification. Our findings show that immigration increases imports from the origin countries of the migrants. This effect is driven by growth along the extensive margin—that is, through the appearance of new products in trade relationships rather than increases in the volumes of goods already traded. More migration increases total imports and substantially expands the range of imported products, whereas the impact on the intensive margin is small or insignificant when a wide array of controls is included.

Our study further uncovers several heterogeneities. We find evidence of non-linearities: the first increments of migration have the largest trade-promoting role, and the effect declines as diasporas grow. Importantly, the pro-trade impact is larger when either the origin or destination country is less economically developed. Expectile estimations show that the effect is broadly stable across the distribution of trade flows, though somewhat stronger at the lower end. The trade-creating effect of immigration could be more pronounced for differentiated goods where information barriers matter substantially and the knowledge of migrants is particularly valuable. Overall, the results remain consistent across different migration measures and econometric approaches, indicating that the trade-enhancing role of migration is both substantial and empirically robust.

Taken together, our findings show that immigration acts as an important catalyst for international economic integration. Migrants networks help reduce information frictions and cultural barriers, enabling firms to enter new markets and diversify the range of traded products. This implies that immigration policies, including regularization programs, can indirectly foster trade by lowering the fixed costs of market entry.

1 Introduction

Research on the empirical link between immigration and trade is extensive and shows a significant positive relationship between immigration and trade, with immigrants boosting trade between their home and host countries through a combination of network effects, shared knowledge, and demand for culturally familiar goods (Gould, 1994; Head and Ries, 1998). This effect is often analyzed using the gravity model, which estimates the impact of immigration on bilateral trade flows (Egger et al., 2012; Elguezabal et al., 2025). Studies have shown that immigration tends to enhance both exports and imports, although the effect varies by country, product type, and the development level of the immigrant's home country. Generally, immigrants' connections and familiarity with their home markets create opportunities for both imports and exports, with network and preference channels being particularly influential (Rauch, 1999).

The seminal study from Gould (1994) pioneered the analysis of the impact of immigrants on trade in the United States (US), concluding that immigrants' knowledge of foreign markets plays a crucial role in driving bilateral trade, especially exports of consumption goods. Head and Ries (1998) further showed that increases in Canada's immigrant population increased exports and imports, reflecting the strong demand for imported goods that immigrants often introduce. Also for the US, Dunlevy and Hutchinson (1999) found that as immigrant populations increased, so did trade with immigrants' home countries. This effect was especially strong in differentiated products, with imports responding more significantly due to shared cultural factors, such as language.

In the 2000s, evidence on the positive link between trade and immigration has been extended to other OECD countries. For example, studies for Italy (Bratti et al., 2014), Spain (Peri and Requena-Silvente, 2010), Sweden (Hatzigeorgiou, 2010), and UK (Girma and Yu, 2002), among other countries, show that migrant stocks have a pro-trade effect. More specifically, Bratti et al. (2014) noted that migration enhances both imports and exports, with slightly larger gains for imports, especially from lower-income countries where unique market insights are crucial for trade growth. Peri and Requena-Silvente (2010) found that immigrant networks in Spain had a more substantial effect on trade with developed countries, with an emphasis on differentiated goods. Hatzigeorgiou (2010) showed for Sweden that as migrant populations grow, both imports and exports increase, particularly for differentiated products that require specialized market information. Some authors found differential effects depending on the degree of economic development of the immigrants origin countries. For instance, Girma and Yu (2002) found that only immigration from non-Commonwealth countries had a significant effect on UK exports, whereas immigration from former colonies reduced imports, perhaps indicating a substitution effect generated by immigrant activities. Differently, White (2007) showed that US immigration from lower-income countries had a strong effect on exports, with a 1 percent rise in immigration linked to a 0.47 (0.15) percent increase in exports (imports). These studies show that networks created by immigrant communities are valuable for trade development, and more so for developing countries.

For groups of receiving countries, Lewer and Van den Berg (2009) showed that immigration to OECD countries stimulated trade mainly through three channels: increasing foreign direct investment in source countries, reinforcing trade networks bilaterally, and raising income in destination countries, whereas Piperakis (2011) found that immigrants to the EU-15 boosted imports and exports, but immigrants from different subgroups of countries showed heterogeneous impacts. Those from Eastern European countries or with common borders to EU countries positively affect European imports and exports. In contrast, immigrants from rich countries showed no effects.

Despite the broad coverage in terms of origin and destination countries and the heterogeneity of the effects, there is a lack of research investigating the impact of immigration on the extensive and intensive margins of trade. Hatzigeorgiou and Lodefalk (2021) in their review of over a hundred papers that examine the nexus between migration and internationalization, identify this gap in the existing literature. Most former studies have focused on the relationship between bilateral migration stocks and trade volumes, which reflects the intensive margin of trade between countries when examining only established trade relationships. However, migrants also play a role in opening trade with new foreign partners, as noted in studies by Hatzigeorgiou and Lodefalk (2015), Peri and Requena-Silvente (2010), Bastos and Silva (2012) and Ferragina et al. (2021). These studies are limited to Swedish, Spanish and Portuguese firm-level data and intra-OECD trade and the effect of education. To our knowledge, there is no macro-level research that investigates the causal effect of immigration on the extensive and intensive margins of trade for bilateral trade and their relative importance.

To fill this gap in the literature, the main contribution of this paper is to focus on trade flows between OECD countries and their trading partners to provide a quantification of the causal effect of immigration on the bilateral trade margins –extensive and intensive– using Hummels and Klenow (2005) definitions and a sound instrumentation strategy. The second contribution consists of exploring the non-linearities and heterogeneity in the relationship between migration and the margins of trade with the help of a control function approach. With this aim, we investigate whether the effect of migration on trade is non-linear and whether it varies by income and across the conditional distribution of the margins of trade. Methodologically, we estimate a structural gravity model of trade using recent advances in econometric methods for panel data (Anderson and Yotov, 2025; Baier and Bergstrand, 2007; Baier et al., 2014, 2019; Baldwin and Taglioni, 2006; Bergstrand et al., 2025). These developments exploit the bilateral nature of trade and migration flows in a non-linear framework that accounts for zero trade flows and accommodates the use of instrumental variables in a control function approach (CFA) to address endogeneity issues (Lin and Wooldridge, 2019).

Our main results are consistent with the hypothesis that migration mainly reduces the trade fixed costs resulting from information and trust frictions across migrant host and source countries, and thus primarily influences the extensive margin of trade. This is consistent with the results found for other determinants of trade flows, such as World Trade Organization (WTO) or currency union (CU) membership (Dutt et al. (2013); Bergin and Lin (2009)).

The rest of the paper is structured as follows. Section 2 describes the closely related literature; Section 3 presents the empirical strategy, data and variables. The main results are presented in Section 4 and a number of robustness tests in Section 5. Finally, Section 6 outlines the conclusions and some policy recommendations.

2 Migration and Trade Margins

At the macroeconomic level, immigration can influence trade through multiple channels. An increase in population driven by immigration tends to expand aggregate demand and output, which, in turn, raises the demand for imported goods and services. Exports may also increase if the presence of immigrants in export-oriented industries reduces unit labor costs, enhances productivity, or improves the overall international competitiveness of the host economy by fostering greater labor mobility and efficiency (Gould, 1994; Head and Ries, 1998; Orefice et al., 2025).

At the microeconomic level, immigrants maintain transnational networks that link their host and origin countries, thereby facilitating information exchange, lowering transaction costs, and reducing informational and cultural barriers to trade (Rauch, 1999). Their familiarity with home-country institutions, legal frameworks, and business practices -as well as their linguistic and cultural knowledge- can improve communication and trust in cross-border commercial relationships. These mechanisms are particularly relevant for the extensive margin of trade, as they can help firms enter new export markets or establish novel product linkages (Figueiredo et al., 2020). In addition, immigrants' consumption preferences often create demand for goods from their home countries, stimulating imports. Over time, this demand can extend to the native population through demonstration and cultural diffusion effects, as seen in the diffusion of ethnic cuisines and cultural products (Peri and Requena-Silvente, 2010).

The trade-facilitating effects of migration therefore operate through both exports and imports, whereas the preference channel primarily affects imports. The relative magnitude of these effects depends on structural and policy factors such as tariff barriers, market access, preferences and the integration of migrants into export-oriented sectors. If migrants disproportionately contribute to export expansion through network externalities and knowledge spillovers, their overall effect on exports may exceed that on imports (Figueiredo et al., 2020). Conversely, in contexts where home-country goods dominate consumption patterns, the import channel may be stronger. Understanding these heterogeneous mechanisms is central to identifying the causal impact of migration on the extensive and intensive margins of trade. Accordingly, the present analysis seeks to dig deeper on these channels by estimating the elasticity of trade flows with respect to migration stocks, distinguishing between trade creation through new product-country pairs (extensive margin) and the deepening of existing trade relationships (intensive margin).

Our research is also related to the international trade literature that analyses how the margins of trade are influenced by specific factors. Among these, the effect of free trade agreements (RTA) has been the targeted determinant in Baier et al. (2014) and Foster et al. (2011). The authors find that both margins are affected by RTA membership, but the magnitude of the effect depends on how the margins are measured, the timing and the depth of the agreements. Tariff effects have instead been examined by Debaere and Mostashari (2010) whose findings indicate that only a small portion of changes in the extensive margin is explained by tariffs. Other researchers have examined the effect of cross-innovation and intellectual property rights on the trade margins (e.g. Martínez-Zarzoso and Santacreu (2025); Foster (2014)). The main results indicate that the positive effect on trade comes mainly through the extensive margin, measured as the number of products exported to a given destination. Also, the role of WTO membership has been investigated by Dutt et al. (2013) finding that it increases the extensive margin, while having a negative impact on the intensive margin of trade. Extensions of this literature have addressed the heterogeneity of the effects for specific target variables that affect trade (Baier et al. (2014); Bergstrand et al. (2025)). In particular, Bergstrand et al. (2025) propose a Poisson-based expectile regressions to estimate the heterogeneous effects of RTAs across the entire conditional distribution of exports. However, they focus mainly on the effect on total trade.

Several measures have been used as proxies for the extensive and intensive margins of trade. For the extensive margin, some authors used the count of the number of products exported from a given origin to a given destination (Fieler and Eaton (2025); Foster et al. (2011); Martínez-Zarzoso and Santacreu (2025); among others). For the intensive margin, the share of trade in a bilateral relationship with respect to total trade has also been used. We use the margins suggested by Hummels and Klenow (2005) as in Bergin and Lin (2009) to estimate the effect of currency unions (CU) or in Baier et al. (2014) for RTA effects. The main advantage is that our margins account for the relative value of the baskets of goods traded.

Focusing on the migration and trade link, there are a few papers for specific host countries that investigate the effects of migration on the probability of exporting and on average exports (White (2007) for the US; Koenig et al. (2010) for France; Bastos and Silva (2012) for Portugal). The main findings are mixed. For US trade effects on extensive and intensive margins are found, whereas only effects on the number of firms are found for French firms and on both margins for Portuguese firms. Orefice et al. (2025) distinguish between the effect of migration on trading or not (potential proxy for the extensive margin) and on positive trade (proxy for the intensive margin) to disentangle the role of migration networks that operate at a bilateral level from that of productivity channels. Their main point is that there is also an export-enhancing effect of migration via productivity increases, apart from the other channels already examined in the literature, arguing that migrants also promote exports to other countries, outside the bilateral relationship. They use differentiated goods in a robustness check to show that in sectors that need more problem-solving skills, the effects of immigrants on total exports is larger than in others.

The empirical applications that examine the effect of immigration on trade flows are usually estimated using the gravity model of trade. Our research relies on recent developments in the estimation of gravity models of trade that isolate the effect on trade of target variables by using Poisson-based regressions that incorporate high-dimensional fixed effects (HDFE) (Yotov, 2024). This is a widely used approach to control for factors that are country-specific and vary over time, as proxies for multilateral resistance terms (MRT) (Anderson and van Wincoop, 2003), and bilateral dummy variables (country-pair fixed effects) that are time-invariant to account for unobserved time invariant factors in each trade pair (Baier and Bergstrand, 2007). The estimated models usually include MRT but exclude pair fixed effects, which are used only in a few studies that focus on the trade-migration link, but do not examine effects on the trade margins (Figueiredo et al. (2020); Elguezabal et al. (2025)). Most recent papers use instrumental variables to claim causality in linear model settings (Orefice et al., 2025). We take the advantage of the control function approach proposed by Lin and Wooldridge (2019) and apply it in a Poisson Pseudo-Maximum Likelihood (PPML) framework that allows for heterogeneous effects across several dimensions (Bergstrand et al., 2025). The next section describes the main specifications of the empirical gravity models used to isolate causal effects and tackle endogeneity and heterogeneity issues.

3 Methodology, Data and Stylized Facts

This section presents the modeling strategies to estimate the effect of migration stocks on the margins of trade in subsections 3.1-3.4. The main gravity model specification is outlined in 3.1, the formulation of the trade margins in 3.2, and the control function approach in 3.3. The specifications that include non-linearities and conditional expectiles are shown in 3.4. Finally, subsection 3.5 describes the main data sources and variables used in the empirical analysis and presents some stylized facts of the target variables.

3.1 Gravity Model Estimation

Our baseline econometric specification is a panel gravity model that takes into account most of the recommendations from the existing applied literature on gravity models of trade [1] and serves as a basis to estimate the effects of migration on the trade margins:

$$Y_{ij,t}^{k} = \exp\left\{\delta^{k} lnMIG_{ij,t} + \beta^{k} X_{ij,t} + \pi_{i,t}^{k} + \chi_{j,t}^{k} + \tau_{i,j}^{k}\right\} \times \varepsilon_{ij,t}^{k}. \tag{1}$$

where $Y_{ij,t}^{k}$ denotes k different dependent variables: (i) the value of bilateral import flows ($Value$) in nominal terms (Baldwin and Taglioni, 2006) from country j to country i at time t. (ii) The corresponding extensive margin (EM) and (iii) the intensive margin (IM) of imports from country j to country i at time t. The definitions of extensive and intensive margins of trade are taken from Hummels and Klenow (2005) and Baier et al. (2014), as detailed in the next subsection.

We use the PPML estimator in order to account for heteroskedasticity and to take advantage of the information contained in zero trade flows (Silva and Tenreyro, 2006, 2021). $lnMIG_{ij,t}$ denotes the natural logarithm of the bilateral stock of migrants that migrated from country j to country i at time t. $X_{ij,t}$ is a vector of control variables, that includes free trade agreement (RTA) and common currency (CU) dummy variables, which vary by origin, destination and time and may influence both bilateral trade and

[1] See, among others, Anderson and van Wincoop (2003); Baier and Bergstrand (2007); Baldwin and Taglioni (2006); Baier et al. (2019); Yotov (2022, 2024).

migration.[2] The next two terms $\pi^k_{i,t}$ and $\chi^k_{j,t}$, are exporter-year and importer-year fixed effects, respectively, which represent the MRT of Anderson and van Wincoop (2003), as well as accounting for any other determinants of trade flows on the exporter and importer sides that are country-time specific (Hummels, 2001; Baldwin and Taglioni, 2006; Olivero and Yotov, 2012). $\tau^k_{i,j}$ represents directional pair fixed effects (Baier et al., 2019), which according to Baier and Bergstrand (2007) mitigate endogeneity concerns.

3.2 Decomposing Trade Margins

The seminal contribution of Hummels and Klenow (2005) (hereafter HK) introduced a tractable and transparent method to decompose trade flows into their *extensive* and *intensive* margins using disaggregated trade data for a wide range of countries. Baier et al. (2014) adapted the framework to investigate bilateral trade flows instead.

Let $X_{i,j.t}$ denote the total value of exports from country i to country j in year t (j's imports from i). Following HK, the **extensive margin (EM)** of exports from i to j in year t is defined as:

$$EM_{ij,t} = \frac{\sum_{g\in\mathcal{G}_{ij,t}} X^g_{Wj,t}}{\sum_{g\in\mathcal{G}_{Wj,t}} X^g_{Wj,t}}, \tag{2}$$

where $X^g_{Wj,t}$ represents country j's imports from the world of product m in year t, $\mathcal{G}_{Wj,t}$ is the set of all products imported by j from the world, and $\mathcal{G}_{ij,t}$ is the subset of products exported from i to j. Thus, $EM_{ij,t}$ measures the fraction of world products that i exports to j, weighted by their importance in j's total imports.

Similarly, the **intensive margin (IM)** of exports from i to j is given by:

$$IM_{ij,t} = \frac{\sum_{g\in\mathcal{G}_{ij,t}} X^g_{ij,t}}{\sum_{g\in\mathcal{G}_{ij,t}} X^g_{Wj,t}}, \tag{3}$$

where $X^g_{ij,t}$ is the value of product g exported from i to j in year t. The intensive margin thus represents i's average market share in j's imports, conditional on the set of products that are actually traded.

An important property of the HK decomposition is that the product of the extensive and intensive margins equals the share of exports from i to j in j's total imports:

[2] In some specifications where we do not use country-pair fixed effects we also include variables of a bilateral nature, but which do not change over time. These are: the natural logarithm of distance between the capital cities, a dummy variable that proxies for whether there has ever been a colonial relationship between country pairs, another that takes the value of one when countries share a border (zero otherwise), and finally a dummy that also takes the value of one when countries share an official language (zero otherwise).

$$\frac{X_{ij,t}}{X_{Wj,t}} = EM_{ij,t} \times IM_{ij,t}, \quad (4)$$

where $X_{Wj,t}$ denotes country j's total imports from the world.

While Hummels and Klenow (2005) applied this methodology to a single cross-section of data, we apply it to a panel framework using annual trade flows over time, as done in the related literature (Baier et al., 2014; Bergin and Lin, 2009).

This framework allows for the identification of both the intensive and extensive trade creation effects attributable to migration, while controlling for exporter-time, importer-time, and bilateral fixed effects. Moreover, by implementing panel estimations with country-pair and time fixed effects, our approach mitigates sample selection and firm heterogeneity biases, as discussed in Helpman et al. (2008). Nevertheless, there are some factors that vary over time and bilaterally and could be generating endogeneity problems that are not solved by the addition of dyadic fixed effects. An example is bilateral labour agreements, which could be correlated with trade and migration simultaneously. Since information for all countries in our sample on these agreements is not available, to account for potential endogeneity bias, we present an empirical strategy in the next subsection based on external instruments that are not trade-related but determine migration flows.

3.3 Dealing with Endogeneity

In this subsection we propose a control function approach (CFA) for panel data models in non-linear settings, given that a typical instrumental variable model is not suitable in a PPML framework as explained in Lin and Wooldridge (2019).

Our aim is to identify the causal effect of $MIG_{i,j.t}$ on $Y_{ij,t}$, that is, on the value of trade, and to decompose effects into the *extensive* and the *intensive* margins. We use a CFA in the spirit of Wooldridge (2005, 2015) and Lin and Wooldridge (2019). The structural gravity model using expected values is given by:

$$\mathbb{E}\big[Y_{ij,t} \mid lnMIG_{ij,t}, X_{ij,t}, \theta_{i,t}, \phi_{j,t}, \eta_{ij}\big] \; = \; \exp\Big(\delta lnMIG_{ij,t} + \beta'_X X_{ij,t} + \theta_{i,t} + \phi_{j,t} + \eta_{ij}\Big), \quad (5)$$

where $lnMIG_{ij,t}$ is defined above and, if it is endogenous, a naive estimate of δ is biased. $X_{ij,t}$ is the same vector of control variables as in specification (1). The CFA augments (5) with an estimate of the migration residual from a first-stage estimation (see below). The second-stage equation is given by:

$$\mathbb{E}\big[Y_{ij,t} \mid \cdot\big] = \exp\Big(\delta lnMIG_{ij,t} + \beta'_X X_{ij,t} + \rho\, \hat{u}_{ij,t} + \theta_{i,t} + \phi_{j,t} + \eta_{ij}\Big). \quad (6)$$

where $\hat{u}_{ij,t}$ is the residual (or generalized residual) from the first-stage migration equation. A statistically significant ρ indicates endogeneity of $lnMIG_{ij,t}$. Under correct specification and valid instruments, δ from (6) is consistent for the structural effect.

First stage: migration equation and instruments.

We estimate a first-stage reduced form for $MIG_{ij,t}$ using exogenous instruments $Z_{ij,t}$. The instruments we propose are:

1. $Z^1_{ij,t}$: the *required months of prior residence* to become eligible for regularization in destination j (a deterrent instrument);

2. $Z^2_{ij,t}$: an interaction between a *migration regularization program* dummy in destination j at time t and the *initial* migrant stock in the country-pair (i, j), capturing that regularization increases attractiveness conditional on pre-existing ties;

3. $Z^3_{ij,t}$: the *average migrant stock* from countries in the origin's region (excluding origin i) residing in destination j ("regional diaspora"), i.e., a spatially aggregated stock that varies by destination, origin region, and time.

A non-linear PPML first-stage specification is given by:

$$MIG_{ij,t} = \exp\left\{\pi'_Z Z_{ij,t} + \pi'_X X_{ij,t} + \varphi_{i,t} + \vartheta_{j,t} + \tau_{ij}\right\} \times u_{ij,t}, \tag{7}$$

A linear first-stage specification is given by:

$$lnMIG_{ij,t} = \pi'_Z Z_{ij,t} + \pi'_X X_{ij,t} + \varphi_{i,t} + \vartheta_{j,t} + \tau_{ij} + e_{ij,t}, \tag{8}$$

where $Z_{ij,t} = \left(Z^1_{ij,t}, Z^2_{ij,t}, Z^3_{ij,t}\right)$ and $X_{ij,t}$ are the vectors of control variables (RTA and CU). In practice, we include the same HDFE (origin-time, destination-time, pair) in (7) and (8); that we use in the structural equation so that identification of π_Z arises from within-pair and time variation in the instruments. The first-stage residual (or generalized residual) is estimated as $\hat{u}_{ij,t} = MIG_{ij,t} - \widehat{MIG}_{ij,t}$.

Under the following assumptions:

-Relevance: $\text{Cov}(Z_{ij,t}, MIG_{ij,t}) \neq 0$ (first stage strong);

-Exclusion: $Z_{ij,t}$ affects $Y_{ij,t}$ only through $MIG_{ij,t}$ (conditional on $X_{ij,t}$ and the fixed effects);

-Conditional separability: the structural error in the trade equation can be written as a function of $u_{ij,t}$ and a residual orthogonal to instruments, The inclusion of $\hat{u}_{ij,t}$ in (6) controls for the endogeneity bias induced by $u_{ij,t}$. More formally, if the structural equation error can be represented as $\varepsilon^*_{i,j,t} = g(u_{ij,t}) + \nu_{ij,t}$, with $\mathbb{E}[\nu_{ij,t} \mid Z, X, \theta] = 0$, then a linear approximation $g(u) \approx \rho u$ justifies the CFA specification. Testing $\rho = 0$ is a test for endogeneity. An F-test (Chi^2 test) can be used to test for the relevance of the instruments. The assumption is that the three proposed instruments are exogenous and satisfy the exclusion restriction. Using shift-share instruments, which are constructed by interacting the initial stock of migrants with bilateral-time-varying variables is a widely accepted approach in the migration literature (Figueiredo et al. (2020); Orefice et al. (2025)).

Our two-stage approach is subject to sampling error in the first-stage estimation, which generates the control function. This sampling error needs to be accounted for in the standard errors of the second-stage estimates. An approach to avoid downward biases in the standard errors is to use bootstrap. We apply bootstrap procedures as robustness checks.

3.4 Dealing with Non-Linearities and Heterogeneity

An issue that could be a source of mis-specification in our empirical model is that non-linearities could be present in our target variable. For instance, it could be that migration is only positively affecting trade and its margins for lower values of migration stocks, but not for larger diasporas. It could also be that the effect changes with the income level of the origin or host countries or that it changes depending on the conditional distribution of exports. In order to test these hypotheses, we take two different avenues. First, we introduce a squared term of migration (alternatively, interactions between income and migration) in the gravity model and second, we will estimate the model across different parts of the distribution of the value of trade and the trade margins.

PPML formulation with squared term and interactions

We first introduce the squared term of migration as regressor:

$$Y^k_{ij,t} = \exp\left\{\delta^k_1 lnMIG_{ij,t} + \delta^k_2 (lnMIG)^2_{ij,t} + \beta^k X_{ij,t} + \pi^k_{i,t} + \chi^k_{j,t} + \tau^k_{ij}\right\} \times \varepsilon^k_{ij,t}. \quad (9)$$

where $lnMIG_{ij,t}$, defined above, is included together with its squared term $(lnMIG)^2_{ij,t}$. The same vector of control variables, $X_{ij,t}$, and HDFE as above are added. The squared term of the logged migration stock, if statistically significant, will indicate that the relationship between migration and trade depends on the specific level of the stock of migrants for a given pair of trading countries.

The next specification adds interactions between GDP per capita in origin and destination countries as regressors:

$$Y^k_{ij,t} = \exp\Big\{\delta^k_1 \ln MIG_{ij,t} + \delta^k_2 (\ln MIG \times \ln GDPH)_{ij,t} + \delta^k_3 (\ln MIG \times \ln GDPH)_{ij,t} \\ + \beta^{k\prime} X_{ij,t} + \pi^k_{i,t} + \chi^k_{j,t} + \tau^k_{ij}\Big\} \times \varepsilon^k_{ij,t}. \tag{10}$$

where $lnGDPH_{i,t}$ and $lnGDPH_{j,t}$ are the Gross Domestic Products of countries i and j, respectively.

In order to deal with endogeneity issues in this setting, we augment the models with additional regressors obtained from (7), to adapt these two specifications to a CFA framework. Following Wooldridge (2015), the additional controls needed are: the error term and its squared form from the first stage specification (7) are added in model (9), and the error term and its interaction with the centered GDPH variables are added in model (10).

Expectiles-PPML formulation

Following Bergstrand et al. (2025), we specify a multiplicative *expectiles* model to explore heterogeneity across the conditional distribution of trade. For a given expectile level $\tau \in (0, 1)$, define the conditional expectile function $\mu_{ij,t}(\tau)$ of $Y_{ij,t}$ given covariates by the solution of an asymmetric squared-loss problem with an exponential predictor:

$$\hat{\boldsymbol{\beta}}(\tau) \;=\; \arg\min_{\boldsymbol{\beta}} \sum_{ij,t} \rho_\tau\Big(Y_{ij,t} - \exp\big(\boldsymbol{Z}'_{ij,t}\boldsymbol{\beta}\big)\Big), \tag{11}$$

where

$$\rho_\tau(u) \;=\; |\tau - \mathbf{1}\{u < 0\}|\, u^2, \qquad \boldsymbol{Z}'_{ij,t}\boldsymbol{\beta} = \beta(\tau)\ln \mathrm{MIG}_{ij,t} + \boldsymbol{\gamma}(\tau)'\mathbf{X}_{ij,t} + \phi_{ij}(\tau) + \alpha_{i,t}(\tau) + \psi_{j,t}(\tau) + \delta_t(\tau).$$

The implied conditional τ-expectile of trade is

$$\widehat{\mu}_{ij,t}(\tau) \;=\; \exp\big(\boldsymbol{Z}'_{ij,t}\widehat{\boldsymbol{\beta}}(\tau)\big).$$

where Y denotes the value of trade, or alternatively, its extensive (intensive) margin, Z is a vector of explanatory variables and β is a vector of conformable parameters. In the gravity model, it describes the effect of the variables included in Z on the conditional mean of Y.

Interpretation. The slope coefficient on $\ln \mathrm{MIG}_{ij,t}$, $\beta(\tau)$, represents the elasticity of the τ-expectile of bilateral trade with respect to migration. It describes how migration relates to trade outcomes at different parts of the conditional distribution (e.g. lower-tail, median-like, upper-tail effects). Comparing $\beta(\tau)$ across τ reveals heterogeneity in the migration-trade relationship.

Estimation. Equation (11) is an *asymmetric least-squares* problem with a nonlinear (exponential) predictor. In practice, this is solved by an iterative numerical algorithm (e.g. iteratively reweighted least squares or gradient-based optimization), absorbing HDFE at each iteration (via within-transformation or alternating projections). Expectile estimation on levels naturally accommodates zeros.

For correct statistical inference, robust standard errors (clustered by dyad) should be computed or bootstrap procedures should be adapted to the expectile objective. We follow the first avenue. Given that we use a framework with three blocs of fixed effects, we rely on the estimation method proposed by Bergstrand et al. (2025) that accounts for incidental-parameter bias and relies on large-T asymptotics.

3.5 Data description

The data used in this study draws from several data sources. Firstly, bilateral trade data was obtained from the BACI dataset available at CEPII and spans from 1995 to 2023, normally available for more than 180 countries. The margins of trade are computed using disaggregated trade data at the six-digit level. Secondly, bilateral migration stock data were retrieved from the OECD (OECDstats) and therefore our sample is mainly determined by the available data. This leaves us with 192 origin countries and 33 destination countries. In our main results we use data for migrants based on the citizenship-concept. In robustness checks, we use migration stock data based on the country-of-birth concept, obtained from the OECD database.[3] As two additional robustness checks, we filled data with the information from country of birth when the citizenship principle was missing, and the mirrored version for the fourth robustness (country of birth filled with citizenship when missing, and the latter available).

[3] We use migration data based on the citizenship-concept due to better data availability.

Thirdly, as trade flows are also shaped by regional trade agreements (RTA), we construct an RTA dummy using the Regional Trade Agreements Database (RTA-IS: https://rtais.wto.org/), the information provided in Jose de Sousa's personal website (http://jdesousa.univ.free.fr/data.htm) and our own update. Similarly, to account for currency unions, we also build on the work from de Sousa and updated it with the information provided in the IMF's AREAER Database. The control variables, namely geographical distance (*lndist*), common colonial past (*colony*), common border (*contig*), and official common language (*comlangoff*), are also sourced from CEPII, while GDPs and population variables come from the World Development Indicators. Finally, to construct the instruments we need information related to the regularization programs in OECD countries. To this end, we use the data from Elguezabal et al. (2025) who kindly provided it to us.

Table 1 reports the summary statistics of the main variables used in the empirical analysis. The trade values are measured in thousands of USD and there is some dispersion, as indicated by their standard deviation. The trade margins, computed as shares, range from 0 to 1, although these extreme values occur only rarely in the dataset. The measure of bilateral migrant stock, in natural logarithms, also exhibits notable variation. On average, about 8% of the observations indicate the presence of a regularization program, and 3 months are needed, as an undocumented migrant in a host country to be eligible for the programs. On average, about 3% of the country pairs in the sample share a common colonial past and a common border, while this proportion doubles for pairs that share an official language. Finally, 38% and 6% represent, respectively, the share of country pairs that participate in a regional trade agreement or share a common currency.

Table 1: Summary Statistics

	N	Mean	SD	Min	Max
Value	82324	2029.41	10185.73	0.00	430008.61
IM	82324	0.01	0.04	0.00	1.00
EM	82324	0.29	0.30	0.00	1.00
lnmig	82324	5.93	3.00	-0.12	16.03
RegularPro	82324	0.08	0.27	0.00	1.00
MonthsReg	82150	3.16	18.52	0.00	192.00
MigStockReg	82012	17.08	103.70	0.00	1882.74
MigRegion	81816	5.65	1.85	0.00	10.71
lnGDPHo	80903	8.58	1.51	4.27	11.80
lnGDPHd	82324	10.46	0.61	8.02	11.80
lndist	81435	8.39	0.92	4.09	9.88
colony	81435	0.03	0.18	0.00	1.00
contig	81435	0.03	0.17	0.00	1.00
comlangoff	81435	0.06	0.23	0.00	1.00
RTA	82324	0.38	0.49	0.00	1.00
CU	82324	0.06	0.23	0.00	1.00

Notes: RegularPro takes the value of 1 if a regularisation program is implemented, zero otherwise; Z^1_{ijt}=MonthsReg is required months of prior irregular residence in the host country to be eligible for the regularization; Z^2_{ijt}=MigStockReg is the interaction between the regularization dummy and the stock of migrants in the first year with data; Z^3_{ijt}= MigRegion denotes the average migrant stock from countries in the origin region. lnGDPHo(lnGDPd) is the Gross Domestic Product per head for origin (destination) country in logs. lndist denotes the log of the geographical distance between capital cities in the country pair. Colony, contig and comlangoff are dummy variables for past or present common colonial links, common border and common official language. RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union).

To motivate the empirical results from the next section, we first plot the natural logarithm of the stock of migrants against the natural logarithm of the trade values. Figure 1 shows that the variables are strongly correlated. This suggests that countries with larger diasporas tend to trade more intensively with the origin countries of the migrants. This serves as a starting point for the next section, where we investigate this formally and also look into the trade margins.

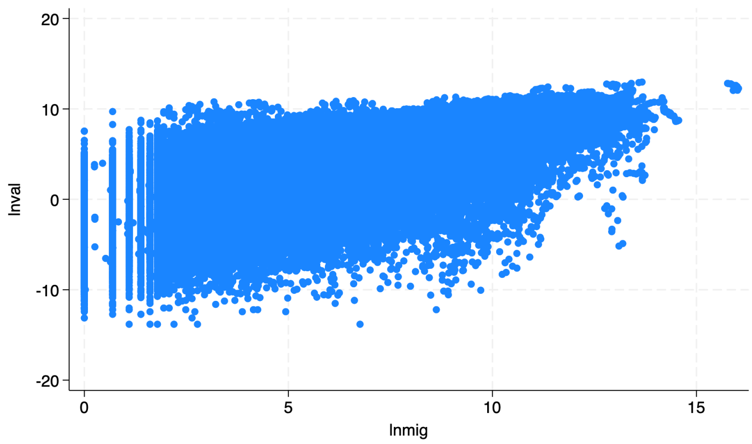


Figure 1: Scatter Plot: Migration Stock and Trade Value. Source: Produced with data from OECDstat for migration and BACI for the value of OECD imports, both in natural logs (1995-2023).

4 Main Results

The main results are presented in Table 2. This table reports estimates of the effects of the stock of migrants on the value of imports (Value), extensive (EM) and intensive margins (IM), which are computed according to equations (2) and (3), respectively. The estimates in columns (1)-(3), obtained from specification (1), indicate that the effect of migration on imports is mainly through the extensive margin. The estimated coefficient indicates that a 10% increase in migration stock increases the extensive margin by 0.23%, which is a sizable increase considering that the mean of the EM is 0.29 in our sample. Nevertheless, these estimates do not account for endogeneity that varies by origin-destination-time. To address this, the estimates in columns (4)-(6) are obtained from specification (6) that is based on the CFA described in the methodology section. When endogeneity coming from factors that vary by country pair and time are addressed, the effect of migration on trade increases considerably, showing that a 10% increase in the migration stock raises trade by almost 1.9% and the extensive margin by 0.7%. Our results indicate that the effects are biased downward in the baseline specification. This could be because migration is positively correlated with some omitted time-varying bilateral variables, such as tariffs or FDI, which negatively affect trade. The first step results are shown in Table A.1 in the Appendix, indicating that the three instruments used are not weak, with all three IVs showing statistically significant coefficients and the corresponding F-statistic and χ^2 test showing reasonable values. In order to compare our results with the related literature in Table 3, we estimate model (1) by replacing the bilateral fixed effects with the usual gravity variables, namely, distance, colonial links, common border, and common language. The estimated coefficients are higher in magnitude and the effect of migration on the intensive margin turns out to be significant, with and without considering the endogeneity of migration. This result indicates that controlling for all bilateral unobserved heterogeneity that is time invariant, as we did in Table 2, wipes out the intensive margin effect. This means that the cultural ties or trust that persist over time between two trading countries could be the reason why migrants from a given origin contribute to increase the trade share with its origin country. Differently, the push given by new migrants to trade new varieties remains significant even after controlling for these persistent bilateral links.

Table 2: PPML and CFA with pair FE

	(1) Value	(2) EM	(3) IM	(4) Value	(5) EM	(6) IM
lnmig	0.025	0.023***	-0.012	0.188**	0.071***	0.097
	(0.016)	(0.004)	(0.018)	(0.082)	(0.022)	(0.091)
RTA	0.093***	0.053***	0.003	0.093***	0.049***	0.008
	(0.025)	(0.010)	(0.052)	(0.025)	(0.010)	(0.052)
CU	-0.022	0.049***	-0.038	-0.022	0.053***	-0.040
	(0.048)	(0.017)	(0.041)	(0.048)	(0.017)	(0.040)
lreshatp				-0.167**	-0.050**	-0.110
				(0.083)	(0.023)	(0.095)
N	82324	82324	82324	81497	81497	81497
r2_p	0.993	0.264	0.260	0.993	0.264	0.261
chi2	16.346	71.437	1.375	25.564	77.065	2.518
rmse	0.187	0.197	0.559	0.186	0.198	0.560

Notes: This table reports estimates of the effects of the stock of migrants (in natural logarithms) on the value of trade (Value), extensive margin (EM) and intensive margin (IM). The EM and IM are computed according to equations (2) and (3), respectively. The estimates in columns (1)-(3) are obtained from specification (1). In addition, the estimates in columns (4)-(6) are obtained from specification (6). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp denotes the residuals obtained from the first step estimation, specification (7). The dependent variables in each specification are in levels and the estimator is PPML. * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications. See text for further details.

Table 3: PPML and CFA with Gravity Controls

	(1) Value	(2) EM	(3) IM	(4) Value	(5) EM	(6) IM
lnmig	0.209***	0.101***	0.072***	0.213***	0.106***	0.070***
	(0.019)	(0.005)	(0.020)	(0.019)	(0.005)	(0.022)
RTA	0.481***	0.091***	0.223*	0.474***	0.083***	0.226*
	(0.083)	(0.023)	(0.117)	(0.085)	(0.023)	(0.119)
CU	0.067	0.001	-0.007	0.068	-0.001	-0.003
	(0.082)	(0.024)	(0.091)	(0.082)	(0.024)	(0.092)
lndistance	-0.369***	-0.113***	-0.484***	-0.370***	-0.114***	-0.484***
	(0.037)	(0.015)	(0.046)	(0.038)	(0.016)	(0.047)
colony	-0.014	0.106***	-0.066	-0.013	0.103***	-0.067
	(0.079)	(0.032)	(0.094)	(0.081)	(0.033)	(0.095)
contiguity	0.461***	-0.183***	0.667***	0.456***	-0.188***	0.669***
	(0.078)	(0.034)	(0.082)	(0.079)	(0.034)	(0.082)
comlanguage	0.023	0.118***	-0.153	0.020	0.107***	-0.145
	(0.090)	(0.030)	(0.116)	(0.090)	(0.030)	(0.116)
lreshatp				-0.071**	-0.043***	0.022
				(0.030)	(0.008)	(0.026)
N	81795	81795	81795	80798	80798	80798
r2_p	0.940	0.243	0.196	0.940	0.243	0.197
chi2	1439.531	1205.432	782.127	1402.308	1224.423	772.884
rmse	0.624	0.364	1.198	0.625	0.363	1.197

Notes: This table reports estimates of the effects of the stock of migrants (in natural logarithms) on the value of trade (Value), extensive margin (EM) and intensive margin (IM). The EM and IM are computed according to equations (2) and (3), respectively. The estimates in columns (1)-(3) are obtained from specification (1) when the pair FE are replaced by gravity controls (the log of geographical distance, and dummy variables for having a colonial relationship, sharing a border and speaking a common language). Estimates in columns (4)-(6) are obtained from specification (6) also with gravity controls. RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp denotes the residuals obtained from the first step estimation, specification (7). The dependent variables in each specification are in levels and the estimator is PPML. * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications. See text for further details.

In Table 4 we introduce the squared term of the logged migration stock, as specified in model 9. In columns (1)-(3), we show the results without the CFA, and in columns (4)-(6) the CFA adding the residuals and the residuals squared to control for endogeneity (Wooldridge, 2015). Only the estimated coefficient for the EM of the squared term (*lnmig2*) is statistically significant and negatively signed, but small in magnitude. This indicates an inverted U-shaped relationship between the EM and the stock of migrants. The marginal effect computed at the mean, taking the estimate in column (5) indicates that for each 10% increase in migration stock, the EM increases by 9.4% (0.186-5.932*0.008*2). The estimated turning point is at almost 158,000 migrants ($exp(11.970)$). It is important to note that less than 2% of our estimation sample lies above the turning point.

Table 4: Gravity Model with CFA and non-linearities in migration

	(1) Value	(2) EM	(3) IM	(4) Value	(5) EM	(6) IM
lnmig	0.088**	0.129***	-0.071	0.264***	0.186***	0.042
	(0.042)	(0.010)	(0.054)	(0.093)	(0.024)	(0.102)
lnmig2	-0.004	-0.007***	0.004	-0.005*	-0.008***	0.004
	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)
RTA	0.093***	0.050***	0.003	0.092***	0.046***	0.009
	(0.025)	(0.009)	(0.052)	(0.025)	(0.010)	(0.052)
CU	-0.023	0.046***	-0.034	-0.023	0.049***	-0.036
	(0.047)	(0.016)	(0.041)	(0.048)	(0.016)	(0.039)
lreshatp				-0.156*	-0.053**	-0.095
				(0.083)	(0.022)	(0.095)
lreshatp2				0.020**	0.001	0.020**
				(0.009)	(0.002)	(0.010)
N	82324	82324	82324	81497	81497	81497
r2_p	0.993	0.264	0.260	0.993	0.264	0.261
chi2	22.717	199.775	3.361	36.309	214.715	11.464
rmse	0.187	0.197	0.559	0.186	0.198	0.560

Notes: This table reports estimates of the effects of the stock of migrants (in natural logarithms) on the value of trade (Value), extensive margin (EM) and intensive margin (IM). The EM and IM are computed according to equations (2) and (3), respectively. The estimates in columns (1)-(3) are obtained from specification (1) with the squared term of the log of migration added (lnmig2). The estimates in columns (4)-(6) are obtained from specification (9). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp (lreshat2) denotes the residuals (its squared term) obtained from the first step estimation, specification (7). The dependent variables in each specification are in levels and the estimator is PPML. * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications. See text for further details.

Another potential heterogeneity that we consider is whether the effect of migration on trade varies by the level of income per capita in the origin or host country. The main results are shown in Table 5. Interestingly, the value of trade, as well as its margins, are affected by this source of heterogeneity. For instance, the effect of migration on the value of trade and EM decreases with both the GDP per capita of the origin and destination countries, whereas the effect on the IM is not statistically significant, as in the model without interactions. Since the GDP variables are centered around their means, the interpretation is as follows. Taking the estimates in column (5), although an increase of 10% of the migration stock increases the EM of trade by almost 0.8% for the average GDPs in origin and destination countries, this effect decreases for countries with GDPs above the mean values in the sample. Migrants seem to serve as a way to counteract weak institutions or weak contract enforcement, usually attributed to lower-income countries according to standard governance indicators, like those from the World Bank (World Governance Indicators).

Table 5: Heterogeneities across income levels

	(1) Value	(2) EM	(3) IM	(4) Value	(5) EM	(6) IM
lnmig	0.036**	0.024***	-0.016	0.181*	0.079***	0.118
	(0.018)	(0.005)	(0.024)	(0.100)	(0.025)	(0.106)
lnmig*GDPHo	-0.017***	-0.026***	0.012	-0.021**	-0.031***	0.013
	(0.007)	(0.002)	(0.009)	(0.009)	(0.003)	(0.015)
lnmig*GDPHd	-0.025**	-0.035***	-0.012	-0.006	-0.038***	-0.011
	(0.012)	(0.003)	(0.016)	(0.016)	(0.005)	(0.022)
RTA	0.097***	0.048***	0.011	0.078***	0.037***	0.016
	(0.024)	(0.009)	(0.052)	(0.025)	(0.009)	(0.041)
CU	-0.020	0.052***	-0.046	0.004	0.064***	0.043
	(0.046)	(0.015)	(0.041)	(0.051)	(0.019)	(0.050)
lreshatp				-0.188*	-0.055**	-0.155
				(0.100)	(0.026)	(0.106)
lreshatpo				0.034**	0.016***	0.010
				(0.015)	(0.005)	(0.019)
lreshatpd				-0.053	0.041***	-0.064
				(0.035)	(0.007)	(0.041)
N	80880	80880	80880	69646	69646	69646
r2_p	0.993	0.262	0.259	0.994	0.260	0.260
chi2	38.531	491.850	2.838	37.340	460.633	14.266
rmse	0.185	0.194	0.554	0.178	0.188	0.517

Notes: This table reports estimates of the effects of the stock of migrants (in natural logarithms) on the value of trade (Value), extensive margin (EM) and intensive margin (IM). The EM and IM are computed according to equations (2) and (3), respectively. The estimates in columns (1)-(3) are obtained from specification (1) with interactions between the log of income per capita in origin (GDPHo) (destination countries, GDPHd) and the log of migration added. The estimates in columns (4)-(6) are obtained from specification (10). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp (o,d) denotes the residuals (its interactions with GDPHo, GDPHd) obtained from the first step estimation, specification (7). The dependent variables in each specification are in levels and the estimator is PPML. * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications. See text for further details.

Next, we estimate the gravity model considering the conditional expectiles of the distribution of imports. The main results obtained after combining the CFA with the expectiles-PPML estimator are shown in Table 6 for the value of trade and in Tables 7 and 8 for the extensive and intensive margins, respectively. Tables A.2 and A.3 show the corresponding results for the value and the EM using PPML, without the CFA. Results for the IM are not shown since they are identical to those in Table 8, where the residuals from the first step are not significant. As can be observed in column (3) the results mimic those of Table 2 in column (4) for the 0.5 conditional expectile of the distribution. However, for lower expectiles, more so for the lowest (0.1) the coefficient is slightly higher and a decreasing path could be observed when moving to higher expectiles. Nevertheless, the variation is small and considering the confident bands the effect of immigration on the value of trade does not differ significantly across the distribution of trade flows. The coefficients estimated for immigration when the EM is the dependent variable show a similar decreasing path across the distribution of trade, as can be seen in Table 7, but in this case the effect on the highest tile of the distribution is significantly lower than in the lowest expectile of the distribution of imports. Finally, no significant effects are observed on the intensive margin, as reported in Table 8.

Table 6: CFA and expectiles (Value)

	(1)	(2)	(3)	(4)	(5)
Expectile	0.1	0.25	0.5	0.75	0.9
lnmig	0.206**	0.192**	0.188**	0.184**	0.177***
	(0.088)	(0.086)	(0.082)	(0.073)	(0.066)
RTA	0.090***	0.092***	0.093***	0.090***	0.088***
	(0.028)	(0.026)	(0.025)	(0.024)	(0.024)
CU	-0.036	-0.031	-0.022	-0.012	-0.003
	(0.050)	(0.051)	(0.048)	(0.043)	(0.039)
lreshatp	-0.189**	-0.173**	-0.167**	-0.161**	-0.153**
	(0.088)	(0.087)	(0.083)	(0.075)	(0.067)
N	81497	81497	81497	81497	81497
r2_p	0.996	0.994	0.993	0.994	0.996
chi2	21.109	22.444	25.564	28.589	28.118
rmse	0.189	0.196	0.186	0.162	0.127

Notes: This table reports estimates of the effects of the stock of migrants (in natural logarithms) on the conditional distribution of the value of trade (Value). The estimates are obtained from specification (11), adding the residuals (lreshatp) from the first step specification (7) and for five different expectiles of the conditional distribution of imports (0.1, 0.25, 0.5, 0.75 and 0.9). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

Table 7: CFA and expectiles (EM)

	(1)	(2)	(3)	(4)	(5)
Expectile	0.1	0.25	0.5	0.75	0.9
lnmig	0.086***	0.078***	0.071***	0.068***	0.066***
	(0.025)	(0.024)	(0.022)	(0.020)	(0.019)
RTA	0.044***	0.048***	0.049***	0.046***	0.041***
	(0.010)	(0.010)	(0.010)	(0.009)	(0.008)
CU	0.056***	0.056***	0.053***	0.049***	0.045***
	(0.019)	(0.019)	(0.017)	(0.015)	(0.013)
lreshatp	-0.061**	-0.055**	-0.050**	-0.049**	-0.049**
	(0.025)	(0.025)	(0.023)	(0.021)	(0.020)
N	81497	81497	81497	81497	81497
r2_p	0.324	0.295	0.264	0.238	0.221
chi2	70.079	73.465	77.065	81.517	80.322
rmse	0.226	0.217	0.198	0.172	0.141

Notes: This table reports estimates of the effects of the stock of migrants (in natural logarithms) on the conditional distribution of the extensive margin (EM) of trade. The EM is computed according to equation (2). The estimates are obtained from specification (11), adding the residuals (lreshatp) from the first step specification (7) and for five different expectiles of the conditional distribution of imports (0.1, 0.25, 0.5, 0.75 and 0.9). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

Table 8: CFA and expectiles (IM)

	(1)	(2)	(3)	(4)	(5)
	0.1	0.25	0.5	0.75	0.9
lnmig	0.113	0.103	0.097	0.087	0.080
	(0.118)	(0.102)	(0.091)	(0.082)	(0.074)
RTA	0.008	0.008	0.008	0.017	0.026
	(0.050)	(0.051)	(0.052)	(0.049)	(0.043)
CU	-0.040	-0.041	-0.040	-0.041	-0.035
	(0.045)	(0.044)	(0.040)	(0.036)	(0.033)
lreshatp	-0.120	-0.114	-0.110	-0.100	-0.091
	(0.120)	(0.106)	(0.095)	(0.085)	(0.078)
N	81497	81497	81497	81497	81497
r2_p	0.291	0.271	0.261	0.260	0.267
chi2	1.750	2.069	2.518	3.071	3.377
rmse	1.267	0.762	0.560	0.425	0.312

Notes: This table reports estimates of the effects of the stock of migrants (in natural logarithms) on the intensive margin (IM) of trade. The IM is computed according to equation (3). The estimates are obtained from specification (11), adding the residuals (lreshatp) from the first step obtained from specification (7). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

Finally, we present estimates by type of good following the classification proposed by Rauch (1999). Goods are classified according to their characteristics as differentiated, referenced priced and homogeneous (goods traded in an organized exchange). The results using a CFA-PPML framework are shown in Table 9 for the value, the EM and the IM for the main specification with pair FE. Results in Table 9 show that the estimated effects on reference priced goods and homogeneous goods for the EM are higher in magnitude that the effect for differentiated goods. The equivalent table but without including the country-pair fixed effects is available in the Appendix (Table A.4).

As argued for example in Felbermayr and Toubal (2012), the coefficient estimates for differentiated, reference-priced and homogeneous goods cannot be directly compared, as the elasticity of substitution differs across these three types of goods. The estimated coefficients are a structural combination of the "true" reactions to trade cost changes (or migration, our case) and the substitution responsiveness. Broda and Weinstein

(2006), for example, highlight that the elasticity of substitution is high for homogeneous goods, comparably low for differentiated goods, with those of reference-priced goods in between. Homogeneous and reference-priced goods have thus *mechanically* higher coefficient estimates than differentiated goods, and this does not necessarily imply that their reaction to changes in trade costs are stronger.

Table 9: CFA and types of Goods

	(1) Value diff	(2) EM diff	(3) IM diff	(4) Value ref	(5) EM ref	(6) IM ref	(7) Value hom	(8) EM hom	(9) IM hom
lnmig	0.038	0.045**	-0.165**	0.362***	0.100***	0.077	0.032	0.154***	0.251*
	(0.071)	(0.023)	(0.070)	(0.111)	(0.033)	(0.113)	(0.186)	(0.054)	(0.141)
RTA	0.103***	0.027***	0.027	0.137***	0.061***	0.019	0.011	0.074**	-0.037
	(0.032)	(0.008)	(0.050)	(0.037)	(0.017)	(0.058)	(0.064)	(0.029)	(0.103)
CU	-0.046	0.035**	0.008	-0.046	0.094***	-0.072	0.006	0.113***	0.001
	(0.053)	(0.017)	(0.032)	(0.041)	(0.022)	(0.048)	(0.107)	(0.039)	(0.098)
lreshatp	-0.011	-0.026	0.148**	-0.307***	-0.075**	-0.078	-0.028	-0.119**	-0.256*
	(0.073)	(0.023)	(0.072)	(0.110)	(0.034)	(0.117)	(0.188)	(0.056)	(0.145)
N	81415	81415	81415	79825	79825	79825	77881	77881	77881
r2_p	0.996	0.278	0.323	0.988	0.343	0.235	0.975	0.268	0.280
chi2	15.201	41.793	7.439	36.244	58.262	3.347	0.101	40.450	3.268
rmse	0.148	0.193	0.512	0.230	0.226	0.789	0.391	0.583	0.957

Notes: This table reports estimates of the effects of the stock of migrants (in natural logarithms) on the value of trade (Value), extensive margin (EM) and intensive margin (IM). The EM and IM are computed according to equations (2) and (3), respectively. The estimates are obtained from specification (6) for three types of goods: differentiated (diff), reference priced (ref) and homogeneous (hom) goods, according to Rauch (1999) classification. RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp denotes the residuals obtained from the first step estimation, specification (7). * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

To facilitate interpretation of the migration coefficients, we compute their tariff-equivalent effect using standard tariff elasticities from the gravity literature. The comparison quantifies the tariff change that would generate the same proportional effect on trade as an increase in bilateral migration. We adopt benchmark elasticities of substitution of 4 for differentiated goods and 25 for reference-priced goods and for homogeneous goods, according to Felbermayr and Toubal (2012). Following the authors, to calculate the tariff equivalent we take the migration coefficient and divide it by the elasticity of substitution minus 1. The elasticity of 0.045 for differentiated goods corresponds to a tariff reduction of 1.5%. For reference-priced goods, the coefficient of 0.10 implies a tariff-equivalent reduction of 0.42%. Finally, for homogeneous goods, the coefficient of 0.154 translates into a tariff-equivalent reduction of 0.64%. These magnitudes indicate that migration flows generate economically meaningful trade cost reductions, with the highest effects observed for differentiated goods, despite their small migration elasticity.

Table 10: Tariff-Equivalent Effects of Migration on the EM

Product Type	**Migration Elasticity** (ϵ_M)	**Elasticity of substitution** (σ_τ)	**Tariff Equivalent**
Differentiated goods	0.045	4	1.50%
Reference-priced goods*	0.100	25	0.42%
Homogeneous goods*	0.154	25	0.64%

Notes: The elasticities of substitution follow the standard values reported in Felbermayr and Toubal (2012). Tariff equivalents are computed as $[\epsilon_M/(\sigma_\tau - 1)] * 100$.
* Following Felbermayr and Toubal (2012), we use the same elasticity of 25 for homogeneous and reference-priced goods. They define homogeneous goods more broadly as a combination of goods traded in an organized exchange plus reference-priced goods.

5 Robustness Checks

To assess the robustness of the baseline results, we conduct a set of alternative estimations: (i) Using different measures of migration; (ii) Adding leads of the migration variable; (iii) using lagged values of migration; (iv) bootstraping standard errors. (v) Estimating an IV linear model. These estimations allow us to validate both the stability and the economic significance of the estimated effects.

Different migration variables. Figure 2 shows the robustness of the coefficient estimates to the inclusion of different migration variables for the central analysis. In addition to the main migration variable, based on citizenship (Cit), we use three other migration variables. First, the citizenship-based variable, augmented by migration data by the country of birth concept, whenever citizenship-based data is missing for all origins of a country ($Cit+$).[4] Second, a migration variable based on the country of birth concept (COB). Third, a migration variable that is based on COB, but augmented by Cit-data, whenever COB-data is missing for all origins of a given destination country. All migration variables are taken from the OECD.

[4] We do not mix the two concepts within a destination country, such that for a given destination country, the source is either by citizenship (default) or by country of birth for all origins.

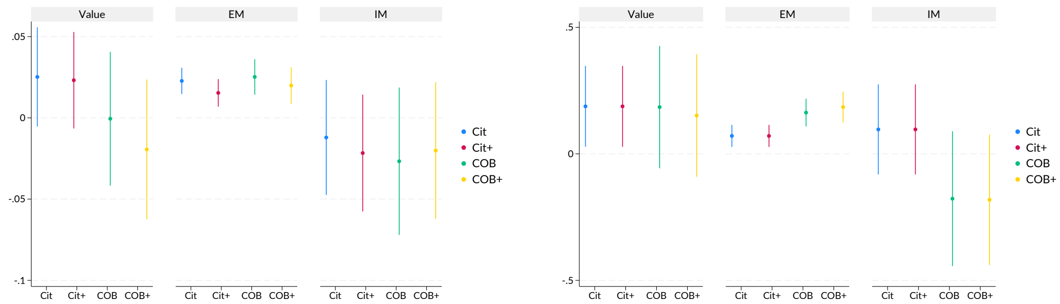
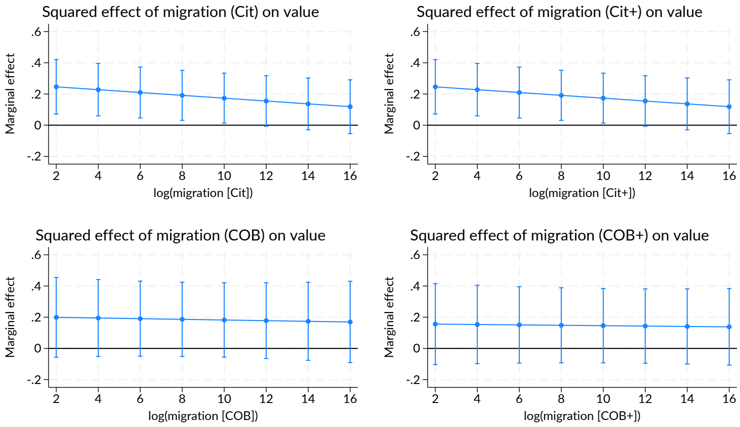
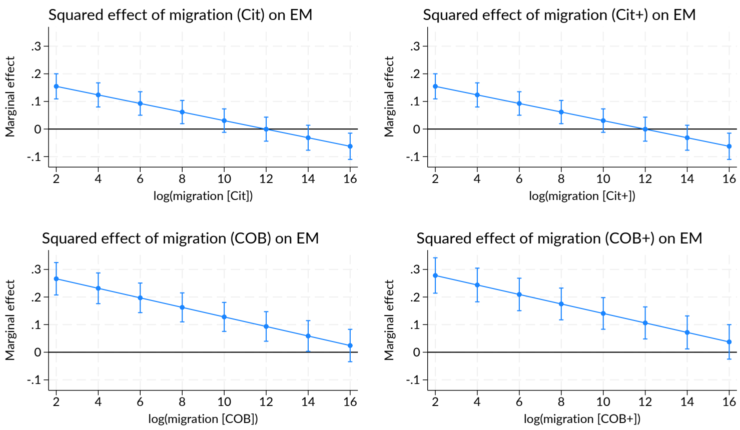


Figure 2: Robustness of results in Table 2 with respect to different migration variables without CFA (left) and with CFA (right). *Cit* is migration based on citizenship, *Cit+* based on citizenship, augmented by country of birth if *Cit* missing for a country, *COB* is migration based on the country of birth, and *COB+* based on country of birth, augmented by citizenship if missing for a country. Spikes indicate 95% confidence intervals.

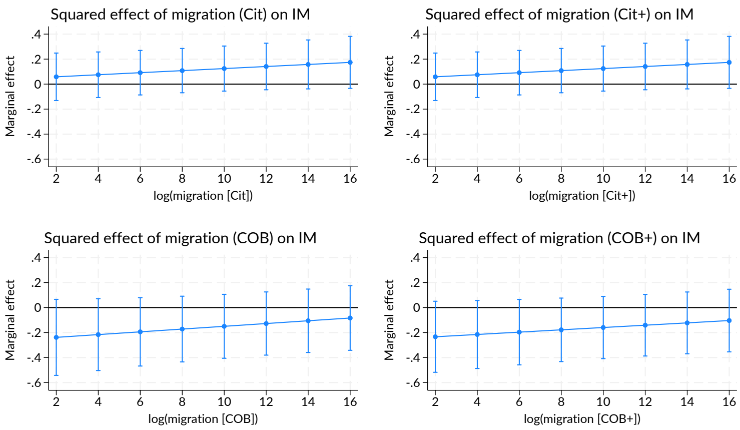
We perform similar robustness checks for the specifications with squared migration variables. Figures 3a to 3c plot the marginal effects of migration on trade values, the extensive and the intensive margin depending on different values of the migration variables. The same charts, but without using the CFA approach, are provided in the Appendix (see figures A.1a to A.1c).



(a) Robustness of marginal effects of migration for trade **values**.



(b) Robustness of marginal effects of migration for the **extensive margin**.



(c) Robustness of marginal effects of migration for the **intensive margin**.

Figure 3: Robustness of marginal effects of migration for the value of trade, the extensive margin and the intensive margin in Table 4 (columns 4 to 6) for different migration variables. Spikes indicate 95% confidence intervals.

Adding migration at t+1. Following Baier and Bergstrand (2007), in order to test for endogeneity of the migration variable, we have added migration leads as regressors to the CFA model. Table A.5 shows in columns (1)-(3) the results when the variable migration at (t+1), or alternatively in columns (4)-(6) migration at (t+2), is added to the main specification in model (6). The leads are not statistically significant in any of the specifications for value, EM and IM. Therefore, this allows us to reject the null hypothesis of reverse causality between imports and migration in our main specification.

Adding lagged migration. Table A.6 shows the results when the variable migration at (t-1) is used, instead of using the contemporaneous value. Although this is by no means a good practice to control for endogeneity with panel data, it has been used in the related literature. Columns (1)-(3) show the results with PPML (no CFA) and columns (4)-(6) with CFA. The sign and magnitude of the estimated coefficient corroborate our main results and remain very similar for the PPML approach. It is worth mentioning that the effect on the value of imports in the CFA is 6 percentage points higher in magnitude, in comparison when using the comteporaneous value of migration.

Bootstraping standard errors. Table A.7 shows the results when bootstraped standard errors are computed for the main specification with the CFA. We bootstrap standard errors with the robust option and, as expected for large samples, this does not change the results in terms of significance levels. For this reason, we do not show the additional estimation outcomes concerning non-linearities and heterogeneities.

Linear specification with instrumental variables. Table A.8 presents the results obtained by estimating a linear specification, estimated with two-stage least squares, with the same external instruments used for the CFA. The advantage of this method is that we can directly test for the validity of the instruments and whether the instruments are weak (Kleibergen-Paap rk Wald F and rk LM statistics). The results of the tests indicate that we strongly reject that the instruments are weak. In addition, the Hansen test probability indicates that we cannot reject the validity of the instruments. The estimated coefficients for value of trade, EM and IM show similar significance levels, with the EM coefficient being higher in magnitude, which could be due to the exclusion of zeroes in the estimations.

6 Conclusions

This paper reexamines the trade-creating effects of migration to OECD countries using a comprehensive dataset of origin countries and new econometric strategies that account for endogeneity, non-linearities, and heterogeneity across product types. We put a special focus on the extensive and intensive margins of trade. By estimating a structural gravity model augmented with migration stocks and implementing a control function approach, we provide robust evidence that immigration significantly promotes bilateral trade, mainly through reductions in fixed export costs, affecting the extensive margin. Our results consistently show that migrant networks matter most for expanding the *extensive margin* of trade, enabling the exchange of new products or creating new bilateral trade relationships, rather than intensifying pre-existing flows. This is especially the case when controlling for time-invariant factors that are specific to each trade relationship, a practice that has not been common in the related literature.

Additional key contributions of this study include the identification of heterogeneous effects across the distribution of trade, non-linear effects, and effects across types of goods. We find non-linear effects with respect to migration stocks only on the extensive margin. The initial migrants provide stronger trade-creation effects, which decrease as the diaspora expands. Expectile estimations indicate that the pro-trade impact of migration remains positive and relatively stable across the trade distribution, though somewhat stronger for lower expectiles of the extensive margin. Our analysis by product classification highlights further heterogeneity: reference-priced and differentiated goods display particularly strong extensive-margin responses. Translating these elasticities into tariff equivalents shows that migration generates trade-cost reductions comparable in magnitude to small tariff cuts, underscoring the economic significance of migrant networks.

Our findings also suggest that the strength of the migration-trade link varies systematically with economic development. Interactions between migration stocks and GDP per capita show that migration has the largest pro-trade impact when either the origin or destination country is less economically developed. This is consistent with the notion that migrants from lower-income economies provide valuable information and network externalities in settings where formal contracting institutions or information channels may be weaker. Importantly, the results remain robust when alternative measures of migration-based on citizenship or country of birth are employed, as well as when lagged values of migration are used, and when accounting for first step uncertainty in our sample.

This study is subject to several limitations that warrant careful interpretation of the results. First, the identification strategy relies on the exclusion restriction that migration regularization policies affect bilateral trade flows only through their impact on migration stocks. While the inclusion of country-time pairs and high-dimensional fixed effects mitigates many concerns related to confounding political or macroeconomic shocks, it cannot fully rule out the possibility that regularization programs coincide with broader policy reforms, such as changes in labor market regulation, social policy, or trade policy, which may independently influence trade outcomes. Second, although the analysis is global in scope, destination countries are restricted to OECD economies due to data availability. While this focus is appropriate given the institutional relevance of regularization programs in these countries, the external validity of the findings for South-South trade or for emerging-market destinations remains an open question. Extending the analysis to a broader set of destination countries as comparable data become available would be a valuable direction for future research. Finally, the international trade literature acknowledges the importance of dynamic gravity specifications (Martínez-Zarzoso and Horsewood (2005), Anderson and Yotov (2025), among others), which allow for persistence in trade relationships, adjustment costs, and gradual responses to policy changes. Incorporating lagged trade flows can capture path dependence arising from sunk export costs, learning-by-doing, and the gradual formation of trade networks. Nevertheless, combining dynamic specifications with high-dimensional fixed effects and endogenous regressors would require alternative estimation strategies and instruments to address potential bias. Therefore we leave this avenue for further research.

Taken together, the evidence supports the view that immigration facilitates global economic integration by lowering informational and cultural barriers to trade. By reducing fixed export costs and helping firms navigate foreign markets, migrant communities enable countries to diversify their export baskets and foster deeper economic ties. For policymakers, these findings highlight the relevance of migration policy as an indirect tool for trade promotion: policies that shape the size or characteristics of migrant populations may have significant implications for international trade patterns.

Several avenues for future research emerge from our findings. First, the use of firm-level data could uncover the microeconomic mechanisms through which migrants support market entry and export expansion, allowing researchers to distinguish between productivity effects and network effects. Second, richer data on migrants' skills, occupations, and legal status, would help to clarify how different types of migrants influence trade. Finally, examining the dynamic persistence of migration-induced trade links would shed light on whether these effects are temporary or long-lasting. Such extensions would deepen our understanding of how migration contributes to international economic integration.

References

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

Anderson, J. E. and Yotov, Y. V. (2025). Gravity in transition. *Review of International Economics*.

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

Baier, S. L., Bergstrand, J. H., and Feng, M. (2014). Economic integration agreements and the margins of international trade. *Journal of International Economics*, 93(2):339–350.

Baier, S. L., Yotov, Y. V., and Zylkin, T. (2019). On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration. *Journal of International Economics*, 116:206–226.

Baldwin, R. and Taglioni, D. (2006). Gravity for dummies and dummies for gravity equations. *NBER Working Paper Series*, (12516).

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

Bergin, P. R. and Lin, C.-Y. (2009). Exchange rate regimes and the extensive margin of trade. In *NBER International Seminar on Macroeconomics*, volume 5, pages 201–227. The University of Chicago Press Chicago, IL.

Bergstrand, J. H., Clance, M. W., and Santos Silva, J. (2025). The tails of gravity: Using expectiles to quantify the trade-margins effects of economic integration agreements. *Journal of International Economics*, 157:104145.

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

Broda, C. and Weinstein, D. E. (2006). Globalization and the gains from variety. *The Quarterly journal of economics*, 121(2):541–585.

Debaere, P. and Mostashari, S. (2010). Do tariffs matter for the extensive margin of international trade? an empirical analysis. *Journal of International Economics*, 81(2):163–169.

Dunlevy, J. A. and Hutchinson, W. K. (1999). The impact of immigration on american import trade in the late nineteenth and early twentieth centuries. *The Journal of Economic History*, 59(4):1043–1062.

Dutt, P., Mihov, I., and Van Zandt, T. (2013). The effect of wto on the extensive and the intensive margins of trade. *Journal of international Economics*, 91(2):204–219.

Egger, P. H., Von Ehrlich, M., and Nelson, D. R. (2012). Migration and trade. *The world economy*, 35(2):216–241.

Elguezabal, P., Martínez-Zarzoso, I., and Nowak-Lehmann, F. (2025). Immigration as a trade facilitator: A review and new insights. *Cuadernos Económicos de ICE*, (109).

Felbermayr, G. J. and Toubal, F. (2012). Revisiting the trade-migration nexus: Evidence from new oecd data. *World Development*, 40(5):928–937.

Ferragina, A. M., Iandolo, S., and Taymaz, E. (2021). Migration and comparative advantages: new evidence on the eu-mena region. *International Journal of Manpower*, 42(5):904–934.

Fieler, A. C. and Eaton, J. (2025). The margins of trade. *Econometrica*, 93(1):129–160.

Figueiredo, E., Lima, L. R., and Orefice, G. (2020). Migration, trade and spillover effects. *Journal of Comparative Economics*, 48(2):405–421.

Foster, N. (2014). Intellectual property rights and the margins of international trade. *The Journal of International Trade & Economic Development*, 23(1):1–30.

Foster, N., Poeschl, J., and Stehrer, R. (2011). The impact of preferential trade agreements on the margins of international trade. *Economic Systems*, 35(1):84–97.

Girma, S. and Yu, Z. (2002). The link between immigration and trade: Evidence from the united kingdom. *Weltwirtschaftliches Archiv*, 138(1):115–130.

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

Hatzigeorgiou, A. (2010). Migration as trade facilitation: assessing the links between international trade and migration. *The BE Journal of Economic Analysis & Policy*, 10(1).

Hatzigeorgiou, A. and Lodefalk, M. (2015). Trade, migration and integration–evidence and policy implications. *The World Economy*, 38(12):2013–2048.

Hatzigeorgiou, A. and Lodefalk, M. (2021). A literature review of the nexus between migration and internationalization. *The Journal of International Trade & Economic Development*, 30(3):319–340.

Head, K. and Ries, J. (1998). Immigration and trade creation: econometric evidence from canada. *Canadian journal of economics*, pages 47–62.

Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The quarterly journal of economics*, 123(2):441–487.

Hummels, D. (2001). Toward a geography of trade costs. *Purdue University Working Paper*.

Hummels, D. and Klenow, P. J. (2005). The variety and quality of a nation's exports. *American economic review*, 95(3):704–723.

Koenig, P., Mayneris, F., and Poncet, S. (2010). Local export spillovers in france. *European Economic Review*, 54(4):622–641.

Lewer, J. J. and Van den Berg, H. (2009). Does immigration stimulate international trade? measuring the channels of influence. *The International Trade Journal*, 23(2):187–230.

Lin, W. and Wooldridge, J. M. (2019). Chapter 2 - testing and correcting for endogeneity in nonlinear unobserved effects models. In Tsionas, M., editor, *Panel Data Econometrics*, pages 21–43. Academic Press.

Martínez-Zarzoso, I. and Horsewood, N. (2005). Regional trading agreements: Dynamic panel data evidence from the gravity model. *University of Birmingham*.

Martínez-Zarzoso, I. and Santacreu, A. M. (2025). Cross-border patenting and the margins of international trade. *Review of International Economics*, 33(2):393–409.

Olivero, M. P. and Yotov, Y. V. (2012). Dynamic gravity: Endogenous country size and asset accumulation. *Canadian Journal of Economics*, 45(1):64–92.

Orefice, G., Rapoport, H., and Santoni, G. (2025). How do immigrants promote exports? *Journal of Development Economics*, 174:103443.

Peri, G. and 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.

Piperakis, A. S. (2011). The impact of immigration on eu-15 trade. *The International Trade Journal*, 25(5):558–580.

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

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

Silva, J. M. C. S. and Tenreyro, S. (2021). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters*, 186:108789.

White, R. (2007). Immigrant-trade links, transplanted home bias and network effects. *Applied Economics*, 39(7):839–852.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of applied econometrics*, 20(1):39–54.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, 50(2):420–445.

Yotov, Y. V. (2022). Proper use of gravity with domestic trade: A comment. *World Economy*, 45(5):1509–1526.

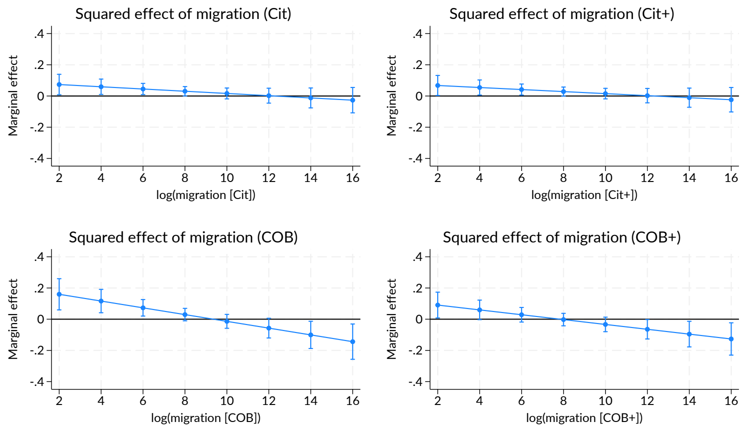
Yotov, Y. V. (2024). The evolution of structural gravity: The workhorse model of trade. *Contemporary economic policy*, 42(4):578–603.

A APPENDIX

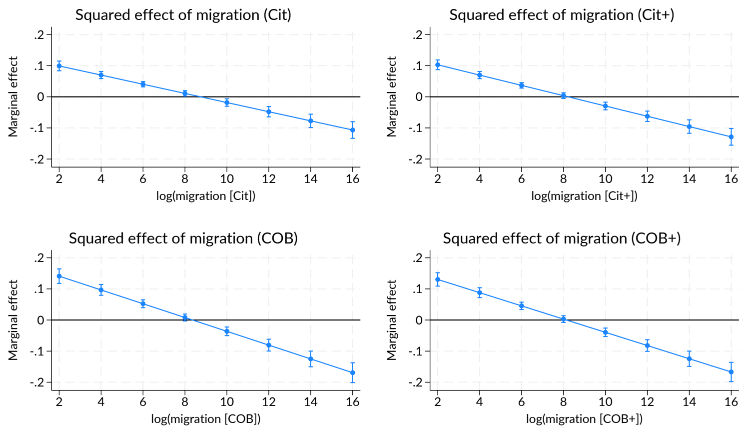
Table A.1: First-step estimations with external instruments

	(1) reg_HDFE	(2) ppml_HDFE
MonthsReg	-0.0062**	-0.0104***
	(0.002)	(0.003)
MigStockReg	-0.0001**	0.0002**
	(0.000)	(0.000)
MigRegion	0.3212***	0.2332***
	(0.022)	(0.045)
RTA	-0.0176	0.0677*
	(0.028)	(0.036)
CU	-0.0833	-0.0576
	(0.059)	(0.077)
N	82113	84357
R^2	0.984	
Pseudo-R^2		0.992
F-statistic	48.771	
χ^2		54.766
RMSE	0.410	0.232

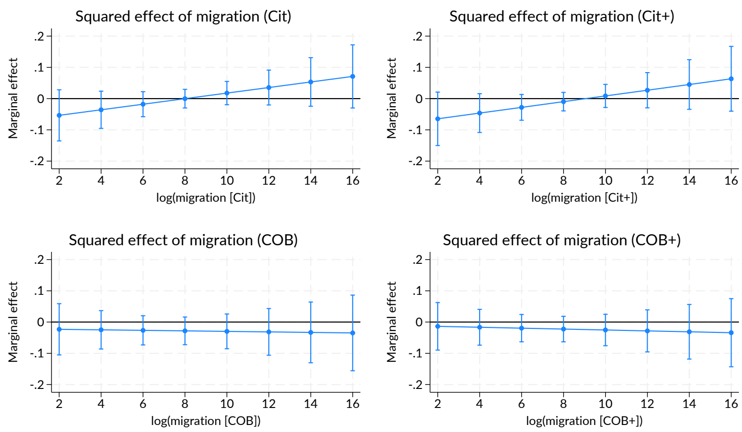
Notes: Dependent variable: Natural log of the bilateral migration stock (lnmig). Column (1) uses an OLS estimator with high dimensional fixed effects (equation 8) and column (2) the PPML estimator (equation 7). * 0.1, ** 0.05, *** 0.01 significance levels. Robust standard errors reported in parentheses. All specifications include origin-year (i,t), destination-year (j,t), and pair (ij) fixed effects (HDFE). External instruments are: Z^1_{ijt}=MonthsReg is required months of prior irregular residence in the host country to be eligible for the regularization; Z^2_{ijt}=MigStockReg is the interaction between the regularization dummy (= 1 if a regularisation program is implemented, zero otherwise) and the stock of migrants in the first year with data; Z^3_{ijt}= MigRegion denotes the average migrant stock from countries in the origin region. Additional control variables: RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union).



(a) Robustness of marginal effects of migration for trade **values**.



(b) Robustness of marginal effects of migration for the **extensive margin**.



(c) Robustness of marginal effects of migration for the **intensive margin**.

Figure A.1: Robustness of marginal effects of migration for the trade value, the extensive and intensive margins in Table 4 (columns 1 to 3) with respect to the inclusion of different migration variables. Spikes indicate 95% confidence intervals.

Table A.2: Expectiles (Value)

	(1)	(2)	(3)	(4)	(5)
Expectile	0.1	0.25	0.5	0.75	0.9
lnmig	0.023***	0.024***	0.025***	0.028***	0.028***
	(0.008)	(0.008)	(0.007)	(0.007)	(0.006)
RTA	0.091***	0.093***	0.093***	0.089***	0.085***
	(0.015)	(0.015)	(0.014)	(0.014)	(0.013)
CU	-0.044*	-0.035	-0.022	-0.011	0.001
	(0.023)	(0.024)	(0.023)	(0.022)	(0.021)
N	82324	82324	82324	82324	82324

Notes: Estimates of the effects of the stock of migrants (in natural logarithms) on the value of trade. The estimates are obtained from specification (11) for five different expectiles of the conditional distribution of imports (0.1, 0.25, 0.5, 0.75 and 0.9). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

Table A.3: Expectiles (EM)

	(1)	(2)	(3)	(4)	(5)
Expectile	0.1	0.25	0.5	0.75	0.9
lnmig	0.027***	0.025***	0.023***	0.021***	0.019***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
RTA	0.050***	0.053***	0.053***	0.049***	0.043***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
CU	0.051***	0.052***	0.049***	0.045***	0.042***
	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)
N	82324	82324	82324	82324	82324

Notes: Estimates of the effects of the stock of migrants (in natural logarithms) on the extensive margin (EM) of trade. The EM is computed according to equation (2). The estimates are obtained from specification (11) for five different expectiles of the conditional distribution of imports (0.1, 0.25, 0.5, 0.75 and 0.9). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

Table A.4: CFA, no pair FE, types of Goods

	(1) Value diff	(2) EM diff	(3) IM diff	(4) Value ref	(5) EM ref	(6) IM ref	(7) Value hom	(8) EM hom	(9) IM hom
lnmig	0.264***	0.102***	0.142***	0.225***	0.152***	0.028	0.120***	0.112***	-0.018
	(0.019)	(0.005)	(0.024)	(0.024)	(0.008)	(0.020)	(0.028)	(0.008)	(0.022)
RTA	0.611***	0.033	0.319***	0.464***	0.187***	0.238**	0.090	0.221***	0.002
	(0.091)	(0.023)	(0.094)	(0.105)	(0.037)	(0.095)	(0.121)	(0.036)	(0.134)
CU	0.005	-0.019	0.164*	0.124	-0.027	0.029	0.288*	-0.041	0.059
	(0.080)	(0.023)	(0.089)	(0.104)	(0.030)	(0.090)	(0.166)	(0.037)	(0.115)
lndist	-0.271***	-0.069***	-0.395***	-0.509***	-0.215***	-0.437***	-0.806***	-0.231***	-0.554***
	(0.047)	(0.015)	(0.053)	(0.046)	(0.022)	(0.050)	(0.067)	(0.022)	(0.065)
colony	-0.050	0.093***	-0.020	0.098	0.154***	0.075	0.123	0.147***	-0.166
	(0.097)	(0.035)	(0.094)	(0.112)	(0.037)	(0.098)	(0.136)	(0.037)	(0.138)
contig	0.328***	-0.225***	0.561***	0.434***	-0.131***	0.750***	0.731***	-0.122***	1.007***
	(0.081)	(0.034)	(0.089)	(0.094)	(0.042)	(0.083)	(0.144)	(0.046)	(0.142)
comlang_off	0.028	0.136***	-0.093	-0.058	-0.021	-0.084	0.177	0.062	0.107
	(0.077)	(0.031)	(0.107)	(0.126)	(0.039)	(0.139)	(0.150)	(0.041)	(0.137)
lreshatp	-0.105***	-0.046***	-0.094***	-0.116***	-0.053***	-0.043	0.051	-0.023*	0.034
	(0.031)	(0.007)	(0.034)	(0.040)	(0.012)	(0.038)	(0.056)	(0.014)	(0.034)
N	80794	80794	80794	80635	80635	80635	80319	80319	80319
r2_p	0.966	0.257	0.280	0.922	0.323	0.150	0.867	0.228	0.175
chi2	1469.762	834.423	800.106	1128.730	1529.296	714.018	829.480	1154.018	504.549
rmse	0.469	0.355	1.017	0.705	0.456	1.510	0.994	0.757	1.688

Notes: The estimates are obtained from specification (6) for three types of goods: differentiated (diff), reference priced (ref) and homogeneous (hom) goods, according to Rauch (1999) classification. RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp denotes the residuals obtained from the first step estimation, specification (7). * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. Origin-time (i,t), destination-time (j,t) fixed effects in all specifications.

Table A.5: CFA adding leads of migration

	(1) Value b/se	(2) EM b/se	(3) IM b/se	(4) Value b/se	(5) EM b/se	(6) IM b/se
lnmig	0.238***	0.066***	0.101	0.264***	0.072***	0.144
	(0.083)	(0.025)	(0.103)	(0.085)	(0.025)	(0.105)
RTA	0.083***	0.038***	0.009	0.076***	0.035***	0.024
	(0.027)	(0.010)	(0.050)	(0.028)	(0.011)	(0.047)
CU	-0.016	0.057***	-0.037	-0.014	0.053***	-0.039
	(0.049)	(0.017)	(0.039)	(0.049)	(0.017)	(0.038)
lreshatp	-0.232***	-0.048*	-0.084	-0.269***	-0.053**	-0.129
	(0.084)	(0.025)	(0.106)	(0.086)	(0.026)	(0.109)
lnmig(t+1)	0.014	0.002	-0.018			
	(0.015)	(0.005)	(0.022)			
lnmig(t+2)				0.019	-0.001	-0.024
				(0.016)	(0.005)	(0.021)
N	73983	73983	73983	69058	69058	69058
r2_p	0.994	0.264	0.260	0.994	0.264	0.259
chi2	24.867	59.174	2.218	23.972	53.370	3.802
rmse	0.176	0.193	0.539	0.171	0.191	0.528

Notes: The estimates are obtained from specification (6) adding leads of the migration variable. * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp denotes the residuals obtained from the first step estimation, specification (7). Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

Table A.6: PPML and CFA using lagged migration

	(1)	(2)	(3)	(4)	(5)	(6)
	Value	EM	IM	Value	EM	IM
	b/se	b/se	b/se	b/se	b/se	b/se
lnmig(t-1)	0.021	0.021***	-0.027	0.246***	0.080***	0.084
	(0.016)	(0.004)	(0.020)	(0.079)	(0.022)	(0.089)
RTA	0.079***	0.052***	0.016	0.077***	0.049***	0.021
	(0.025)	(0.009)	(0.043)	(0.025)	(0.009)	(0.043)
CU	-0.022	0.059***	-0.027	-0.022	0.061***	-0.028
	(0.048)	(0.019)	(0.045)	(0.048)	(0.019)	(0.045)
lreshatp(t-1)				-0.232***	-0.062***	-0.114
				(0.081)	(0.023)	(0.092)
N	78596	78596	78596	77895	77895	77895
r2_p	0.993	0.261	0.260	0.993	0.262	0.261
chi2	11.497	68.722	2.992	24.888	75.067	4.324
rmse	0.185	0.194	0.531	0.185	0.195	0.532

Notes: The estimates are obtained from specification (6) taking 1 lag of the logged migration variable. * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp (t-1) denotes the lagged residuals obtained from the first step estimation, specification (7). Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

Table A.7: CFA with bootstrapped standard errors

	(1)	(2)	(3)
	Value	EM	IM
	b/se	b/se	b/se
lnmig	0.188***	0.071***	0.097
	(0.048)	(0.012)	(0.064)
RTA	0.093***	0.049***	0.008
	(0.019)	(0.006)	(0.040)
CU	-0.022	0.053***	-0.040
	(0.030)	(0.008)	(0.030)
lreshatp	-0.167***	-0.050***	-0.110
	(0.050)	(0.012)	(0.069)
N	81497	81497	81497
r2_p	0.993	0.264	0.261
chi2	56.648	225.101	5.034
rmse	0.186	0.198	0.559

Notes: The estimates are obtained from specification (6) using 1000 replications for the bootstrap. * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses). RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). lresthatp denotes the residuals obtained from the first step estimation, specification (7). Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications.

Table A.8: Gravity Model Linear Specification with Instruments

	(1)	(2)	(3)
	ln Value	ln EM	ln IM
	b/se	b/se	b/se
lnmig	0.139*	0.177***	-0.041
	(0.080)	(0.052)	(0.088)
RTA	0.075*	0.047*	0.028
	(0.043)	(0.025)	(0.040)
CU	0.009	0.106***	-0.097*
	(0.053)	(0.037)	(0.053)
N	79385	79385	79385
rkf	81.861	81.861	81.861
idstat	193.751	193.751	193.751
Hansen stat	0.426	1.212	1.104
Hansen prob	0.808	0.546	0.576

Notes: The estimates are obtained from a linear specification of the gravity model estimated with the stata command ivreghdfe and using the same instrumental variables for migration as in the control function approach: Z^1_{ijt}=MonthsReg is required months of prior irregular residence in the host country to be eligible for the regularization; Z^2_{ijt}=MigStockReg is the interaction between the regularization dummy (= 1 if a regularisation program is implemented, zero otherwise) and the stock of migrants in the first year with data; Z^3_{ijt}= MigRegion denotes the average migrant stock from countries in the origin region. * 0.1 ** 0.05 *** 0.01 significance levels, standard errors (in parentheses) clustered at the country pair level. RTA (CU) takes the value of 1 if countries are members of the same free trade agreement (currency union). Origin-time (i,t), destination-time (j,t) and pair (ij) fixed effects in all specifications. rkf denotes Kleibergen-Paap rk Wald F statistic and idstat is the Kleibergen-Paap rk LM statistic. See column (1) in Table A.1 for the first step linear estimation.

The Working Paper series of the Oesterreichische Nationalbank is designed to disseminate and to provide a platform for discussion of either work of the staff of the OeNB economists or outside contributors on topics which are of special interest to the OeNB. To ensure the high quality of their content, the contributions are subjected to an international refereeing process. The opinions are strictly those of the authors and do in no way commit the OeNB.

The Working Papers are also available on our website (http://www.oenb.at) and they are indexed in RePEc (http://repec.org/).

Publisher and editor	Oesterreichische Nationalbank Otto-Wagner-Platz 3, 1090 Vienna, Austria PO Box 61, 1011 Vienna, Austria www.oenb.at oenb.info@oenb.at Phone (+43-1) 40420-6666
Editor	Martin Summer
Cover Design	Information Management and Services Division

DVR 0031577

ISSN 2310-533X (Online)