UNIVERSITÀ DEGLI STUDI FIRENZE

DISEI
DIPARTIMENTO DI SCIENZE PER L'ECONOMIA E L'IMPRESA

Working Papers - Economics

Population Density and Countries' Export Performance: A Supply-Side Structural Gravity with Unilateral Variables

Luca Lodi

Working Paper N. 11/2024

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

Population Density and Countries' Export Performance: A Supply-Side Structural Gravity with Unilateral Variables

Luca Lodi

Università degli Studi di Firenze

July 15, 2024

Abstract

This paper analyzes the effect of population density on international trade using a theoretical and empirical framework. It builds on the works of Allen and Arkolakis (2014), Allen et al. (2020) and Freeman et al. (2021) to derive a structural gravity model that identifies the impact of country-specific features on bilateral exports. The study interprets Heid et al. (2020) and Freeman et al. (2021) empirical approaches. Focusing on population density as a component of productivity and agglomeration, it explores how density influences country specialization and comparative advantages in labor-intensive or natural resource-dependent industries. The research suggests that population density significantly impacts manufacturing but negatively affects mining, with further investigation needed for agriculture, forestry, and fisheries

Keywords: structural gravity model, economic geography, population density

JEL: **F00 F12 F14 F19**

Acknowledgement: I am grateful to Prof. Luca De Benedictis, Prof. Ariell Reshef and Dr. Gianluca Santoni for the help and mentoring in developing and writing this piece of my PhD thesis, to Prof. Gaia Narciso, Prof. Mario Larch and Prof. Lionel Fontagné for the comments and suggestions during the revision and the dissertation of the thesis. And also to people organising and participating in the GSIE seminar (Paris 1), the ETSG at the University of Surrey, the 15th FIW conference at the University of Wien and the DISEI Seminars at the University of Florence (in particular Prof. Giorgio Ricchiuti). All errors are mine.

1 Introduction

The main objective of the paper is to describe theoretically and empirically the effect of population density on international trade and does it through a theoretical framework based on Allen and Arkolakis (2014) and Allen et al. (2020) to derive a structural gravity setting that identifies the effect of country-specific features on bilateral exports and assess their contribution through the computation of specific parameter. The model provides a theoretical interpretation of the empirical approach developed by Heid et al. (2021) to measure unilateral policy variables' effects in a theoretically grounded structural gravity model that overcomes the perfect collinearity with the importer and exporter fixed effects, needed to control for multilateral resistance terms (Baldwin and Taglioni, 2006). Moreover, our theoretical framework describes how domestic and international trade shares affect the value of the estimated coefficient.

The focus on population density as a component of productivity and as a determinant of agglomeration forces,[1] allows testing *the hypothesis*, that population density also affects country specialization. The abundance of production factors and their spatial distribution within a country have consequences on comparative advantages and gains from trade.[2] Consequently, countries with great population density specialize in more labour-intensive economic activities, while others either specialize in more land (natural resources)-intensive industries or diversify their production.

The analysis investigates the contribution of population concentration on the supply side of the domestic economy, enriching the debate of density and agglomeration (mainly studied at the urban and regional levels) to understand the implication of the spatial distribution of the production factors at the macro level for different sectors, and for trade flows specialization. We propose a method to measure density sensitivity as in Moscona and Levy (2022), but our approach uses bilateral trade flows, and it also allows us to quantify the contribution of population density as productivity

[1] Duranton and Puga (2004),Rosenthal and Strange (2004), Allen and Arkolakis (2014), Bakker et al. (2021), Moscona and Levy (2022).

[2] Courant and Deardorff (1992)Courant and Deardorff (1993).

fundamental. This yields results in line with Combes et al. (2012) considering the macro-level and the international trade flows.

The framework has elements in common with the *New Economic Geography*, [3] but models geography, as Allen and Arkolakis (2014) and Allen et al. (2020). One of the originalities of the proposed approach is to bridge the literature on structural gravity to estimate unilateral variables,[4] the *New Economic Geography* and quantitative spatial economics (Redding and Rossi-Hansberg, 2017).

The findings suggest that population density matters more in labour-intensive industries (Manufacturing). A 1% change in population density leads to a 0.3% of exports with respect to domestic sales, while the direct effect is slightly higher at 0.5%. Whereas natural resources that depend on activities, such as mining, show a negative correlation, both for relative impact and for the direct effect. Furthermore, we evaluate the potential biases of the aggregate manufacturing trade by examining each industry in this sector. Our theoretical interpretation helps us to assess whether aggregation generates biases, as discussed in Redding and Weinstein (2019) and Breinlich et al. (2022). In this work, using total bilateral trade in manufacturing does not change the main findings. However, disaggregated data provide a more complete understanding of the nexus of exports and population density. The contribution to Agriculture, Forestry and Fisheries is not robustly estimated since signs and significance are not consistent. These controversial results offer the opportunity for a new stand-alone analysis looking at differences at the industry level.

The discussion starts with a review of the contribution to the related literature (Section 2) and continues with the description of the theoretical model (Section 4) which is the guideline for the description of the empirical strategy (Section 5) and the dis-

[3] This branch of the literature included scale effects related to Marshallian externalities, the cost of moving goods between locations and different market structures. Hence, the modelling features regard CES preferences as Dixit and Stiglitz (1977), icebergs trade costs Samuelson (1952), and the evolution of the Computer (meaning the possibility to perform computer simulation or solution of the model, especially when the complexity of the frameworks grows). In a vast literature, some of the relevant references are Krugman (1979), Krugman (1980) Fujita et al. (1999)

[4] Sellner (2019), Heid et al. (2021), Freeman et al. (2021).

cussion of the results in Section 7.

2 Literature Review

The literature examining the intricate interplay between geography, population, and trade reveals a complex set of relationships. While existing research lacks a focused exploration of population density, it becomes imperative to elucidate how density intersects with agglomeration economies, productivity dynamics, and the repercussions of spatially uneven distribution of production factors.

The intricate nexus between population and exports unfurls a tapestry of conflicting findings. Different levels of analysis may produce heterogeneous results since some works use trade as its total and others refer to the bilateral flows from country to country. Econometric specifications with dyadic data offer the possibility to consider translational linkages, which is crucial in international trade, but it is also more challenging from the econometric point of view. Our work wants to provide a theoretical framework and empirical strategies that may help overcome the issues of measuring and interpret empirical evidence obtained from bilateral trade flows.

Among the articles looking at total country trade, a seminal framework introduced by Redding and Venables (2004a) stands as a pivotal theoretical and empirical cornerstone, delving into the impact of geography on a country's exports. This framework skillfully navigates multilateral resistances and provides a consistent methodology to discern country-specific features. Notably, Redding and Venables (2004b) presents a comparable analysis centered on inequality. The significance of this framework lies in its ability to untangle the intricate threads of multilateral resistances and isolate country-specific attributes, marking a watershed moment in research.[5] Lately, using a similar approach, Bleaney and Neaves (2013) attempted to unpack the enigma of density's impact on trade openness. Although their cross-sectional analysis of the

[5] In these works, bilateral trade flows serve to measure market access measures, both for importer and exporters, which are, subsequently, included in the main estimates where trade is not dyadic

country-level effect doesn't account for temporal variations, it intriguingly contests the expected negative influence of population concentration on trade.

An article that uses bilateral trade to measure the effect of trade is Yamarik and Ghosh (2005), where the authors assessed new potential determinants of international trade through a naive gravity using bilateral trade flows. They used a measure of relative density (the difference in absolute values between the exporter and importer density) included together with variables concerning development levels, linguistic and colonial ties, geography, common currency and regional trade agreements.[6] This difference represents the relative land endowments between the two countries and positively impacts bilateral trade. However, the absence of theoretical underpinnings leaves this relationship in a realm of empirical intrigue. Sellner (2019) warn against using differences as bilateral measures because these may not represent pure dyadic variables. We add that in the case of population density, looking at the differences between importers and exporters levels does not distinguish between the supply and demand side effects of this country feature. On the exporters' (supply) side it is also associated with agglomeration forces, while for the importers (demand), it relates to market absorption (more people, larger markets means more buyers/consumers).

Recently, Query (2022) studied the interplay of population density and border effects, discerning no statistically significant outcomes in the context of inter-regional and intra-regional product trade between Canadian provinces and U.S. states. In this article, the empirical strategy, similar to ours, is the one from Heid et al. (2021). The difference with our analysis is that Query (2022) extend the exercise of Anderson and van Wincoop (2003), to add density impact on the role of the international and national border on trade between two countries. Our exercise focuses on density as a deterministic productivity factor and uses international borders to control the divide between domestic and international sales at the country level and consider the

[6] The authors mentioned: the Central American Common Market (CACM), Caribbean Community (Caricom), Mercado Común del Sur (Mercosur), Australia-New Zealand Closer Economic Relations Trade Agreement (ANZCERTA), and Asian Pacific Economic Cooperation (APEC).

international trade network.

In the last few years, several relevant contributions on the linkages between trade and agglomeration have been released, but they are still unpublished, including works by Moscona and Levy (2022) and Bakker et al. (2021). The former, strongly connected to our paper, explores how domestic economic geography influences trade patterns, highlighting a country's population distribution as a significant factor in its comparative advantage. Moscona and Levy (2022) introduce a model of quantitative spatial economics (Redding, 2016), which formalize subnational level and rationalize aggregation,[7] demonstrating how variations in productivity within a country can influence its export patterns in different industries, highlighting two essential factors: differing productivity levels across regions within a country and varying benefits of clustering different industries. Moreover, they introduce methods to assess the *population density affinity* of industries and the population concentration of regions. Findings reveal that both US states and countries with concentrated populations tend to export sectors aligned with high population density affinity. This study uses trade at the country level and not bilateral flows, we provide an alternative approach to obtain density sensitivity of different sectors by exploiting dyadic data and the properties of the structural gravity model. Our results confirm the main intuition of this paper.

We want to explore the channel by which density proxies agglomeration forces. Then, we also add country level evidence to a broad literature on population density and agglomeration economies from urban and regional studies. Combes et al. (2012) found that density and large cities are crucial in determining locations' total factor productivity, but they also point out that urban density is a source of advantages and disadvantages for the economy (Duranton and Puga, 2020). Therefore, it stimulates productivity and innovation, guarantees access to decent goods and services, reduces commuting distances, fosters energy-efficient housing and transport, and makes it easier to share scarce amenities. However, density generates congestion due to crowding, high living and travel costs, greater pollution levels and more likely spread of disease.

[7] A similar modelling strategy to Ramondo et al. (2016)

Duranton and Puga (2004) describes the mechanism that leads to agglomeration (at the micro level). It arises when three conditions happen *1) sharing*: splitting the cost of indivisible facilities, assuming share risk and having a common network of buyers and sellers; *2) matching* labour supply and demand avoiding hold-up problems;*3) learning* knowledge creation and diffusion.

At the macro level, density fits with the definition of *fundamental productivity* as discussed in Costinot et al. (2012)[8] and enhances its role in determining trade patterns and specialization, comparative advantages and heterogeneous gains from trade. Moreover, the role of density captures the consequences of the uneven distribution of production factors defined as *lumpiness* by Courant and Deardorff (1992) and Courant and Deardorff (1993) by which the concentration of production factors endowments within a country matters as the abundance of them. These theoretical underpinnings examine aggregate country trade considering the spatial allocation of resources. However, the original formal setting does not allow a multi-country analysis.

Our work furnishes the tool to extract the contribution of density to productivity, similar to Combes et al. (2012) but at the macro level. And also, the country level implication of density on sector-specific productivity. Furthermore, we discuss that population density is associated with heterogeneity across industries. Evidence from Rosenthal and Strange (2004) together with the three forces above mentioned, also natural advantages, home market effects, consumption opportunities, and rent-seeking all play a part in how agglomeration and density affect trade. The article of Faggio et al. (2017) emphasizes the considerable heterogeneity across industries in the micro-foundations of agglomeration economies.

Empirically, most of the analysis focuses on the manufacturing sector [9] and the advan-

[8] Costinot et al. (2012) defines *fundamental productivity*: "captures factors such as climate, infrastructure, and institutions that affect the productivity of all producers in a given country and industry"(p.582). According to other works, (Allen and Arkolakis, 2014; Allen et al., 2020; Bakker et al., 2021; Moscona and Levy, 2022; Rosenthal and Strange, 2004) population density is part of productivity.

[9] Nakamura (1985),Rosenthal and Strange (2004),Bakker et al. (2021), Moscona and Levy

tages of dense areas. Less explored is the nexus with more natural or land-intensive sectors. Focusing on agriculture, Ricker-Gilbert et al. (2014) and Josephson et al. (2014), analyse Ethiopian and Malawian agricultural sectors, and find out the limits of Boserupian intensification.[10] Thus, higher rural densities concern smaller farms size and lower farm wages, revenue per hectare and farm income are not increasing in population density. Concerning the impact on trade, not many works deal with the role of population density on agricultural commodity exports, most of them focus on the demand side. For instance, Morrison (1984) states that physiological density[11] is a significant long-run factor explaining cereal imports by developing countries.

3 Background and Stylized Facts

Before presenting the theoretical framework, here are shown some relevant facts on population density in general and its correlation with international trade.

Population Density, in particular at the country level, is a very slow-moving variable and it takes decades to change and often by a small amount. More interesting is to explore the features of population density and its distribution around the world and the contribution of its two main components (land area and population).

Figure 1 points to two obvious facts: 1) a larger country area implies more population but 2) density (represented by bubbles size in the graph) is higher in smaller countries (i.e. island on the bottom left of the graph). However, the relationship is not linear, the two extremes of the distribution present outliers and variability around the spline function that capture the local correlation between the two variables. On the bottom left, small islands like Turks and Caicos (TCA) and Faroe's Islands (FRO) are less inhabited, and on the top left larger and more populous countries are both highly

(2022).

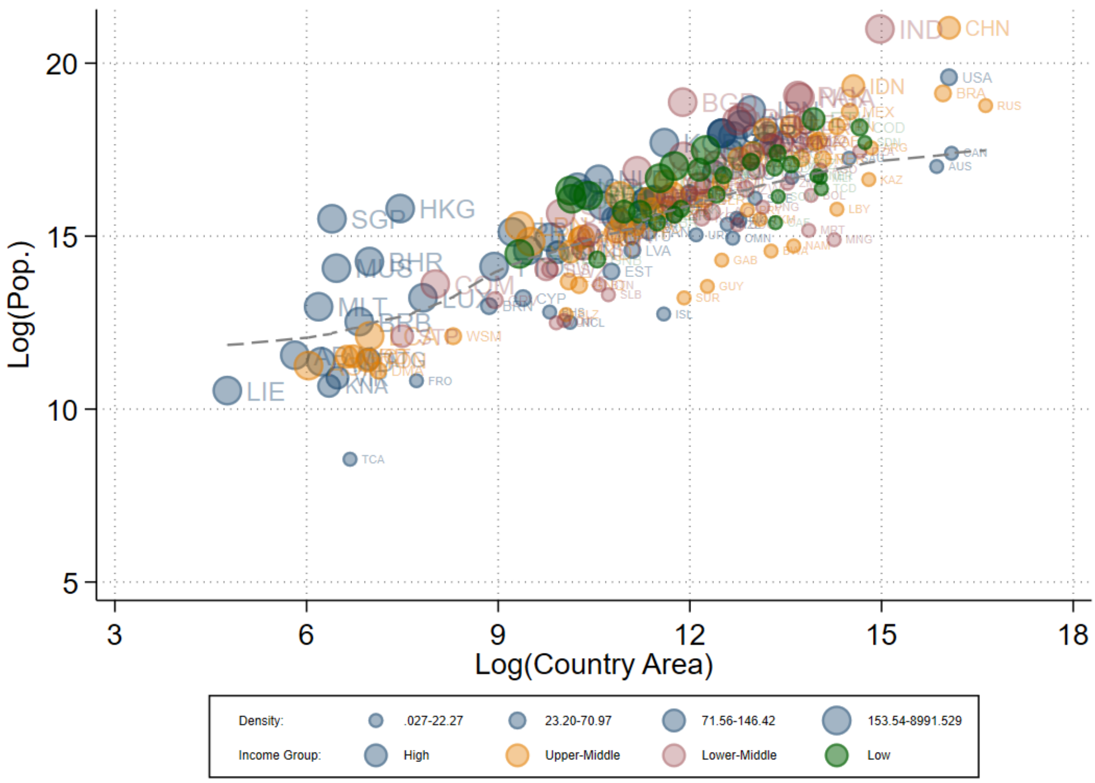
[10] In brief, Boserupian theory refers to population growth as the prime cause of agricultural change. Boserup assessed that population growth does not necessarily lead to a total depletion of food (crops in particular), but people overcome issues through technological advances able to satisfy their needs.

[11] population density on arable land.

dense (China, CHN and India, IND) but also low-dense like the US and Russia (RUS). Also, the world income distribution is heterogeneous.

Therefore, this complex view of population density across countries suggests that it could have a different impact on the export level and specialization in different industries and different counties. Country size and population concentration would have a specific effect either on sectors that are more labour or natural resources intensive, concerning also domestic specialisation. Moreover, other determinants of trade, such as bilateral trade costs and multilateral resistances, impact the volume and the margins of trade.

Figure 1: Population Density and Country Area by Income Groups



Source: Author's elaboration on population density (2015) and country area from HYDE 3.2, income groups classification is taken from the World Development Indicator of the World Bank. The non-linear fit is a lowess with running-mean smooth and tricube weighting function.

Figure 2 the correlation between total export and density (on the left) and between this and the share of exports over total production(on the right). The first column of graphs shows that population density and total exports have different nexus according to industries. Natural resources and/or land-intensive sectors (*Agriculture, Forestry and Fisheries* and *Mining*) have a negative correlation while *Manufacturing* which is (relatively) labour-intensive is positively related to density.

Different is the situation when considering relative export specialization, the graphs in the right column show a slightly different scenario. *Manufacturing* and *Mining* have the same relationship, respectively, positive and negative with population density. While *Agriculture, Forestry and Fisheries* in this case turns out to be positive, meaning that denser countries sell abroad most of their output.

Figure 2 shows some outliers in the correlation between the ratio of export on production and population density. In particular, some small (in terms of areas) countries with high-density sell abroad more than half ($\geq 50\%$) of their production.

In *Agriculture, Forestry and Fisheries*, the export shares of Luxembourg,[12] Malta, Singapore and Malta exceed 50% of the total production. These countries produce and trade mostly processed products and just not raw materials (i.e., synthetic rubbers, oilseeds for Agriculture and board and plywood for Forestry products). Having more disaggregated data on production would help to understand better these specialization patterns. Moreover, Singapore it is been developing urban vertical farming since 2005.[13]

In *Manufacturing* as well these countries are involved in global value chains, so some manufacturing productions are settled in a place which does not necessarily furnish the domestic market. The mix of advantages of having high population density and the geographical strategic position make these countries ideal to locate a specific branch of production of manufactured goods. In this case, other outliers are some

[12] see `https://resourcetrade.earth/?year=2015&exporter=442&category=1&units=value&autozoom=1`

[13] `https://ourworld.unu.edu/en/farming-in-the-sky-in-singapore`

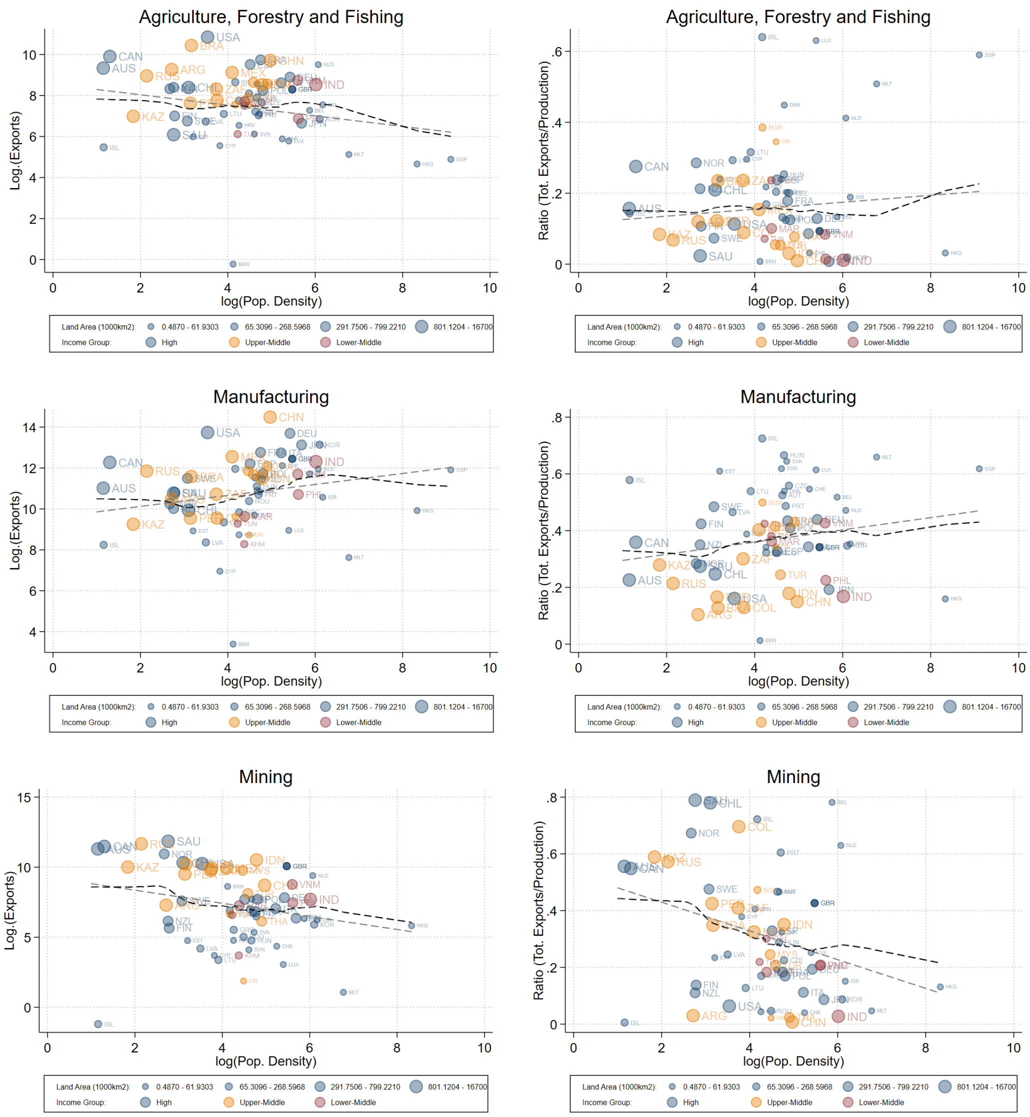
Eastern European countries (ie. Hungary, Estonia) where some automotive factories are placed.

This is important because suggests looking also at the country area and not only at its population concentration to understand the role of the spatial distribution of production factors. To sum up some interesting evidence from these pictures are:

- large countries with low population density export a high volume of goods in the natural resource of land-intensive goods;
- large countries with low population density countries are relevant exporters in all the broad sectors;
- large countries with high population density specialize in manufacturing:
- small countries with high-density export more manufactured goods;
- for small countries with high density sell abroad more than 40% of what they produce, but not for *Mining*;
- population density and country area capture also the effect of domestic demand. Agriculture, Forestry and Fisheries' output is absorbed in larger countries, while in small but dense countries is not true.

To sum up, labour-intensive activities are often associated with high population density, while industries relying on natural resources benefit from a more scattered concentration of people within a country. These findings align with the results obtained from the econometric specification. Additionally, graphs in Figure 2, according to De Benedictis et al. (2009), illustrate that countries do not necessarily specialize in just one sector but rather diversify. Large countries can benefit from both urban agglomeration and available land, allowing them to become top exporters across multiple sectors. On the other hand, even small countries with high population density, despite not being top exporters, are able to sell a significant portion of their production abroad. This is due to the positive influence of population density on productivity, which prevails over the effect of domestic demand absorption.

Figure 2: Countries' Exports, Export Shares and Density by Sectors



Source: Author's elaboration on population density (2015) and country area from HYDE 3.2, gross exports from TiVA 2018, and income groups classification are taken from the World Development Indicator of the World Bank. The non-linear fit is a lowess with running-mean smooth and tricube weighting function.

4 Theory

In this section, we present the theoretical framework and the derivation of the structural gravity model and the main equations that interpret the empirical strategy based on Heid et al. (2021) and Freeman et al. (2021).

Set Up. The economy consists, in a multi-country setting, of $N \times N$ countries, where i are the exporters $i = 1, ..., N$ and j the importers $j = 1, ..., N$. Each country produces a tradable good with infinite varieties[14] $\omega \in \Omega \equiv 1, ..., +\infty$ using just one immobile production factor, labour, where L_i is the number of workers in country i and the unit cost of labour, wages, w_i.

Preferences. Using CES assumptions (Dixit and Stiglitz, 1977), the utility of the representative consumer:

$$u(x_j(\omega)) = \sum_{j}^{N}(x_j(\omega)^{\frac{\sigma-1}{\sigma}}), \tag{1}$$

where $\sigma > 1 + \theta$, is the elasticity of substitution between varieties. The maximization problem leads to the demand (expenditure) for the varieties ω in the country j:

$$x_j = \left[\frac{p_j(\omega)}{p_j}\right]^{1-\sigma} \alpha_j w_j L_j, \tag{2}$$

where α_j is the consumption share of country j and then $\alpha_j w_j L_j$ represents the expenditure of country j.

Trade Costs. Moving goods from country i to country j is costly. According to the iceberg trade costs assumption, For each unit of good shipped from country i to country j, only $\frac{1}{\tau_{ij}} \leq 1$ units arrive, selling domestically is costless, $\tau_{ii} = 1$. For τ_{ij} hold the triangle inequality such that $\tau_{ij} \leq \tau_{il}\tau_{jl}$

[14] we use infinite varieties as in Costinot et al. (2012), as the authors pointed this does not differ substantially by a continuum of goods, this eludes the technical complication of implementing the law of large numbers with a continuum of i.i.d. variables, and the number of varieties per industry is exogenously given. Since we build on their model using their main functional forms, we keep this assumption for coherence with the framework.

Market Structure. The market is characterized by perfect competition. In any country j, the price $p_j(\omega)$ paid by buyers of a variety ω the lowest:

$$p_j(\omega) = \min_{1 \leq i \leq I}[c_{ij}(\omega)] \tag{3}$$

,

where $c_{ij}(\omega) = \frac{\tau_{ij} w_i}{A_i} > 0$ is the cost of producing and delivering one unit of this variety from country i to the country j.

Technology. Using Costinot et al. (2012) Assumption 1, for all countries i and their varieties ω, productivity $A_i(\omega)$ is a random variable, drawn independently from a Frechét distribution $F_i(.)$ such that;

$$F_i(A) = e^{-\left(\frac{A}{A_i}\right)^{-\theta}} \tag{4}$$

where $A_i > 0$ is the *fundamental productivity* (deterministic) which represents also the absolute advantages of a country- $\theta > 1$ is the *intra-industry heterogeneity* (stochastic) that parameterizes the impact of changes in fundamental productivity level A_i, and capture comaparative advantages between countries.

In the deterministic productivity component we add, as in Allen and Arkolakis (2014) and Allen et al. (2020), a specific element for labour contribution and its sensitivity to population density, η:

$$A_i = \bar{A}_i(L_i)^{\eta} \tag{5}$$

where $\bar{A}_i$ represents the exogenous country productivity, $\eta \in R$ is the extent by which agglomeration (population density) affects productivity, η is specific for each sector as in Moscona and Levy (2022), we assume that different industries do not benefit by population density in the same way.

At the moment there is not any specific assumption on the value of density sensitivity. Therefore, η, as in Allen and Arkolakis (2014), could be either positive or negative: $\eta > 0$ means that a certain industry benefits from the scale effects of population agglomeration: on the other hand, $\eta < 0$ indicates that an excessive number of workers imply diseconomies related to an excessive population level according to specific industries that rely mostly on other factors (i.e. natural resources for raw materials and intermediates) for their production process.

Expenditure Share - Trade. Given the price from 3 and the expenditure share 2 obtain, and the productivity function from equation 5:

$$X_{ij} = \frac{\left(\frac{w_i \tau_{ij}}{\bar{A}_i L_i^{\eta}}\right)^{-\theta}}{\sum_{j=1}^{N} \left(\frac{w_i \tau_{ij}}{\bar{A}_i L_i^{\eta}}\right)^{-\theta}} \alpha_j w_j L_j, \tag{6}$$

the first part of equation 6, $\pi_{ij} = \left(\frac{w_i \tau_{ij}}{\bar{A}_i L_i^{\eta}}\right)^{-\theta} / \sum_{j=1}^{N} \left(\frac{w_i \tau_{ij}}{\bar{A}_i L_i^{\eta}}\right)^{-\theta}$ is the *trade share*, representing the probability that country i supply goods at the minimum price in country j. The second term is the expenditure of country j, $E_j = \alpha_i w_j L_j$.

This equation helps to understand why we choose the functional form of productivity from Costinot et al. (2012). First, we have $(w_i \tau_{ij}/A_i)^{-\theta}$, which stress the role of A_i affecting the domestic cost (w_i) and the cost of selling goods abroad (τ_{ij}). Moreover, in this case, η and θ interact in the contribution of L_i to productivity, and as shown when estimating the effect of sensitivity to density the role of the comparative advantages parameter is crucial to not overestimate its impact.

Market Clearing. In equilibrium, the model assumes that Goods Market is cleared when:

$$Y_i = \sum_{j=1}^{N} x_{ij}, \tag{7}$$

meaning that the domestic output contains both the amount of produced goods shipped and sold to j and also the part for the domestic market. On the production side, Labour Market clears when:

$$Y_i = w_i L_i, \tag{8}$$

Price Distribution. From 3 we obtain the price distribution from a Frechét (Eaton and Kortum, 2002). The cheapest good in country j will have a price lower than p unless each price of i is greater than p. So if j buys at a lower price than p, the distribution is:

$$G_j(p) = Pr[P_j \leq p] = 1 - \prod_{j=1}^{J} [1 - G_{ij}(p)], \tag{9}$$

The equation gives the price parameter;

$$\Phi_j = \sum_{j=1}^{N} (A_i)^{\theta} (w_i \tau_{ij})^{-\theta} \tag{10}$$

Φ_j, as in Eaton and Kortum (2002),[15] concerns the world's state of the technology , and the geographic features that determine prices in each country j. The exact price index is

$$P_j = \gamma(\Phi_j)^{-\frac{1}{\theta}} \ ; \ \Phi_j = \gamma^{\theta}(P_j)^{-\theta} \tag{11}$$

[15] the model of Eaton and Kortum (2002) considers also intermediate inputs, in their framework these price parameter

The exact price index and the price distribution parameter are proportional and this helps to derive the multilateral resistance terms.

Multilateral Resistance Terms. Once price distribution, price parameter and the related exact price index are defined, it is possible to derive the *Multilateral Resistance Terms*. These are the structural terms defined by Anderson and van Wincoop (2003), that capture market importer (*inward*) and exporter (*outward*) access determinants

From Anderson and van Wincoop (2003) define the Outward Multilateral Resistance Term (OMR):[16]

$$\Pi_i = \sum_{j=1}^{N} \left(\frac{\tau_{ij}}{P_j} \right)^{-\theta} \frac{E_j}{Y}, \tag{12}$$

and the Inward Multilateral Resistance Term (IMR):

$$P_i = \sum_{i=1}^{N} \left(\frac{\tau_{ij}}{\Pi_i} \right)^{-\theta} \frac{Y_i}{Y}, \tag{13}$$

As shown in the appendix, now define the factory gate price or wage:

$$w_i = \left(\frac{Y_i/Y}{(\Pi_i)^{-\theta}(A_i)^{\theta}} \right)^{-\frac{1}{\theta}}, \tag{14}$$

this equation is different from its typical formalization because here includes also the productivity of the country i and not just the costs of exporting captured by Π_i. The Outward Multilateral resistance term also proxies unobservable congestion forces operating in each country i. Then domestic prices are lower if productivity is higher, and also higher cost of reaching a foreign market (Π_i) obliges countries to lower production costs for being competitive in the global markets. Remembering

[16] see Appendix A for derivation.

that productivity is given by 5, substitute w_i into equation 8,

$$Y_i = (\bar{A}_i)^{\frac{\theta}{1+\theta}} (L_i)^{\frac{\theta(1+\eta)}{1+\theta}} (\Pi_i)^{-\frac{\theta}{1+\theta}} (Y)^{\frac{1}{1+\theta}}, \tag{15}$$

Then adding the 15 into the main gravity equation 47, gives the extended gravity equation with exporters' specific variables:

$$X_{ij} = \frac{(\bar{A}_i)^{\frac{\theta}{1+\theta}} (L_i)^{\frac{\theta(1+\eta)}{1+\theta}} E_j (\tau_{ij})^{-\theta}}{(Y)^{-\frac{1}{1+\theta}} (\Pi_i)^{\theta-\frac{\theta}{1+\theta}} (P_j)^{-\theta}}, \tag{16}$$

5 Empirical Strategy

This section discusses how to bridge theory and econometrics. To do so we need to look at the literature focused on the estimation of unilateral variables into a structural gravity model.

Firstly looking at the published work of Heid et al. (2021), where is provided a solution to solve the perfect collinearity issues arising in including country-specific variables in a gravity model with importer and exporter fixed effect (a framework not falling in the *gold medal mistake* (Baldwin and Taglioni, 2006). Solving this issue needs to include domestic sales[17] and multiplying the unilateral variable of interest by the international borders dummy $INTL_{ij}$.[18] Even though it allows the inclusion of any country-specific or unilateral policy measure, this approach has limitations in the interpretation of the results since the coefficient besides the impact of the covariate of interest contains also its differential effect on international trade with respect to domestic sales. This article presents the empirical solutions to it but it does not include a theoretical interpretation of the results.

[17] total production minus exports

[18] $INTL_{ij} = 1$ for $i \neq j$ and $INTL_{ij} = 0$ otherwise.

The other approach, from Freeman et al. (2021), wants to build a framework and a methodology to measure the direct effect of unilateral variables on international trade through a structural gravity model.

Following this paper, the logarithm and an exponential transformation of equation 16, gives the empirical equation for a Poisson Pseudo Maximum Likelihood (PPML),

$$X_{ij} = exp[\beta_1 ln(\bar{A}_i) + \beta_2 ln(L_i) + \beta_3 ln(E_j) + \beta_4 ln(\tau_{ij}) \beta_5 ln(\Pi_i) + \beta_6 ln(P_j) + \beta_7 ln(Y)] \times \varepsilon_{ij}, \tag{17}$$

and the coefficient can be interpreted thanks to the parameter associated with each variable in the theoretical exports function,16. Therefore, $\beta_1 = \frac{\theta}{1+\theta}$, $\beta_2 = \frac{\theta(1+\eta)}{1+\theta}$, $\beta_3 = 1$, $\beta_4 = -\theta$, $\beta_5 = \theta - \frac{\theta}{1+\theta}$, $\beta_6 = -\theta$, $\beta_7 = -\frac{1}{1+\theta}$. The most important coefficient is β_2, which includes both trade elasticities $\frac{\theta}{1+\theta}$ and also the agglomeration/scale effect captured by density, $1 + \eta$.

The work of Freeman et al. (2021) proposes an alternative estimation method which overcomes the identification issues related to source and destination fixed effect but still uses a theoretically grounded gravity model. They developed a two-stage procedure in which multilateral resistance terms are proxyed by two indices measured from the origin and destination fixed effects. Here we propose an application of the two methodologies to a cross-sectional setting to compare baseline results.

5.1 Method 1: Heid et al. (2021)

Applying the approach of Heid et al. (2021) to the present theoretical framework, the reduced form for estimating the effect of population density on exports is:

$$X_{ij} = exp[\beta_2 ln(L_i) \times INTL_{ij} + \delta_0 INTL_{ij} + \beta_4 ln(\tau_{ij}) + \mu_i + \chi_j] \times \varepsilon_{ij}, \tag{18}$$

where, as already wrote above, $INTL_{ij} = 1$ for $i \neq j$ and $INTL_{ij} = 0$ for $i = j$, $ln(\tau_{ij})$ concerns bilateral trade barriers. For the moment, exogenous productivity, $\bar{A}_i$ is omitted, it is reasonable to assume that it is contained in the export fixed effects, μ_i, and it is considered just the cross-sectional setting. The last term, χ_j, is the destinations/importers fixed effect which controls for all the costs of importing, for country j expenditure and trade imbalances.

The purpose of the empirical section is the focus on β_2, the coefficient capturing the density impact. However, as pointed out previously, using the method, of multiplying the main variable for the international border dummy, is going to measure the effect of population density on international trade with respect to domestic sales.

Therefore, in the following paragraph, it is provided with a theoretical interpretation of Heid et al. (2021) method. This is useful for two reasons: 1) because the main variable is not considered as just a unilateral trade cost but also as a productivity component and it contains an additional parameter, η, that needs to be explained and interpreted properly; 2) data contains both international and domestic flows,[19] so is needed as well an interpretation of the role of these two components and how they drive the results.

Moreover, to test the robustness and the interpretation of the density coefficient, we replicate the analysis following the method of Freeman et al. (2021). This new approach allows the estimation of the direct effect of the country-specific variable. Also in this case, we provide a theoretical discussion of the interpretation when domestic and international flows are both in the sample.

[19] Using data with both the dimension is better both for merely empirical work and to run general equilibrium analysis. The advantages of using a complete dataset are widely explained in Yotov (2021).

5.2 Theoretical Interpretation of Method 1

The trade share, π_{ij}, with the price parameter a la Eaton and Kortum (2002), ϕ_j, is log-transformed:

$$log(\pi_{ij}) = \theta log(\bar{A}_i) + \eta\theta log(L_i) - \theta(log(w_i) + log(\tau_{ij}) - log(\phi_j)). \tag{19}$$

This equation allows a better theoretical interpretation of the coefficient of interest. Obtaining the partial effect of $log(L_i)$ in percentage changes and in changes respectively:[20]

$$\frac{\partial log(\pi_{ij})}{\partial log(L_i)} = \eta\theta(1 - \pi_{ij}) \tag{20}$$

$$\frac{\partial \pi_{ij}}{\partial log(L_i)} = \eta\theta(1 - \pi_{ij})\pi_{ij} \tag{21}$$

Now we focus on equation 20, the first general interpretation of the per cent change in population density is that:

- for large π_{ij} the effect on international sales is smaller, while is greater for domestic sales.
- Positive or negative changes are related to η.

However, the exact coefficient takes into account the differential effect between external and internal dimensions which is formalized as:

$$\beta_2 = \frac{\partial log(\pi_{ij})}{\partial log(L_i)} - \frac{\partial log(\pi_{ij})}{\partial log(L_i)} = \eta\theta(1 - \pi_{ij}) - \eta\theta(1 - \pi_{jj}) = \eta\theta\pi_{jj} - \eta\theta\pi_{ij} \tag{22}$$

[20] See proof in appendix A.3 for the derivation.

This means that following Heid et al. (2021), it is likely to assume that the model measures:

$$\beta_2 = \eta\theta(\pi_{jj} - \pi_{ij}) \tag{23}$$

where η and $(\pi_{jj}-\pi_{ij})$ drive the sign. From the literature[21] the parameter θ is positive. Also from literature and empirical evidence, the proportion between domestic and international trade share, $\pi_{jj} - \pi_{ij} > 0$, is that the domestic component is higher than the whole international sales.[22] More precisely, for aggregate trade, the two shares are almost balanced (close to 50%), while for sectoral trade it depends on industries and the differences can be larger.

The interpretation of η must consider that the coefficient measures only the relative effect of density on international trade. Then, η is the sensitivity to the density of international trade with respect to domestic sales. When $\pi_{jj} - \pi_{ij} > 0$ holds:

- $\eta > 0$ ($\beta_2 > 0$); the marginal contribution in the differential effect of population density is more sensitive to international sales. It contributes more to reaching foreign markets (smaller trade share than to the domestic market) than internal markets. In this case, the supply effect from productivity is clearly evident.

- $\eta < 0$ ($\beta_2 < 0$); the marginal contribution in the differential effect of population density is more sensitive to domestic sales. Even if the internal share is larger, the domestic market absorbs the density effect on trade. There are two possible explanations for this: the first is that the supply-side effect related to productivity works on economic integration (this would be the case for a sample of developing countries), and the second is that the domestic demand effect is

[21] Eaton and Kortum (2002) for aggregate trade and Costinot et al. (2012) in a multisectoral setting.
[22] The trade shares here considered are the average value of the sample, for each country this condition is not always true, as shown in the graph, as in the case of Singapore and Malta

higher than the supply one. Then, a large domestic market absorbs the greatest part of the production, as for *minings* and the extraction of fuels and other energetic sources.

To make our formalization more complete, considering the case in which $\pi_{jj} - \pi_{ij} < 0$ would provide a different interpretation of η as the sensitivity to the density of international trade with respect to domestic sales. We will discuss better this point using empirical evidence from Table 3.

Nevertheless, even if θ has been widely studied in former contributions, a correct specification of this component in the empirical part is crucial because it might affect seriously the results. This point is going to be discussed in the following paragraph.

5.3 The Role of the International Border Dummy

In the previous section, equation 18 represents our main empirical model and it contains an international border dummy, $INTL_{ij}$, both multiplied with the main variable of interest (population density) and also alone. Therefore, after the theoretical interpretation of the main coefficient $\beta_2 L_i \times INTL_{ij}$, now we focus on the role of the dummy that allows us to identify the main coefficient. The implications of the variable for international borders, $INTL_{ij}$, is widely debated in the literature in seminal work as (Anderson and van Wincoop, 2003) and (Balistreri and Hillberry, 2007) and in our baseline estimates it is used both with the country-specific variable, population density, and alone. The baseline equation is:

$$X_{ij} = exp[\beta_2 ln(L_i) \times INTL_{ij} + \delta_0 INTL_{ij} - \theta ln(\tau_{ij}) + \mu_i + \chi_j] \times \varepsilon_{ij} \tag{24}$$

The $INTL_{ij}$ dummy is crucial in the model specification, especially in cross-sectional settings. It is exogenous by construction and captures the effects of all possible determinants of trade not modelled explicitly, along with gravity covariates (geography and language), making domestic and international sales. On the other hand, this cannot catch the heterogeneous effect of international borders across countries and does not allow it to break up into its determinants. Analyze the model with country-specific variables multiplying the international border dummy and the dummy itself. We follow the formalization of an econometric model with dummy variables, defining the model when the international border is equal to one, as follows:

$$E(X_{ij}|INTL_{ij} = 1, ln(L_i), ...) = \beta_2 ln(L_i) + \delta_0 - \theta ln(\tau_{ij}) + \mu_i + \chi_j \tag{25}$$

when $INTL_{ij} = 0$, the model refers to the domestic component of the data, both trade and explanatory variables:

$$E(X_{jj}|INTLij = 0, ln(L_i), ...) = -\theta ln(\tau_{jj}) + \mu_i + \chi_j \tag{26}$$

The difference between 25 and 26 is the estimated model which takes the following form:

$$\begin{aligned}\hat{X}_{ij} &= E(X_{ij}|INTL = 1, ln(L_i), ...) - E(X_{ij}|INTL_{ij} = 0, ln(L_i), ...) = \\ &= \beta_2 ln(L_i) + \delta_0 - \theta(ln(\tau_{ij}) - ln(\tau_{jj}))\end{aligned} \tag{27}$$

Now it is clear that the effect captured by the coefficient for borders dummy, δ_0, is affecting the model specification. But to explain better its role, we need to go back to the equation 23 in the previous section and substitute β_2 and δ_0[23] with their theoretical interpretation:

[23] see Appendix A.3 for its definition.

$$\hat{X}_{ij} = \eta\theta(\pi_{jj} - \pi_{ij})ln(L_i) - \theta(\pi_{jj} - \pi_{ij}) - \theta(ln(\tau_{ij}) - ln(\tau_{jj})) \tag{28}$$

this equation shows the content of the international border dummy as in Yotov et al. (2016). It controls all possible exogenous sources of trade frictions and the wedge between domestic and international sales. Moreover, it allows controlling for potential bias arising from the difference between the two dimensions, $(\pi_{jj} - \pi_{ij})$, when the effect of density (or any other country-specific variables representing a component of fundamental productivity) is measured.

5.4 Method 2: Freeman et al. (2021)

An alternative method (hereafter called Method 2) to identify the effect of country-specific variables is provided by Freeman et al. (2021).

It consists in a two-step procedure, where the first stage is a basic gravity estimated with a PPML with panel data:

$$X_{ij,t} = exp[\mu_{i,t} + \chi_{j,t} + \tau_{ij}] \times \varepsilon_{ij,t} \tag{29}$$

where μ_i are the exporter fixed effect, χ_j the importer fixed effect and τ_{ij} the country-pair fixed effect. This estimation is useful to obtain the source and destination fixed effect to compute the related indexes for the estimated multilateral resistance terms:

$$\hat{\Pi}_i = \frac{Y_i}{exp(\hat{\mu}_i)} \times \frac{E_0}{Y} \; ; \; \hat{P}_i = \frac{E_j}{exp(\hat{\chi}_j)} \times \frac{1}{E_0} \tag{30}$$

where E_0 is the expenditure of the numeraire country. These terms are added in the second stage which is done with cross-section data to compare better the results of the two methods:

$$X_{ij} = exp[\beta_1 ln(\bar{A}_i) + \beta_2 ln(L_i) + \beta_3 ln(E_j) + \beta_4 ln(\tau_i)\beta_5 ln(\hat{\Pi}_i) + \beta_6 ln(\hat{P}_j)] \times \varepsilon_{ij} \quad (31)$$

5.5 Theoretical interpretation of Method 2

The interpretation of β_2, in this case, would be following the Freeman et al. (2021) from equation 16:

$$\beta_2 = \frac{\partial log(\pi_{ij})}{\partial log(L_i)} = (1+\eta)\frac{\theta}{1+\theta} \quad (32)$$

this statement is true if the data contains only international flows. Following the formalization proposed before but expressed in levels:

$$\beta_2 = \frac{\pi_{ij}}{\partial log(L_i)} = \eta\theta(1-\pi_{ij})\pi_{ij} \quad (33)$$

the interpretation is the same as provided in the previous section. Since θ is positive by the literature, η drives the sign of the effect. The magnitude is affected by the variance of trade shares, $(1-\pi_{ij})\pi_{ij}$, which is always positive.

Using this method, we can obtain the simple density sensitivity easier to interpret since here is ignored the differential effect of the two dimensions and then η represents the contribution of density to productivity. In the case of $\eta > 0$, we have an agglomeration effect due to the natural advantages of having a large population concentration. By contrast, when $\eta < 0$, technology and market size have a positive impact. Moreover, the natural advantages are mostly related to the natural endowments, and an increase in population density may generate diseconomies.

6 Data

This section describes the variables and the data sources used for the estimation; data on exports and production by sectors, population density and the different measures we test and the unilateral and bilateral geographic controls.

The dependent variable is bilateral exports which accounts also for domestic sales. Data are from TiVA (version 2018); international trade includes gross exports and domestic flows, the latter is the difference between gross production and total exports (Yotov, 2022). These data are grouped to obtain three broad sectors i) Agriculture, Forestry and Fisheries,[24], ii) Manufacturing iii) and Mining.[25] The sample is a $N \times N$ matrix (64 X 64 countries)[26] for each year from 2005 to 2015.

We use the TiVA sample because it is a balanced trade matrix (in terms of linkages and year) that allows us to have international and domestic flows in the same unit of measure (gross terms). It has a slightly greater country coverage than other data sources having values for each sector for trade and production. Moreover, the 64 countries in it represent a heterogeneous composition regarding geographical and economic features. We leave out services in this analysis since we find stronger evidence related to the ratio of population and land endowments.

Population density is computed from the *History Database of the Global Environment*

[24] TiVA does not contain disaggregated sector for it

[25] This sector has both energy and non-energy products. It is grouped maintaining the category *Mining support service activities.* which does not change significantly results

[26] The countries and their ISO3-CODE are: Argentina (ARG), Australia (AUS), Austria (AUT), Belgium (BEL), Bulgaria (BGR), Brazil (BRA), Brunei (BRN), Canada (CAN), Switzerland (CHE), Chile (CHL), China (CHN), Colombia (COL), Costa Rica (CRI), Cyprus (CYP), Czech Rep. (CZE), Germany (DEU), Denmark (DNK), Spain (ESP), Estonia (EST), Finland (FIN), France (FRA), Switzerland(CHE), United Kingdom (GBR), Greece(GRC), China (CHN) , Hong Kong SAR (HKG), Croatia (HRV), Hungary (HUN), Indonesia (IDN), India (IND), Ireland(IRL), Iceland (ISL), Israel (ISR), Italy (ITA), Japan(JPN), Kazakhstan (KAZ), Cambodia (KHM), Rep. of Korea (KOR), Lithuania (LTU), Luxembourg (LUX), Latvia (LVA), Morocco (MAR), Mexico (MEX), Malta (MLT), Malaysia (MYS), Netherlands (NLD), Norway (NOR), New Zealand (NZL), Peru (PER), Philippines (PHL), Poland (POL), Portugal (PRT), Romania (ROU), Russian Fed. (RUS), Saudi Arabia (SAU), Singapore (SGP), Slovakia (SVK), Slovenia (SVN), Sweden (SWE), Thailand (THA), Tunisia (TUN), Turkey (TUR), Taiwan (TWN), United States (USA), Viet Nam (VNM), South Africa (ZAF).

(HYDE 3.2) (Klein Goldewijk et al., 2017). This data set combines updated population (grid) estimates and land use for the past and for a more contemporary range of time. It classifies land into several categories by different crop and irrigation systems and other anthromes. The population is also split into total, urban and rural. The results rely on different measures of density to test for robustness:

1. **Population Density**: the standard measures of $\frac{Population}{Area(km^2)}$. The area does not count lakes.

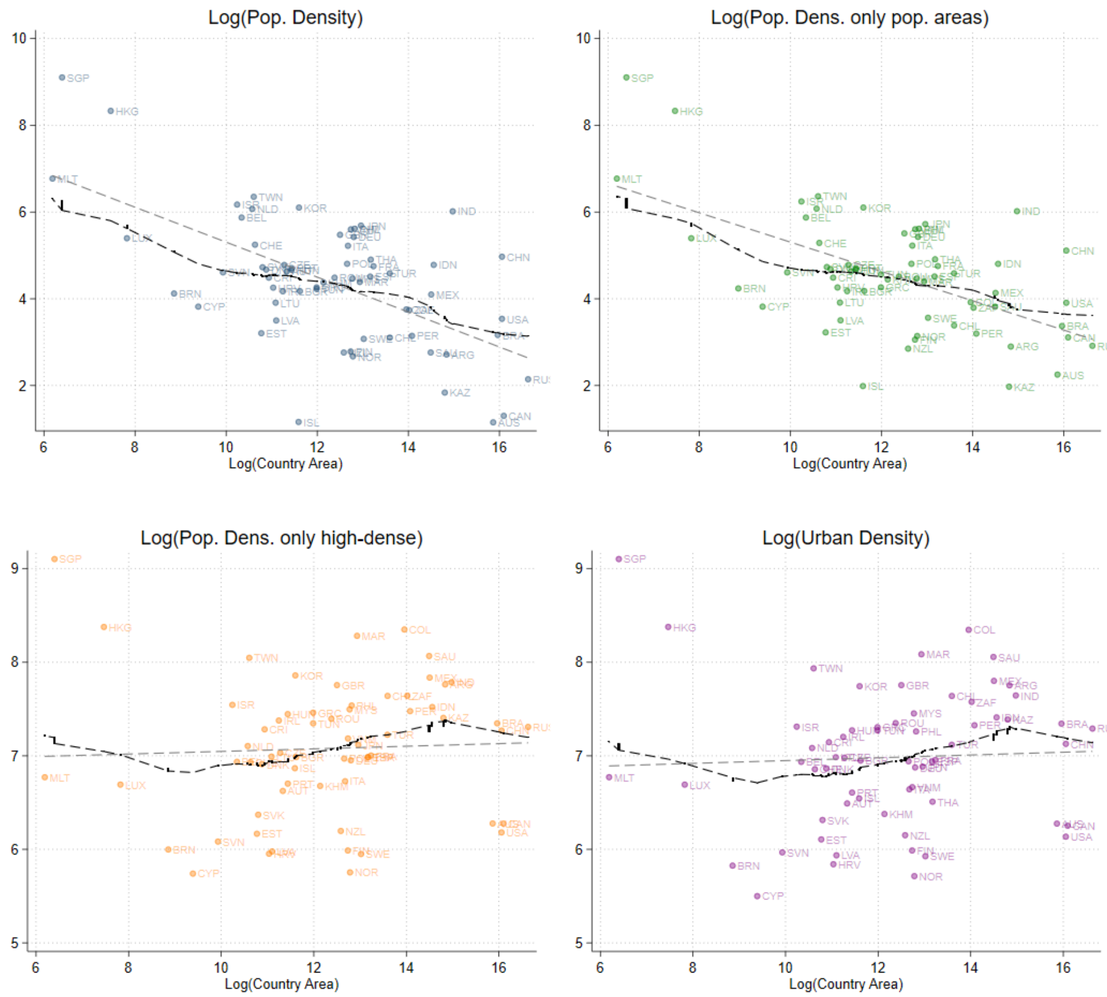
2. **Population Density (only populated cells)**: considers the area of the cells where the population is greater than zero

3. **Population Density (high density cells)**: consider total population and area only from cells classified as *Urban* and *Dense Settlements*

4. **Urban Density (urban cells)**: consider just urban population and are only from cells classified as *Urban*

The first two measures are similar (see Figure 3) larger countries (in terms of area) have less density and vice versa. These capture the uneven distribution of the population with respect to land. Differently, Population Density (high-density cells) and Urban Density (high-density cells) have the opposite relation with size and also less variability.

Figure 3: Comparing Population Density Measures



Source: Author's elaboration with HYDE 3.2 data

Note: The country area in the x-axis is the original measure (the one used for population density) and it is the same in all the graphs. It is done to compare the heterogeneity of these variables. The linear and non-linear fit regress the different density measures on the total country area. The non-linear fit is a lowess with running-mean smooth and tricube weighting function.

The last set of unilateral variables represents further controls for geographical features that might affect production, productivity, and density. These are taken from the seminal work of Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (Gem diamond extraction 1958-2000 (1000 carats)), *near coast* (percentage Within 100 km. of ice-free coast). To make a robustness check on potential omitted variables we use nominal GDP per capita from CEPII, and total employment and human capital from Penn's World Table.

The bilateral covariates are the weighted distance, contiguity, official common language and colonial links taken from the *GeoDist Database* (Mayer and Zignago, 2011). An alternative measure of distance concern sailing length. The dyadic components (distance from country A to country B are provided by the CERDI Sea Distances dataset (Bertoli et al., 2016). This measure does not include internal distances which are computed by the author using the Router Project Open Street Map. To do so, according to the assumption of domestic trade costs from Ramondo et al. (2016), these are the average distance related to country size. Here including country size in the internal distance is considered as a starting point for the centroid of each country. Then the road distances using Open Street Maps tools measure the kilometres to reach the main port according to the CERDI data. For landlocked countries, internal distances are imputed regressing weighted distances on the road distance. The imputation is done to avoid strong assumptions on the geographic domestic frictions of those countries.

7 Results

First, we present the baseline results, a cross-section for 2015, based on Heid et al. (2021). The robustness checks are made using the sea distance measure integrated with domestic road distance from the country's centroid to the main port. Then, we show the effect of using different density measures and including other variables to check if there is a problem with omitted variables. With a focus on manufacturing, we run the same analysis for each industry contained in the aggregate trade, to check potential biases related to aggregation as suggested in Redding and Weinstein (2019) and Breinlich et al. (2022). Always using this methodology (Heid et al., 2021) we estimate also the effect of density on international trade with respect to domestic sales in a panel setting (2005-2015) and check bilateral determinants firstly with gravity covariates and also with country pairs fixed effects. Thereafter, we explore the implication of a cross-section version of the approach in Freeman et al. (2021), which measure the direct effect of density on overall trade.

7.1 Method 1: Heid et al. (2021).

7.2 Results: Cross-Section

Baseline estimates refer to 2015. The time dimension is not examined in the theoretical part and it is shown later, the panel analysis yields slightly different and less robust results. The main cross-sectional equation is:

$$X_{ij} = exp[\beta_2 ln(L_i) \times INTL_{ij} + \delta_0 INTL_{ij} + \\ + \boldsymbol{\beta} GEO\_CONTROLS_i \times INTL_{ij} - \theta ln(\tau_{ij}) + \mu_i + \chi j] \times \varepsilon_{ij} \tag{34}$$

Using a PPML to estimate the effect of density, $ln(L_i) \times INTL_{ij}$, on exports with respect to domestic sales, and including gravity covariates $ln(\tau_{ij})$ including different measures of distance, controlling for multilateral resistance terms with the exporter, μ_i, and importer, χ_i, fixed effects, and controlling for international and internal trade with the border dummy and exporter-specific geographic features, $GEO\_CONTROLS_i$, from Nunn and Puga (2012).

Table 1: Baseline Estimates, PPML, Cross-Section: Gross Exports, 2015

VARIABLES	(1) Manufacturing	(2) Agric. Forest. and Fish.	(3) Mining
$Log(Density) \times INTL_{ij}$	0.3200***	-0.0156	-0.2703**
	(0.0525)	(0.0767)	(0.1295)
Observations	4,096	4,050	3,670
Exporter FE	YES	YES	YES
Importer FE	YES	YES	YES
GRAVITY	YES	YES	YES
$INTL_{ij}$	YES	YES	YES
GEO Control X $INTL_{ij}$	YES	YES	YES
Clusters	Pair	Pair	Pair

Note: The difference in sample size in different sectors is due to singletons and duplicates which in Agriculture et al. and Mining are dropped by the importer and exporter fixed effect. $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (Gem diamond extraction 1958-2000 (1000 carats)), *near coast* (percentage Within 100 km. of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$. Clustered by pair (exporter-importer, non-symmetric) robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

The first set of results, Table 1, measures population density elasticities for the three broad sectors. Manufacturing gross exports increase by 0.32% with respect to domestic sales consequently to a 1% increase in population density. Drawing from the theoretical interpretation presented in Section 5.2, as well as the trade shares data in Table 11, the density sensitivity, denoted as η, positively influences export performance with respect to domestic trade.

The outcomes for agricultural sectors lack significance due to inherent industry heterogeneity that we cannot explore given data classification. For instance, forestry relies mainly on natural resources, while other products like horticulture goods come from more technology intensive activities, and in many cases, grown in greenhouses, requiring less land and natural resources. However, the diversity among products makes it challenging to properly identify the impact of density on agricultural trade using the aggregation provided by TiVA, as indicated by the parameter η in the model.

Mining exhibits a negative and statistically significant coefficient in accordance with the theoretical interpretation of the β_2. Thus, the mining export elasticities with respect to domestic sales lead to a variation of -0.27% of trade if population density rise by 1%. Consequently, based on our theoretical prediction, a negative η suggests that an increase in the labour force reduces exports, possibly due to diseconomies resulting from population concentration or the sector's higher dependence on natural resources.

The role of the parameter θ also emerges from the theoretical prediction concerning the coefficient β_2. In the literature, it has positive value and, as in Eaton and Kortum (2002), is estimated using geographic barriers. To test the sensitivity of these results to this parameter, we used a different measure of shipping distances regarding sea travel. In Table 2, coefficients do not change by significance and sign. In manufacturing, the elasticity slightly increases from 0.32% to 0.33%.

Table 2: Alternative distances measure: PPML, Cross-Section: Gross Exports, 2015

	(1)	(2)	(3)
VARIABLES	Manufacturing	Agric. Forest. and Fish.	Mining
$Log(Density) \times INTL_{ij}$	0.3377***	0.1311*	-0.1554
	(0.0587)	(0.0790)	(0.1416)
Log(Sea Distances) includes domestic	-0.4460***	-0.5901***	-0.8049***
	(0.0233)	(0.0443)	(0.0778)
Observations	4,096	4,050	3,670
Exporter FE	YES	YES	YES
Importer FE	YES	YES	YES
GRAVITY	YES	YES	YES
$INTL_{ij}$	YES	YES	YES
GEO Control X $INTL_{ij}$	YES	YES	YES
Clusters	Pair	Pair	Pair

Note: The difference in sample size in different sectors is due to singletons and duplicates which in Agriculture et al. and Mining are dropped by the importer and exporter fixed effect. $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (Gem diamond extraction 1958-2000 (1000 carats)), *near coast* (percentage Within 100 km. of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$. *Sea Distance* is from Bertoli et al. (2016) plus author value on domestic road distance(from centroid to main port): landlocked distances are imputed.Clustered by pair (exporter-importer, non-symmetric) robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

7.3 Robustness check: Aggregation

The works of Redding and Weinstein (2019) and Breinlich et al. (2022) discussed the potential biases arising from different aggregation (or disaggregation levels) according to various sectors and product classifications.

The sample we use allows us to check if the results for the aggregate manufacturing obtained in Table 1 show potential bias due to considering manufacturing as a broad sector. Thus in Table 8, we run the same analysis for each industry. Most of them have a positive and significant sign and suggest that the η is positive in all the cases. The different coefficient magnitudes imply different levels of sensitivity to agglomeration forces of exports.

Table 3: Coefficients, trade shares and density sensitivity by Manufacturing Industries (TiVA, 2015)

Industry	$\beta \times INTL_{ij}$	π_{jj}	π_{ij}		
Chemicals	0.2461 ***	0.5090	0.4910	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Petroleum	0.5303***	0.6496	0.3504	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Fabricated metal	0.2928***	0.6569	0.3431	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Food	0.2010***	0.7950	0.2050	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Other manufactur.	0.4954 ***	0.6650	0.3350	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Other non-metall.	0.4513 ***	0.7631	0.2369	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Paper products	0.2317 ***	0.7283	0.2717	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Rubber and plast.	0.2015**	0.54567	0.4543	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Wood and product	0.3608***	0.7338	0.2662	$\pi_{jj} - \pi_{ij} > 0$	$\eta > 0$
Other transport	0.3804***	0.3655	0.6345	$\pi_{jj} - \pi_{ij} < 0$	$\eta < 0$
Textiles and wearing	0.3670***	0.4766	0.5234	$\pi_{jj} - \pi_{ij} < 0$	$\eta < 0$
Machinery and eq.	0.3384***	0.3898	0.6102	$\pi_{jj} - \pi_{ij} < 0$	$\eta < 0$
Basic metals	0.0388	0.4853	0.5147	$\pi_{jj} - \pi_{ij} < 0$	$\eta < 0$
Computer and electr.	0.2259	0.3569	0.6431	$\pi_{jj} - \pi_{ij} < 0$	$\eta < 0$
Electrical equip.	0.2203	0.4091	0.5909	$\pi_{jj} - \pi_{ij} < 0$	$\eta < 0$
Motor vehicles	-0.0535	0.4085	0.5915	$\pi_{jj} - \pi_{ij} < 0$	$\eta > 0$
Total average	0.2830	0.5586	0.4414		

Note: Full results are in Table 8. Manufacturing sectors are the one provided by TiVA (version 2018) $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (Gem diamond extraction 1958-2000 (1000 carats)), *near coast* (percentage Within 100 km. of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$. Clustered by pair (exporter-importer, non-symmetric) robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

Our theoretical discussion helps to integrate the work of Breinlich et al. (2022). This article discusses how differences in parameters and trade costs may produce different results from aggregate and disaggregate trade. We add the implication of domestic and international trade shares thanks to our theoretical framework. In Table 8, we compare the coefficients of each industry with their average trade shares (international and domestic) of the sample. At first, we notice that even if trade shares are slightly different from the baseline estimates' sample. However, even while holding the condition $\pi_{jj} - \pi_{ij} > 0$, it's notable that the average coefficient across all industries is approximately 0.28% (total average), which isn't significantly different from the main results.

When looking at the coefficients for individual industries, most exhibit positive differentials between domestic and international trade shares. However, a subset does not satisfy this condition—namely, *Basic metals*, *Computer, elctr.*, *Electrical equip.*, and *Motor vehicles*—rendering their results statistically insignificant. Conversely, industries such as *Textiles and wearing*, *Other transport*, and *Machinery and equipment* validate the condition $\pi_{jj} - \pi_{ij} < 0$. For these, international sales surpass domestic ones. This aligns with our predictions; the positive coefficients imply $\eta < 0$, suggesting that sensitivity to the density of international sales concerning domestic flows influences a country's domestic market. A plausible explanation is that agglomeration forces, via density sensitivity, impact domestic flows, consequently increasing total production and, in turn, fostering exports. Hence, a negative η, which measures the relative effect of density, fosters economic integration. An alternative interpretation could be that these economic activities are influenced by their positions within global value chains. They might be strategically located in specific regions for particular reasons, which could lead to distinct roles for domestic markets. In such cases, incorporating input-output linkages could provide a clearer understanding of how population density contributes to these dynamics.

To better quantify θ, we would require the θ value for each manufacturing activity. Unfortunately, only a few works offer such detailed information. Caliendo and Parro

(2015) has the measures of the dispersion of productivity estimates, which has a value for each industry, but unfortunately, these do not precisely match our classification.

To conclude, within manufacturing, there are differences in the parameters and the trade share. All these generate different coefficients, but the overall effect is not biased. However, looking just at aggregate trade may hide the heterogeneity and specificity of each economic activity.

7.4 Robustness check: Different measures of density

Some robustness checks are done in Table 4 using the alternative density measure presented in Section 6. The coefficient of density on just populated cells, the effect on manufacturing is similar to the baseline, 0.38 instead of 0.32 from the baseline. While the other two measures do not generate any statistically significant results for the manufacturing sectors. For the other industries, the first attempt does not yield relevant outcomes. Although, the measures of the density of *highly dense* and *urban* areas produce negative and significant results for both the indicators and both the sectors (*Agriculture, Forestry and Fisheries* and *Mining*. Therefore, these robustness checks suggest that the effect captured by density at the aggregate level is related to the specialization and the performance due to countries' spatial distribution of production factors, similar to the concept of *lumpiness* of Courant and Deardorff (1992) and Courant and Deardorff (1993). As Figure 3 shows, considering urban density, the heterogeneity between countries' density almost disappears, and country areas do not matter. Hence, we can state that agglomeration forces related to urbanization are not just a matter of the number of inhabitants.

Table 4: Robustness check with alternative measures of density, Cross-Section(2015)

VARIABLES	(1) Manufacturing	(2) Agric. Forest. and Fish.	(3) Mining
Log(Pop. Density (only populated cells)) × $INTL_{ij}$	0.3806***	-0.0388	-0.0802
	(0.0676)	(0.0979)	(0.1546)
Observations	4,096	4,050	3,670
VARIABLES	(4) Manufacturing	(5) Agric. Forest. and Fish.	(6) Mining
Log(Pop. Dens. Only high dense areas) × $INTL_{ij}$	0.0331	-0.4850***	-0.6420***
	(0.1129)	(0.1437)	(0.2079)
Observations	4,096	4,050	3,670
VARIABLES	(7) Manufacturing	(8) Agric. Forest. and Fish.	(9) Mining
Log(Urban Density) × $INTL_{ij}$	-0.0048	-0.4523***	-0.5763***
	(0.1125)	(0.1500)	(0.2134)
Observations	4,096	4,050	3,670
Exporter FE	YES	YES	YES
Importer FE	YES	YES	YES
GRAVITY	YES	YES	YES
$INTL_{ij}$	YES	YES	YES
GEO Control X $INTL_{ij}$	YES	YES	YES
Clusters	Pair	Pair	Pair

Note: The difference in sample size in different sectors is due to singletons and duplicates which in Agriculture et al. and Mining are dropped by the importer and exporter fixed effect. $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (Gem diamond extraction 1958-2000 (1000 carats)), *near coast* (percentage Within 100 km. of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$. Clustered by pair (exporter-importer, non-symmetric) robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

7.5 Robustness check: omitted variable(s)

In the previous analysis, we used geographic determinants. Table 5, focusing on the manufacturing sector, we include variables related to the development level, in particular, GDP per capita (current US$), overall employment (not by sectors) and human capital (the last two from Penn's World Table).

We include these variables in the regression as they are because they are easier to interpret, and we are not interested in quantifying their effect. This table wants to verify if our baseline results are sensitive if adding other country-specific variables.

The main result is in column (4), where including all these variables, which are also all statistically significant and determine a relative positive effect on international trade, the elasticity of population density on gross exports is similar in magnitude and sign to our baseline results.

Table 5: Robustness Check: Other Development Features

VARIABLES	(1) Manufacturing	(2) Manufacturing	(3) Manufacturing	(4) Manufacturing
$Log(Density) \times INTL_{ij}$	0.4507***	0.3278***	0.4315***	0.3317***
	(0.0476)	(0.0518)	(0.0438)	(0.0407)
GDP per capita × $INTL_{ij}$	0.0252***			0.0167***
	(0.0037)			(0.0039)
Employment × $INTL_{ij}$		-0.0001		0.0014***
		(0.0003)		(0.0003)
Human Capital × $INTL_{ij}$			1.0992***	0.9657***
			(0.1098)	(0.1519)
Observations	4,096	4,096	4,096	4,096
Exporter FE	YES	YES	YES	YES
Importer FE	YES	YES	YES	YES
GRAVITY	YES	YES	YES	YES
$INTL_{ij}$	YES	YES	YES	YES
GEO Control × $INTL_{ij}$	YES	YES	YES	YES
Clusters	Pair	Pair	Pair	Pair

Note: The difference in sample size in different sectors is due to singletons and duplicates which in Agriculture et al. and Mining are dropped by the importer and exporter fixed effect. $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (Gem diamond extraction 1958-2000 (1000 carats)), *near coast* (percentage Within 100 km. of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$. Clustered by pair (exporter-importer, non-symmetric) robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

7.6 Results: Panel

Even if the model does not consider a dynamic setting, from the literature on structural gravity, it is possible to extend the static setting to a panel one without extending the theoretical framework. Then the equations for the next set of estimates are, for table 6:

$$X_{ij,t} = exp[\beta_2 ln(L_{i,t}) \times INTL_{ij} + \\ + \boldsymbol{\beta} GEO\_CONTROLS_i \times INTL_{ij} - \theta ln(\tau_{ij}) + \mu_{i,t} + \chi j, t] \times \varepsilon_{ij,t} \tag{35}$$

and for Table 10 (in Appendix B):

$$X_{ij,t} = exp[\beta_2 ln(L_{i,t}) \times INTL_{ij} + \gamma_{ij} + \mu_{i,t} + \chi j, t] \times \varepsilon_{ij,t} \tag{36}$$

The outcomes in Table 6 align with cross-section estimates: the density elasticity is 0.32% in manufacturing, insignificant for agriculture, and negative and significant for mining (value of -0.27%).

To check the robustness, instead of including the gravity covariates, we added pair fixed effects, γ_{ij}, in equation 36. These absorb all the bilateral and unilateral not time-varying variables (such as $GEO\_CONTROLS_i$) and control for all possible bilateral trade frictions between countries and each country's wedge of domestic and foreign sales. In other words, the second equation aims to check if other models' specifications are affected by latent bilateral variables or other pair-specific issues.

These fixed effects lead to varying results (as seen in Table 10 in Appendix B). While manufacturing remains consistent, the coefficient increases by nearly three times, agriculture is positive and significant and mining loses statistical significance. The difference in the results is probably due to the absence of geographical bilateral determinants because these are crucial to identify the impact of population density.

Not including specifically the bilateral and unilateral determinants makes hard to determine which are the interplays of geographical and other kinds of features that determine bilateral trade flows and the nexus with population density. Moreover, density is country-specific and a very slow-moving trend, controlling for bilateral dimensions may not be the best approach to compare the magnitude of the coefficients.

Table 6: PPML: Gross Exports, 2005-2015

VARIABLES	(1) Manufacturing	(2) Agric. Forest. and Fish.	(3) Mining
$Log(Density) \times INTL_{ij}$	0.3246***	-0.0538	-0.2737***
	(0.0877)	(0.1168)	(0.1021)
Observations	44,671	42,284	36,740
Exporter X Time FE	YES	YES	YES
Importer X Time FE	YES	YES	YES
Pair FEs	NO	NO	NO
GRAVITY	YES	YES	YES
$INTL_{ij}$	YES	YES	YES
GEO Control × $INTL_{ij}$	YES	YES	YES
Clusters	Exporter × Importer	Exporter × Importer	Exporter × Importer

Note: The difference in sample size in different sectors is due to singleton and duplicates which in Agriculture et al. and Mining are dropped by the importer-time and exporter-time fixed effect. $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (Gem diamond extraction 1958-2000 (1000 carats)), *near coast* (percentage Within 100 km. of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$ and are time-invariant. 2-way clustered robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

7.7 Method 2: Freeman et al. (2021)

As discussed in Sub-section 5.4, Freeman et al. (2021) proposes a new methodology, theoretically grounded, that allows to estimate unilateral variables with multilateral resistance terms but not necessarily using importer and importer fixed effects in the same regression. Then it is possible to identify the direct coefficient of country-specific variables and avoiding any perfect collinearity issues.

To make this consistent, they propose a two-stage procedure: the first stage regression (see equation 29) provides the fixed effects used to compute the indices of the outward and inward multilateral resistance terms. We run a first stage on the panel setting (2005-2015) to check potential differences across years for our cross-section specification described below.

Therefore, using the novel approach of Freeman et al. (2021), which drives the estimation of the following equation:

$$X_{ij} = exp[\beta_2 ln(L_i) + \delta_0 INTL_{ij} + \beta GEO\_CONTROLS_i + ln(\hat{\Pi}_i) + \chi_j] \times \varepsilon_{ij}. \quad (37)$$

The difference with the previous tables is that $INTL_{ij}$ is a stand-alone control and does not interact with unilateral explanatory variables. Only importer fixed effects, χ_j are included, while for the exporter's side, the Outward Multilateral Resistance Terms is the index estimated as in the section above. The main point of this method is to estimate the direct effect on levels of the variable of interest.

Table 7: Alternative Cross Section Estimates 2015: method of Freeman et al. (2021)

VARIABLES	(1) Manufacturing	(2) Agric. Forest. and Fish.	(3) Mining
$Log(Pop.Density)$	0.5619***	-0.0996	-0.5413***
	(0.0633)	(0.0628)	(0.1078)
$Log(\hat{\Pi}_i)$	-1.3402***	0.0932	0.2396**
	(0.1200)	(0.1640)	(0.1187)
Observations	4,061	3,844	3,340
Exporter FE	NO	NO	NO
Importer FE	YES	YES	YES
GRAVITY	YES	YES	YES
$INTL_{ij}$	YES	YES	YES
GEO Control	YES	YES	YES
Clusters	Pair	Pair	Pair

Note: The difference in sample size in different sectors is due to singleton and duplicates which in Agriculture et al and Mining are dropped by the importer and exporter fixed effects. The small reduction of the observations in all samples is due to the first-stage estimates. $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (gem diamond extraction 1958-2000, 1000 carats), *near coast* (percentage within 100 km of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$. $OMR(i)$ is computed as in Freeman et al. (2021). Clustered by pair (exporter-importer, non-symmetric) robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1.

7.8 Comparison

The main difference between the two methods is that conceptually with method 1 (Heid et al., 2021) we measure the effect of the density of international trade with respect to internal trade and as well as the density sensitivity. Method 2 (Freeman et al., 2021) furnishes a straightforward method to assess density impact. Although the first approach provides more robust results.

The results in Table 7 confirm the findings in Table 1: the positive and significant coefficient for manufacturing, no statistically relevant effect on agricultural and related goods and negative and significant results for minings. The difference is that the coefficients are greater than the baseline. The interpretation is that both the dimensions (international and domestic) are influenced in the same way by density. In line with the theoretical interpretation, the parameter η determines the sign and then the type of impact density has on different sectors.

A further explanation is needed for the OMR index, in column 1 of Table 7, the sign is negative as expected since it represents a cost term. While in column 3 this is not verified, a plausible explanation is that when including domestic flows in the estimation, the effect of internal frictions operating in the domestic market is higher than international. Hence, trade barriers in domestic markets are lower than in foreign ones. The domestic demand absorbs the largest part of the output the selling costs are lower overall.

Table 9 provides a further check of the two approaches. In this case, we combine the two methods, to control for exporter multilateral resistance, we include the OMR index instead of fixed effects, and the dependent variable is the log of population density multiplied by the international border dummy. The main difference is that the agricultural sector here is negative, as expected, and significant. This reinforces the idea that better disaggregation is needed to have robust and coherent results for it. Mining is in line with all the other results, also here negative and significant, and the magnitude does not differ much from the other outcome. Moreover, manufacturing has the same behaviour as the previous estimates, but the coefficient of 0.48 is slightly lower than the direct impact estimates in Table 7 and larger than the coefficient of 0.32 of the baseline results. Finally, all these attempts confirm the mechanism described in the theory but the application of the new method of Freeman et al. (2021) could be improved.

7.9 Assessing the value of the density sensitivity parameter, η

The theoretical discussion above helps the interpretation of the coefficients, and it can also isolate the effect of η and quantify it. We propose two ways to do it, one following Heid et al. (2021),

$$\eta = \hat{\beta}_2 \Big(\frac{1}{\theta(\pi_{jj} - \pi_{ij})} \Big), \tag{38}$$

and the other using the methods in Freeman et al. (2021),

$$\eta = \hat{\beta}_2 \Big(\frac{1}{\theta \pi_{jj} \pi_{ij}} \Big). \tag{39}$$

The value of η is the share by which density sensitivity contributes to the overall population density elasticities on trade. As already stated, θ is measured by previous contributions, as Eaton and Kortum (2002) and Costinot et al. (2012). We focus on the manufacturing sector since it shows robust results. From Figure 4 and Figure 5, [27] there are no relevant differences. In general, a higher value of the technology parameter,[28] θ reduces the magnitude of density sensitivity. To test the sensitivity of the parameter and enhance the role of θ, when it is equal to one, high heterogeneity

[27] Table 12 and Table 13 sum up the detailed results.
[28] this means that $\lim_{\theta \to \infty} \eta(\theta) = 0$.

over varieties, comparative advantages affect mostly trade more than geographic barriers. In this case, the effect is all on *eta* it may overestimate its contribution. The relevant values are when $6 \leq \theta \leq 8$, then $0.26 \geq \eta \geq 0.20$.

According to these values of θ, the increase of 0.32% (see coefficient from Table 1 obtained using Heid et al. (2021) method) in manufacturing gross exports with respect to domestic sales the specific contribution to the density is between the 20% and the 26% of the overall effect (which is also determined by the difference between average trade shares of the sample and the technology parameter). And also, the larger the technology impact lower is the effect of the density. Slightly larger is the pure density effect of on trade measured using Freeman et al. (2021), between 0.29 and 0.39 for θ respectively of value 8 and 6. The only difference is the interpretation since in this case, we have the direct effect instead of the relative one.

Using both methods, the values obtained are acceptable since the effect of labour on output is $1 + \eta$, then scaling of the production factor on overall country output is admissible. Furthermore, these results represent the country-level version of the findings of Combes et al. (2012) on the contribution of large cities to productivity. Our work differs from the unit of analysis, both at the administrative level and by sector classification.

Figure 4: η (from Heid et al. (2021) based estimates) values according to θ measures in literature

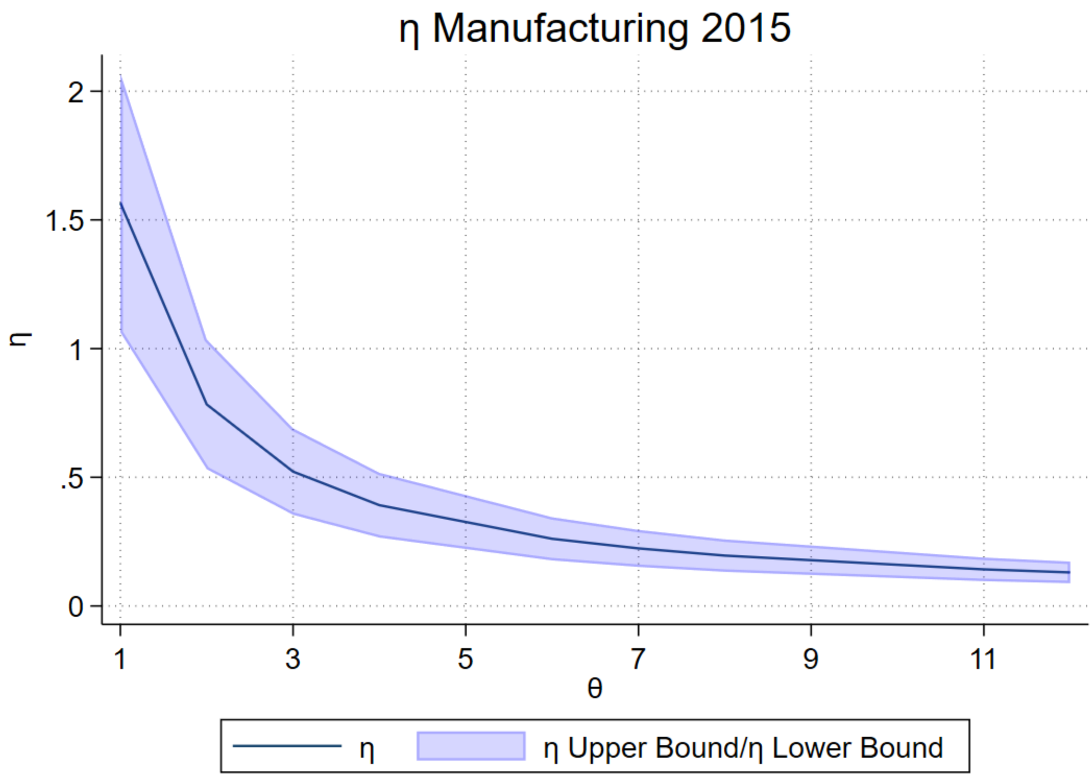
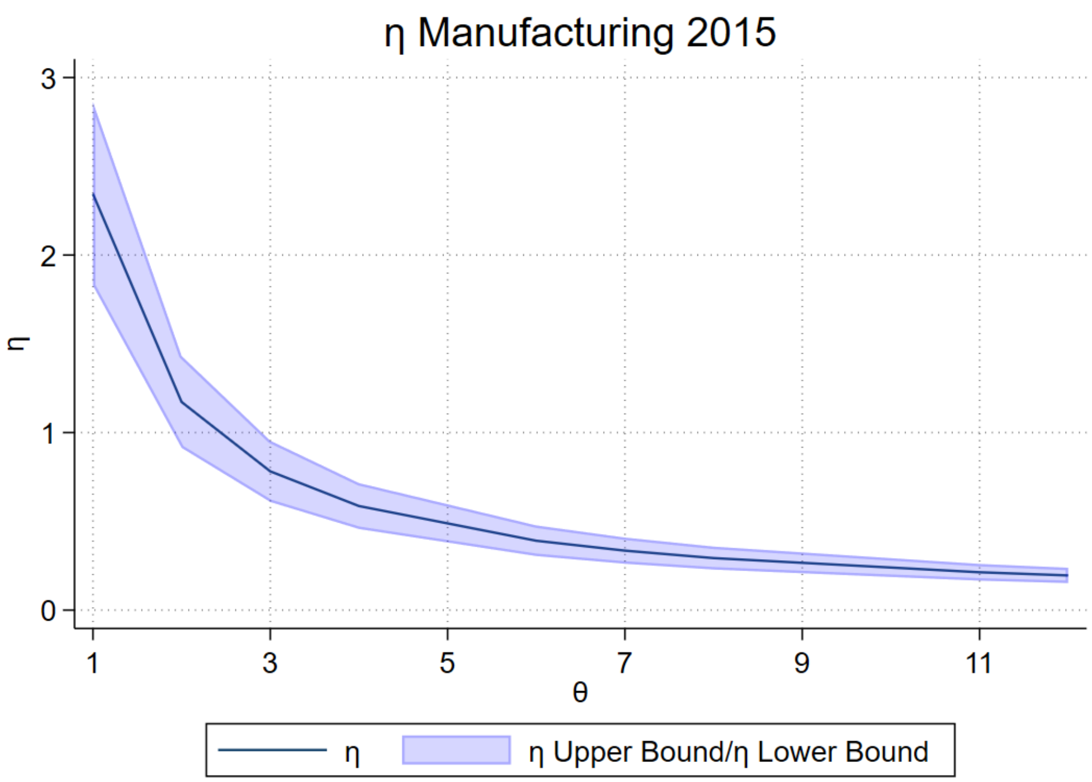


Figure 5: η (from Freeman et al. (2021) based estimates) values according to θ measures in literature



8 Conclusion

This work assesses the impact of density, a fundamental productivity component, on exports. From the theoretical point of view, it includes labour contribution to productivity and allows quantifying the possible scale effects or not. The approach wants to merge theory from Allen and Arkolakis (2014) and Allen et al. (2020) to measure structural gravity framework with country-specific geography.

Moreover, we provide a theoretical interpretation of the approach proposed by Heid et al. (2021) which allows extending this approach not only to trade frictions but also to other variables that are affecting both domestic and international dimensions. This is important because following this method we can also design counterfactual and policy experiments. The flexibility of this framework interprets the analysis of Freeman et al. (2021), which is important to quantify the direct effect of unilateral variables and policy in a theoretically grounded structural gravity model. We give evidence that the manufacturing sector (and its industries) benefits from population concentration in the country area. While other sectors, which are less labour-intensive and more natural resources dependent, uninhabited land area is more important. Agriculture, Forestry and Fisheries may need a deeper and more specific analysis: thinking better about land uses, technology heterogeneity within their industries and the differences between markets.

Looking at different measures of density, the traditional way to measure it produces a variable that captures resources endowments and distribution, while considering only the urbanization is slightly different. Urbanization in numbers may not vary or not captures proper heterogeneity across countries. What differs in cities is the size, quality, and how agglomeration and congestion forces work.

Congestion forces in Allen and Arkolakis (2014) and Allen et al. (2020) are explicitly modelled in the demand. These are more relevant in general equilibrium, while our work provides a partial equilibrium analysis. Moreover, multilateral resistance terms control these country-specific forces that we do not model explicitly.

This work gives several opportunities for further research as 1) developing a model to run counterfactual analysis including density sensitivity parameters (as adapting Dekle et al. (2008)) it is a starting point to measure how population dynamics as transitional growth [29] affect growth and trade and includes path dependency and persistence (Allen and Donaldson, 2020), 3) applies to sub-national analysis, and it may address policy evaluation related to economic geography implications and also the linkage between regional and country-level units as in Ramondo et al. (2016).

[29] This can be done modifying Anderson et al. (2020).

References

Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3):1085–1140.

Allen, T., Arkolakis, C., and Takahashi, Y. (2020). Universal Gravity. *Journal of Political Economy*, 128(2):393–433.

Allen, T. and Donaldson, D. (2020). Persistence and Path Dependence in the Spatial Economy. NBER Working Papers 28059, National Bureau of Economic Research, Inc.

Anderson, J. E., Larch, M., and Yotov, Y. V. (2020). Transitional Growth and Trade with Frictions: A Structural Estimation Framework. *The Economic Journal*, 130(630):1583–1607.

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

Bakker, J. D., Garcia-Marin, A., Potlogea, A., Voigtländer, N., and Yang, Y. (2021). Cities, heterogeneous firms, and trade.

Baldwin, R. and Taglioni, D. (2006). Gravity for dummies and dummies for gravity equations. Working Paper 12516, National Bureau of Economic Research.

Balistreri, E. J. and Hillberry, R. H. (2007). Structural estimation and the border puzzle. *Journal of International Economics*, 72(2):451–463.

Bertoli, S., Goujon, M., and Santoni, O. (2016). The CERDI-seadistance database. Working Papers halshs-01288748, HAL.

Bleaney, M. and Neaves, A. (2013). Declining distance effects in international trade: Some country-level evidence. *World Economy*, 36(8):1029–1040. cited By 9.

Breinlich, H., Novy, D., and Santos Silva, J. (2022). Trade, gravity and aggregation. *Review of Economics and Statistics*, pages 1–29.

Caliendo, L. and Parro, F. (2015). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies*, 82(1 (290)):1–44.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6):2543–2594.

Conte, M., Cotterlaz, P., and Mayer, T. (2022). The CEPII Gravity Database. Working Papers 2022-05, CEPII research center.

Costinot, A., Donaldson, D., and Komunjer, I. (2012). What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas. *Review of Economic Studies*, 79(2):581–608.

Courant, P. N. and Deardorff, A. V. (1992). International trade with lumpy countries. *Journal of Political Economy*, 100(1):198–210.

Courant, P. N. and Deardorff, A. V. (1993). Amenities, nontraded goods, and the trade of lumpy countries. *Journal of Urban Economics*, 34(2):299–317.

De Benedictis, L., Gallegati, M., and Tamberi, M. (2009). Overall trade specialization and economic development: countries diversify. *Review of World Economics*, 145(1):37–55.

Dekle, R., Eaton, J., and Kortum, S. (2008). Global Rebalancing with Gravity: Measuring the Burden of Adjustment. *IMF Staff Papers*, 55(3):511–540.

Dixit, A. K. and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American Economic Review*, 67(3):297–308.

Duranton, G. and Puga, D. (2004). Chapter 48 - micro-foundations of urban agglomeration economies. In Henderson, J. V. and Thisse, J.-F., editors, *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2063–2117. Elsevier.

Duranton, G. and Puga, D. (2020). The Economics of Urban Density. *Journal of Economic Perspectives*, 34(3):3–26.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Faggio, G., Silva, O., and Strange, W. C. (2017). Heterogeneous Agglomeration. *The Review of Economics and Statistics*, 99(1):80–94.

Freeman, R., Larch, M., Theodorakopoulos, A., and Yotov, Y. (2021). Unlocking new methods to estimate country-specific trade costs and trade elasticitie. School of Economics Working Paper Series 2021-17, LeBow College of Business, Drexel

University.

Fujita, M., Krugman, P. R., and Venables, A. (1999). *The spatial economy: Cities, regions, and international trade*. MIT press.

Heid, B., Larch, M., and Yotov, Y. V. (2020). Estimating the Effects of Non-discriminatory Trade Policies within Structural Gravity Models. School of Economics Working Papers 2020-06, University of Adelaide, School of Economics.

Heid, B., Larch, M., and Yotov, Y. V. (2021). Estimating the effects of non-discriminatory trade policies within structural gravity models. *Canadian Journal of Economics/Revue canadienne d'économique*, 54(1):376–409.

Josephson, A., Ricker-Gilbert, J., and Florax, R. (2014). How does population density influence agricultural intensification and productivity? evidence from ethiopia. *Food Policy*, 48:142–152. cited By 84.

Klein Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E. (2017). Anthropogenic land use estimates for the holocene – hyde 3.2. *Earth System Science Data*, 9(2):927–953.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5):950–959.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics*, 9(4):469–479.

Mayer, T. and Zignago, S. (2011). Notes on cepii's distances measures: The geodist database. Working Papers 2011-25, CEPII.

Morrison, T. (1984). Cereal imports by developing countries. trends and determinants. *Food Policy*, 9(1):13–26. cited By 10.

Moscona, J. and Levy, A. (2022). Specializing in cities: Density and the pattern of trade. *Available at SSRN: https://ssrn.com/abstract=4259355 or http://dx.doi.org/10.2139/ssrn.4259355*.

Nakamura, R. (1985). Agglomeration economies in urban manufacturing industries: A case of japanese cities. *Journal of Urban Economics*, 17(1):108–124.

Nunn, N. and Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics*, 94(1):20–36.

Query, J. (2022). Gross product, population distribution and heterogeneity of border effects in gravity models of trade. *Open Economies Review*. cited By 0.

Ramondo, N., Rodríguez-Clare, A., and Saborío-Rodríguez, M. (2016). Trade, domestic frictions, and scale effects. *American Economic Review*, 106(10):3159–84.

Redding, S. and Venables, A. (2004a). Geography and export performance: external market access and internal supply capacity. In *Challenges to globalization: Analyzing the economics*, pages 95–130. University of Chicago Press.

Redding, S. and Venables, A. J. (2004b). Economic geography and international inequality. *Journal of International Economics*, 62(1):53–82.

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics*, 101(C):148–167.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9(1):21–58.

Redding, S. J. and Weinstein, D. E. (2019). Aggregation and the gravity equation. In *AEA Papers and Proceedings*, volume 109, pages 450–455. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Ricker-Gilbert, J., Jumbe, C., and Chamberlin, J. (2014). How does population density influence agricultural intensification and productivity? evidence from malawi. *Food Policy*, 48:114–128. cited By 65.

Rosenthal, S. S. and Strange, W. C. (2004). Chapter 49 - evidence on the nature and sources of agglomeration economies. In Henderson, J. V. and Thisse, J.-F., editors, *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2119–2171. Elsevier.

Samuelson, P. A. (1952). The transfer problem and transport costs: The terms of trade when impediments are absent. *The Economic Journal*, 62(246):278–304.

Sellner, R. (2019). Non-discriminatory trade policies in panel structural gravity models: Evidence from monte carlo simulations. *Review of International Economics*.

Yamarik, S. and Ghosh, S. (2005). A sensitivity analysis of the gravity model. *International Trade Journal*, 19(1):83–126. cited By 29.

Yotov, Y. (2021). The variation of gravity within countries (or 15 reasons why gravity should be estimated with domestic trade flows) by yoto yotov :: Ssrn. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3842321. (Accessed on 07/15/2021).

Yotov, Y. V. (2022). On the role of domestic trade flows for estimating the gravity model of trade. *Contemporary Economic Policy*, 40(3):526–540.

Yotov, Y. V., Piermartini, R., Monteiro, J.-A., and Larch, M. (2016). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. WTO.

A Appendix A

A.1 Multilateral Resistance Terms Derivation

To define these terms in theory, we start from goods market clearing 7 and including 6

$$Y_i = (\frac{w_i}{A_i})^{-\theta} \sum_{j=1}^{N} \frac{\left(\tau_{ij}\right)^{-\theta}}{\sum_{j=1}^{N} \left(\frac{w_i \tau_{ij}}{A_i}\right)^{-\theta}} w_j L_j \tag{40}$$

Normalize 40 by world income as in (Freeman et al., 2021), $\sum_{i=1}^{N} Y_i = Y$, and substitute the denominator with the price parameter of the price distribution 10 and $E_j = w_j L_j$:

$$\frac{Y_i}{Y} = (\frac{w_i}{A_i})^{-\theta} \sum_{j=1}^{N} \frac{\left(\tau_{ij}\right)^{-\theta}}{\left(\Phi_j\right)^{-\theta}} \frac{E_j}{Y} \tag{41}$$

As stated in equation 11, price index P_j is proportional to Φ_j and equation 41 takes the form:

$$\frac{Y_i}{Y} = (\frac{w_i}{A_i})^{-\theta} \sum_{j=1}^{N} \frac{\left(\tau_{ij}\right)^{-\theta}}{\gamma^{\theta}(P_j)^{-\theta}} \frac{E_j}{Y} \tag{42}$$

and then it is possible to obtain Multilateral resistance terms, the Outward (OMR):

$$\Pi_i = \sum_{j=1}^{N} \left(\frac{\tau_{ij}}{P_j}\right)^{-\theta} \frac{E_j}{Y} \tag{43}$$

and the Inward Multilateral Resistance Term (IMR):

$$P_i = \sum_{i=1}^{N} \left(\frac{\tau_{ij}}{\Pi_i}\right)^{-\theta} \frac{Y_i}{Y} \tag{44}$$

A.2 Obtain wages including country productivity and output function to add in the gravity

Rewrite the trade equation, 6:

$$X_{ij} = \frac{[(A_i)^{\theta}(w_i)^{-\theta}]}{(P_j)^{-\theta}} E_i \tag{45}$$

Combine equation 42 with the OMR, P_j, terms and solve for $[(A_i)^{\theta}(w_i)^{-\theta}]$ and obtain:

$$[(A_i)^{\theta}(w_i)^{-\theta}] = \frac{Y_i/Y}{(\Pi_i)^{-\theta}} \tag{46}$$

To obtain the standard structural gravity equation substitute 46 in 45:

$$X_{ij} = \frac{Y_i E_j}{Y} \left(\frac{\tau_{ij}}{\Pi_i P_j}\right)^{-\theta} \tag{47}$$

A.3 Derivation of the theoretical interpretation of density coefficient and the international border dummy.

Here is presented a generalization of the problem, deriving all the elements contained in the productivity A_i, adding a specific parameter like η does not change the algebra to obtain the results in section

Starting from the log-transformed trade shares:

$$ln(\pi_{ij}) = \theta ln(A_i) - \theta(ln(w_i - ln(\tau_{ij})) - ln\Phi_j \tag{48}$$

assume $ln(A_i) = t$, then $ln(\pi_{ij}) = \theta t - ... - ln\Phi_j(t)$:

$$\frac{\partial ln(\pi_{ij})}{\partial t} = \theta - \frac{1}{\Phi_j}\frac{d}{dt}(e^{\theta t})(w_i\tau_{ij})^{-\theta} \tag{49}$$

$$\frac{\partial ln(\pi_{ij})}{\partial t} = \theta - \frac{1}{\Phi_j}\theta(e^{\theta t})(w_i\tau_{ij})^{-\theta} \tag{50}$$

$e^{\theta t} = e^{\theta ln(A_i)} = A_i^{\theta}$, $\Phi_{ij} = A_i^{\theta}(w_i\tau_{ij})^{-\theta}$ and $\pi_{ij} = \Phi_{ij}/\Phi_j$, meaning that:

$$\frac{\partial ln(\pi_{ij})}{\partial ln(A_i)} = \theta - \theta\frac{\Phi_{ij}}{\Phi_j} = \theta(1 - \pi_{ij}) \tag{51}$$

The same procedure applies to $ln(\tau_{ij})$, the result is different because of $-\theta$ and yields:

$$\frac{\partial ln(\pi_{ij})}{\partial ln(\tau_{ij})} = \theta(\pi_{ij} - 1) < 0 \tag{52}$$

Defining the differential effect of international trade costs with respect to domestic trade costs:

$$\frac{\partial ln(\pi_{ij})}{\partial ln(\tau_{ij})} - \frac{\partial ln(\pi_{jj})}{\partial ln(\tau_{jj}^k)} = \theta(\pi_{ij} - \pi_{jj}) \tag{53}$$

B Appendix B

B.1 Cross-Section

Table 8: PPML Results 1: Within Manufacturing (Cross-Section)

	(1)	(2)	(3)	(4)
	Food/Bev./Tob.	Textiles	Wood paper prod.	Wood cork prod.
$Log(Density) \times INTL_{ij}$	0.201***	0.367***	0.361***	0.232***
	(0.04)	(0.08)	(0.10)	(0.06)
N	4096	4096	4096	4096
	(5)	(6)	(7)	(8)
	Coke/petr. prod.	Chemic./Pharma	Rubber/Plast.	Other non-metal.
$Log(Density) \times INTL_{ij}$	0.530***	0.246***	0.202**	0.451***
	(0.10)	(0.05)	(0.06)	(0.06)
N	4096	4096	4096	4096
	(9)	(10)	(11)	(12)
	Basic metal	Fabric. metal	Computer/electro.	Electric. equip.
$Log(Density) \times INTL_{ij}$	0.039	0.293***	0.226	0.220
	(0.06)	(0.06)	(0.16)	(0.11)
N	4096	4096	4096	4096
	(13)	(14)	(15)	(16)
	Machin.	Motor veich.	Other trans.	Other manuf.
$Log(Density) \times INTL_{ij}$	0.338***	-0.054	0.380***	0.495***
	(0.07)	(0.12)	(0.11)	(0.07)
N	4096	4096	4096	4096
Exporter FE	✓	✓	✓	✓
Importer FE	✓	✓	✓	✓
GRAVITY	✓	✓	✓	✓
$INTL_{ij}$	✓	✓	✓	✓
GEO Controls X $INTL_{ij}$	✓	✓	✓	✓
Clusters	Pair	Pair	Pair	Pair

Note: Manufacturing sectors are the one provided by TiVA (version 2018) $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (Gem diamond extraction 1958-2000 (1000 carats)), *near coast* (percentage Within 100 km. of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$. Clustered by pair (exporter-importer, non-symmetric) robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

Table 9: Alternative Cross Section Estimates 2015: method of Freeman et al. (2021) - Robustness

VARIABLES	(1) Manufacturing	(2) Agric. Forest. and Fish.	(3) Mining
$Log(Density) \times INTL_{ij}$	0.4840***	-0.1413**	-0.3241***
	(0.0576)	(0.0628)	(0.1130)
$Log(\hat{\Pi}_i)$	-1.2246***	0.1460	0.3282***
	(0.0992)	(0.1549)	(0.1262)
Observations	4,061	3,844	3,340
Exporter FE	NO	NO	NO
Importer FE	YES	YES	YES
GRAVITY	YES	YES	YES
$INTL_{ij}$	YES	YES	YES
GEO Control	YES	YES	YES
Clusters	Pair	Pair	Pair

Note: The difference in sample size in different sectors is due to singleton and duplicates which in Agriculture et al and Mining are dropped by the importer and exporter fixed effects. The small reduction of the observations in all samples is due to the first-stage estimates. $GRAVITY$ concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). $GEO\_CONTROL$ contains the variables from Nunn and Puga (2012) and they are *ruggedness*, *soil fertility* (percentage of land), *tropical climate*, *desert percentage of land*, *gemstones* (gem diamond extraction 1958-2000, 1000 carats), *near coast* (percentage within 100 km of ice-free coast), all are multiplied by the international border dummy $INTL_{ij}$. $OMR(i)$ is computed as in Freeman et al. (2021). Clustered by pair (exporter-importer, non-symmetric) robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1.

B.2 Panel

Table 10: PPML Gross Exports (Pair FE), 2005-2015

VARIABLES	(1) Manufacturing	(2) Agric. Forest. and Fish.	(3) Mining
$Log(Density) \times INTL_{ij}$	0.9616*	0.9704*	-1.2013
	(0.5557)	(0.5545)	(1.3494)
Observations	44,671	42,284	36,740
Exporter X Time FE	YES	YES	YES
Importer X Time FE	YES	YES	YES
Pair FEs	YES	YES	YES
GRAVITY	NO	NO	NO
$INTL_{ij}$	NO	NO	NO
GEO Control X $INTL_{ij}$	NO	NO	NO
Clusters	Exporter × Importer	Exporter × Importer	Exporter × Importer

Note: The difference in sample size in different sectors is due to singleton and duplicates which in Agriculture et al. are dropped by the importer-time and exporter-time fixed effect. *GRAVITY* concerns *log. of weighted distance*, *contiguity* (dummy) and *common official language* (dummy) from Conte et al. (2022). 2-way clustered robust standard errors in parentheses. Significance level *** p<0.01, ** p<0.05, * p<0.1

Table 11: Trade Share by broad sectors in TiVA

	2015					
	Manufacturing		Agriculture, Forestry and Fisheries		Mining	
	group by i	grouped by j	group by i	grouped by j	group by i	grouped by j
π_{jj}	0.6483	0.6021	0.8412	0.8469	0.7131	0.5548
π_{ij}	0.3516	0.3979	0.1588	0.1531	0.2869	0.4452
Total	0.3563	0.4011	0.1695	0.1640	0.2936	0.4469

Table 12: η values from *Method 1* (Heid et al., 2021)

θ	η	η Upper Bound	η Lower Bound
1	1.57	2.07	1.06
2	0.78	1.04	0.53
3	0.52	0.69	0.35
4	0.39	0.52	0.27
6	0.26	0.35	0.18
7	0.22	0.30	0.15
8	0.20	0.26	0.13
11	0.14	0.19	0.10
12	0.13	0.17	0.09

Table 13: η values from *Method 2* (Freeman et al., 2021)

θ	η	η Upper Bound	η Lower Bound
1	2.35	2.86	1.83
2	1.17	1.43	0.91
3	0.78	0.95	0.61
4	0.59	0.72	0.46
6	0.39	0.48	0.3
7	0.34	0.41	0.26
8	0.29	0.36	0.23
11	0.21	0.26	0.17
12	0.2	0.24	0.15