PPML, Gravity, and Heterogeneous Trade Elasticities*

Vladimir Tyazhelnikov† Xinbei Zhou‡

September 2023

Abstract

The gravity equation, the most popular empirical tool in International Trade, is usually estimated by the OLS or Poisson Pseudo-Maximum Likelihood (PPML), with PPML being robust to the heteroskedasticity of an error term. We find that when the trade elasticity is heterogeneous between country pairs, OLS and PPML estimates have different interpretations: OLS estimates the average elasticity and PPML the elasticity of the average. We show that most of the differences between the PPML and OLS estimates are explained by the difference in the interpretation of the coefficients, with only 8-30% of the differences explained by the heteroskedasticity channel.

JEL Classification: F10, F14, C13, C21, C50

Keywords: elasticity, heterogeneity, heteroskedasticity, gravity model, misspecification

---

*We thank Peter Exterkate, Scott French, Nezih Guner, Vasiliy Korovkin, Ye Lu, Rachael Meager, Mark Melatos, John Romalis, and Xuetao Shi for early suggestions.

†School of Economics, The University of Sydney, Sydney, Australia 2006. Corresponding author. (vladimir.tyazhelnikov@sydney.edu.au)

‡HKU Business School, The University of Hong Kong, Hong Kong SAR, China. (xbeizhou@connect.hku.hk).

1 Introduction

In this paper, we focus on the behaviour of the techniques commonly used to estimate constant elasticity models in the presence of unobserved heterogeneity. Wooldridge (2005) shows that in the case that heterogeneity is independent of the covariates, the ordinary least squares (OLS) estimator generates estimates that can be interpreted as an average partial effect (APE). It follows then that the coefficient from a log-log OLS regression can be interpreted as average elasticity in the case of coefficient heterogeneity.

While average elasticity is a meaningful object, depending on the question, researchers might be interested in the elasticity of the average, the effect of a one per cent change in one variable on the percentage change of the average value of another variable.[1] In this paper, we use the gravity equation, one of the most widely used empirical tools in the field of international trade, to illustrate this difference. We show that the estimates obtained from Poisson pseudo maximum likelihood (PPML), the most common econometric technique used to estimate the gravity equation, can be interpreted as the elasticity of the average.[2]

Previously, all the differences between PPML and OLS coefficients in the gravity equation have been wrongly attributed to the fact that, unlike PPML, OLS estimates are biased when there is heteroskedasticity in an error term, a point that Santos Silva and Tenreyro (2006) (henceforth SST) raised.

In this paper, we obtain evidence of heterogeneity in the effect of distance on trade volume. Furthermore, we show, both theoretically and with Monte Carlo simulations, that the presence of such heterogeneity causes bias in the estimation of the elasticity of the average by the OLS.[3] We introduce a weighted least squares (WLS) estimator and show that both the PPML and WLS estimators are robust to heterogeneity. Still, unlike PPML and similar

[1] Consider the simple numerical example below. In a hypothetical country, there are only two individuals, one rich and one poor. The rich individual has an income of \$1,000,000, while the poor individual has an income of only \$1,000, and they face heterogeneous tax rates of 10% and 1%, respectively. The average tax rate is calculated as 5.5%, but the total tax revenue is \$100, 000 + \$10 = \$100, 010. The percentage of total tax revenue in terms of total income is $\frac{100{,}010}{1{,}001{,}000} = 9.991\%$, which can be interpreted as the "tax rate on average". The latter is the true effect of tax rates on the whole economy. The unweighted average of tax rates differs from the true effect because it neglects the different weights that the two individuals have in the economy and includes both of them in the computation of average elasticity with equal weights. To accurately calculate the "tax rate on average", the heterogeneous tax rates for the two individuals should be assigned weights according to their share of the total income of the whole country: $10\% * \frac{1{,}000{,}000}{1{,}001{,}000} + 1\% * \frac{1{,}000}{1{,}001{,}000} = 9.991\%$. Similarly, countries of different sizes would have various degrees of impact on global trade volume and should enter into the gravity equation computation with different weights. Otherwise, using the OLS estimator to estimate the effects of a uniform decrease in trade costs on global trade volumes, such as a worldwide reduction in gasoline prices, will lead to biased results.

[2] Ciani and Fisher (2019) make a related point and show that the interpretation of OLS and PPML dif-in-dif estimates of a multiplicative model differ.

[3] In this paper, we focus on the estimation of the elasticity of the average; from this point of view, OLS estimates are biased. Alternatively, we can say that OLS and PPML estimates have different interpretations.

to OLS, WLS is not robust to the heteroskedasticity of an error term.

The fourth estimation procedure we consider is Gamma pseudo maximum likelihood (GPML), which, similar to PPML, is consistent when the error term is heteroskedastic but is often overlooked because of its potentially lower efficiency compared to PPML. We show that, similar to OLS, the GPML estimates can be interpreted as the average elasticity, which explains why GPML and PPML estimates often differ in practice.

A comparison of OLS, PPML, WLS, and GPML estimates allows us to decompose the bias caused by using the OLS estimator with the log-linearized gravity equation into two different sources — heteroskedasticity in the error term and heterogeneity across country pairs — using PPML estimate as a benchmark. We document that the share of bias in estimating the distance elasticity caused by heterogeneity is much larger than that caused by heteroskedasticity. In particular, the heteroskedasticity bias accounts for approximately 8% of the total difference between OLS and PPML for the estimates of the average elasticity and 30% for the estimates of the elasticity of the average.

We further contribute to the gravity literature and examine heterogeneous trade shocks in the context of heterogeneous trade elasticities. Expectedly, the effect of a given shock is contingent on which observations were treated respective intensity of such treatment. We show that a standard PPML estimator can be interpreted as the average effect of a uniform 1% decrease in trade costs on the world economy. We show that a PPML estimator augmented with treatment intenstity weights, we call WPPML, consistently estimates the effect of a given arbitrary trade shock.

While the topic of heterogeneous country-pair trade elasticity has rarely been discussed, a few empirical studies examine the distance elasticity across countries. Fratianni and Kang (2006) first use the log form of the gravity equation to estimate the distance elasticity of the full sample, which yields a significant estimate of -1.17. Then, they test whether the homogeneity assumption for different groups of countries holds with two tests: one for Organisation for Economic Co-operation and Development (OECD) countries and non-OECD countries and the other for Christian and Islamic countries; both tests reject the null hypothesis of distance homogeneity at the 1% level. They find that the distance elasticity is much smaller for OECD member countries in absolute values than for nonmember countries. They also find that the trading cost represented by the distance is the largest when trade occurs between an Islamic and Christian country, with a distance elasticity of -1.47. Magerman et al. (2016) summarize multiple studies on the distance and border effects in international trade that perform sensitivity tests of the effects for various countries, regions, and periods with different methods; the authors confirm the presence of heterogeneity in distance and trade elasticities across different country pairs.

A discussion of the mechanisms behind the heterogeneity of trade and distance elasticity is beyond the scope of this paper. Still, several studies provide microfoundations of the gravity equation with country-pair heterogeneity. Fieler (2011) first confirms that low-income countries trade less than rich countries, both with each other and with the rest of the world: in 2000, transactions to and from the 12 Western European countries accounted for 45% of global trade, while the 57 African countries accounted for only 4.2%. This fact suggests that pairs of large countries dominate global trade flows and, therefore, should be given more weight in calculating the "elasticity of the average". The author relaxes two of the traditional assumptions of trade models that generate the gravity equation by allowing for non-homothetic preferences and different quality of goods to be produced in high- and low-income countries. With this new model, the author can explain the prevalence of large trade flows among developed countries and small trade flows among developing countries. Novy (2013) shows that the gravity equation based on the translog demand system, unlike the standard CES gravity model, generates heterogeneous country-pairs trade elasticities. Bas et al. (2017) make a related point but focus on the supply-side mechanism of the heterogeneity. They introduce an extension of the Melitz (2003) model with the productivity distribution of heterogeneous firms following log-normal instead of the commonly used Pareto distribution and show that this assumption leads to bilateral-specific aggregate trade elasticity for each country pair.

Our paper fits into a large strand of literature devoted to the estimation of the gravity equation. A detailed review of this literature can be found in Head and Mayer (2014) and Yotov et al. (2016). The former briefly considers the case of heterogeneous elasticities in the Monte Carlo simulations and documents that OLS and GPML successfully estimate the average elasticity, while PPML fails to do so, which is consistent with our findings. However, the authors interpret the difference as the lack of robustness to misspecification and attribute it to the different weights that PPML and GPML assign to various observations. Mayer et al. (2019) make a similar point that PPML puts more weight on country pairs with larger trade flows; they, nevertheless, do not consider the presence of country-pair heterogeneous elasticity and do not discuss the difference in interpretation of the estimates.[4]

Finally, we contribute to the strand of literature devoted to the issues associated with the aggregation of trade data. Costinot and Rodríguez-Clare (2014), Imbs and Mejean (2015) Kehoe et al. (2017), and French (2017) address issues caused by between-industry heterogeneity. Imbs and Mejean (2017) focus on the issues that arise when aggregating sectoral elasticities in case these elasticities exhibit heterogeneity on a country level. Coughlin

[4]Eaton et al. (2012) also discuss that the differences between PPML and OLS can be driven by different weights applied to trade flows.

and Novy (2016) analyze the consequences of the spatial aggregation of trade data. Larch et al. (2019) show that OLS and PPML estimates diverge when there is a large number of small countries in the data. Arvis and Shepherd (2013) find that the sum of predicted and actual trade flows are equal only for the PPML. Redding and Weinstein (2019) compare aggregate and sectoral gravity equations and derive an exact Jensen's inequality correction term, which drives the difference between estimates from those two equations.

The rest of the paper is organized as follows: we discuss the methodology and theoretical background in Section 2. In Section 3, we present the results of Monte Carlo simulations. In Section 4, we employ the gravity equation to prove the existence of country-pair heterogeneous elasticities and decompose the differences between OLS and PPML estimates. In Section 5, we conclude the paper.

2 Methodology

Heterogeneity

In this paper, we focus on the case of unobserved heterogeneity. The reason is that if the variation in observed independent variables drives the heterogeneity, this issue can be addressed by including the interaction term between the variable of interest and the variable that causes heterogeneity. If the unobserved variable is correlated with other covariates, then a suitable proxy variable can address this issue, while the rest of the procedure would remain the same.

We rely on the approach by Wooldridge (2005) to handle the unobserved heterogeneity and incorporate it into a constant elasticity model. A general formulation is

$$y = \beta_0 x^{\beta} x^{q\gamma} v,$$

where y and x are the dependent and control variables, respectively: q is an unobserved variable; and v is an error term that satisfies the unit conditional mean assumption: $E(v|\mathbf{x}, q) = 1$.[5]

The log-linearized version of this expression is then:

$$\log y = \beta_0 + \beta \log x + \gamma q \log x + \log v$$

The partial effect of $\log x$ on $E(\log y|\mathbf{x}, q)$ is then

[5] Here we focus on the heterogeneity driven bias, so, for now, we assume that the error term is homodeskedastic.

$$\frac{\partial E\left(\log y|\log x,q\right)}{\partial \log x}=\left(\beta+\gamma q\right).$$

As the unobserved term q differs from observation to observation, this partial effect is nonconstant, and we interpret it as the observation-specific elasticity. The formulation above can be interpreted as a random coefficient model with the observation-specific elasticity $\theta_i \equiv \beta + \gamma q$; to simplify the notation from now on, we will rely on this interpretation of the unobserved heterogeneity model.

Wooldridge (2005) shows that in the case that q is independent of $\mathbf{x}$, $E(q) = \mu$, OLS estimates the average partial effect (APE) of $\log x$ on $E(\log y|\mathbf{x}, q)$, which is equal to $\beta + \gamma\mu$. This estimate can be interpreted as the average elasticity, as it averages out the partial effects. Because of Jensen's inequality, however, the average elasticity and elasticity of the average are not equal. If there are N observations with values of y_i and corresponding elasticities θ_i, then the average elasticity is simply $\frac{\sum_{i=1}^{N}\theta_i}{N}$. The elasticity of the average can be computed as a weighted average of elasticities with weights equal to the share of y_i in $\sum_{i=1}^{N} y_i$: $\frac{\sum_{i=1}^{N}\theta_i y_i}{\sum_{i=1}^{N} y_i}$.[6] It follows that to calculate the elasticity of the average, we do not need to know the values of unobserved variable q, the weights are proportional to the observed values of y, and thus, an estimator that applies these weights can be interpreted as the elasticity of the average.

The gravity equation and heteroskedastic errors

In this section, we use the "naive form" of the gravity equation in Head and Mayer (2014) to be consistent with the constant-elasticity models in SST and constrain the trade elasticity to be distance elasticity:[7]

$$Y_{ij}=\frac{GDP_i^{\beta_1}*GDP_j^{\beta_2}}{Distance_{ij}^{\theta}}\varepsilon_{ij} \tag{1}$$

where Y_{ij} is the trade flow between country i and j, $Distance_{ij}$ stands for bilateral distance, and ε_{ij} is the error term.

Taking logs on both sides of the equation yields the following:

$$\log Y_{ij}=\beta_1\log gdp_i+\beta_2\log gdp_j+\theta\log distance_{ij}+\log\varepsilon_{ij} \tag{2}$$

where $\theta < 0$ is the trade elasticity.[8]

[6] Note that this Jensen's inequality differs from one described in SST. In the general case, with both heteroskedasticity and heterogeneity, there are two sources of bias driven by different Jensen's inequalities.

[7] In this paper, we follow SST as close as possible, to model heterogeneous trade elasticities, however, we have to deviate and use a more complex data generating process.

[8] The multiplicative error term ε_{ij} used in our model is derived from the "true" additive error term

To obtain consistent estimates of the coefficients in Equation 1 using the log-linearized form of Equation 2 by OLS, it is necessary for $\mathbb{E}[\log \varepsilon_{ij}|\mathbf{X}]$ to be constant. According to SST, however, ε_{ij} is generally heteroskedastic, so running a simple OLS regression of $\log Y_{ij}$ on $\mathbf{X}$ leads to inconsistent estimates of θ. Moreover, when there is heteroskedasticity in the error term, the log transformation of the gravity equation potentially leads to the violation of the exogeneity assumption and biases the OLS estimator. From Equation 2:

$$\hat{\log} \varepsilon_{ij} = \log Y_{ij} - \hat{\log} Y_{ij}$$

where $\hat{\log} Y_{ij} = \mathbf{X}\hat{\beta}$ is the predicted log of bilateral trade volume. Even though the exogeneity assumption $Cov(\varepsilon_{ij}, \mathbf{X}) = 0$ is satisfied in Equation 2, with heteroskedasticity in ε_{ij}, $Cov(\log \varepsilon_{ij}, \mathbf{X})$ does not necessarily equal 0. Thus, the OLS estimator is biased.

Country-pair heterogeneity

To focus on estimating the aggregate elasticity that reflects the effect of a uniform change in the trade costs on bilateral trade volumes, i.e., the elasticity the of average, we incorporate heterogeneous country-pair trade elasticities into Equation 2:

$$\log Y_{ij} = \beta_1 \log gdp_i + \beta_2 \log gdp_j + \theta_{ij} \log distance_{ij} + \log \varepsilon_{ij} \tag{3}$$

In the case that the error term is homoskedastic, the OLS estimator $\hat{\theta}$ is a consistent estimator for APE: $\overline{\theta}^{APE} = E(\theta_{ij})$ given that the assumption of exogeneity with heterogeneity is satisfied: $E(\theta_{ij} \mid \log distance_{ij}) = E(\theta_{ij})$. This estimator was defined previously as the "average trade elasticity". However, using the OLS to estimate the "elasticity of the average" leads to bias with heterogeneity because $\overline{\theta}^{APE}$ assigns equal weight to every individual θ_{ij}, while different trading country pairs have a differential impact on the world's aggregate trade flows.

A natural solution candidate to the weighting problem is the WLS estimator. The share of the bilateral trade flows between countries i and j over the global trade flows, denoted by $\frac{Y_{ij}}{\mathbf{Y}}$, can be used as the weights. The formal justification of this method follows the solutions to the endogenous stratified sampling issue proposed by Hausman and Wise (1981), where they use the Gary Income Maintenance Experiment as an example to show the extent of selection bias due to endogenous stratified sampling and demonstrate that the bias can be corrected by both the maximum likelihood estimator (MLE) and WLS estimator.

In the context of this study, trade elasticity can be interpreted as the effect of trade

$\eta_{ij} = Y_{ij} - \hat{Y_{ij}}$ where $\varepsilon_{ij} = 1 + \frac{\eta_{ij}}{\exp(\mathbf{X} * \beta)}$ and $\mathbb{E}[\varepsilon_{ij}|\mathbf{X}] = 1$.

liberalization on the representative dollar of world trade flows. If every country pair is treated as a stratum, then the probability of a representative dollar falling in any stratum is $\frac{1}{M}$, where M is the number of country pairs. The share of the bilateral trade flows between a country pair is $\frac{Y_{ij}}{\mathbf{Y}}$. The dollars from a country pair are underrepresented when $\frac{Y_{ij}}{\mathbf{Y}} > \frac{1}{M}$. The relative probability of a particular dollar falling in a stratum is given by $\frac{1}{M} / \frac{Y_{ij}}{\mathbf{Y}} = \frac{\mathbf{Y}}{M * Y_{ij}}$. Therefore, with normalization, the weight used in this study is $\frac{Y_{ij}}{\mathbf{Y}}$.[9]

We follow Hausman and Wise (1981) by using weights that include an error term; to ensure that endogenous weights are not driving our results, we propose an alternative 2-stage procedure, which we refer as to feasible weighted least squares (FWLS), where in the first stage, we run an OLS regression and generate predicted trade flows. In the second stage, we run a WLS regression with the fitted values of trade flows from the previous stage as weights. We also use PPML and GPML estimators in this study following the method discussed by SST.

Table 1 below summarises the empirical moment conditions for each of the estimators discussed in this study.

TABLE 1
Empirical Moment Conditions

OLS	$\sum_{i,j \in \Omega} X_{ij} \left(\log Y_{ij} - \widehat{\log Y_{ij}} \right) = 0$	GPML	$\sum_{i,j \in \Omega} X_{ij} \left(\frac{Y_{ij}}{\hat{Y}_{ij}} - 1 \right) = 0$
WLS	$\sum_{i,j \in \Omega} X_{ij} \left(\log Y_{ij} - \widehat{\log Y_{ij}} \right) Y_{ij} = 0$	PPML	$\sum_{i,j \in \Omega} X_{ij} \left(Y_{ij} - \hat{Y}_{ij} \right) = 0$
FWLS	$\sum_{i,j \in \Omega} X_{ij} \left(\log Y_{ij} - \widehat{\log Y_{ij}} \right) \hat{Y}_{ij} = 0$		

Note: $\hat{Y}_{ij}$ and $\widehat{\log Y_{ij}}$ are the predicted bilateral and log of trade volume correspondingly. $\mathbf{X_{ij}}$ is the vector of covariates for the country pair ij. Ω is the set of all country pairs ij.

From these conditions, one can see that OLS involves log deviations of Y_{ij} from its predicted value. Since percent deviations are approximately equal to log deviations, multiplying the log deviations by actual trade volumes leads to a result close to the level deviations, implied by the WLS first-order condition.

The moment condition of PPML indicates that it also involves the level deviations of Y_{ij} from its predicted value; hence, similar to WLS, it addresses the heterogeneity issue. Conversely, the moment condition of GPML relies on the per cent deviations of Y_{ij} from its predicted value, so we can expect that the GPML and OLS estimator will yield similar estimates.

[9] Mayer et al. (2019) estimate the gravity equation and apply weights proportional to the value of trade costs, which is similar to our WLS estimator. They, however, do not discuss the interpretation of the estimates and the presence of country-pair heterogeneous elasticities.

3 Simulations

In this section, we simulate bilateral trade flows under the assumptions of two types of countries and heterogeneous trade elasticities across country pairs. The simulation results indicate that using the OLS estimator under such conditions yields biased estimates for the parameter of interest (the elasticity of the average) and that the application of the PPML and WLS estimators corrects this bias.

Data generating process

We incorporate the heterogeneous elasticities into SST's specifications of heteroskedastic errors in the data generating process (DGP):

$$Y_{ij} = \exp(\beta_1 \log gdp_i + \beta_2 \log gdp_j + \theta_{ij} \log distance_{ij}) * \varepsilon_{ij} \tag{4}$$

where ε_{ij} is a log-normal random variable with mean 1 and variance σ^2.

Following SST, we consider the following three specifications of σ^2 to assess the performance of different estimators under different patterns of heteroskedasticity.[10]

Case 1: $\sigma^2 = \exp(\log Y_{ij})^{-2}$; $V[Y_{ij}|\mathbf{X}] = 1$.

Case 2: $\sigma^2 = \exp(\log Y_{ij})^{-1}$; $V[Y_{ij}|\mathbf{X}] = \exp(\log Y_{ij})$.

Case 3: $\sigma^2 = 1$; $V[Y_{ij}|\mathbf{X}] = \exp(\log Y_{ij})^2$.

We consider a simple economy in which there are two types of countries, large and small, with $GDP_L = 100$ and $GDP_S = 10$, respectively. The total number of countries in the economy is $N = 100$. The effect of GDP on bilateral trade flows is assumed to be 1, i.e., $\beta_1 = \beta_2 = 1$. The distance between a country pair is randomly generated from the uniform distribution $U(2, 3)$, and a counterfactual case with a 1% increase in each pairwise distance is also generated to reflect the changes in trade costs.

In terms of heterogeneity, the value of trade elasticity between two countries θ_{ij} can be either -1 or -0.5. We first examine the scenario where $Pr[\theta_{ij} = -1] = Pr[\theta_{ij} = -0.5] = 0.5$. This satisfies the exogeneity with heterogeneity assumption in Wooldridge (2005) that the unobserved heterogeneity is independent of the structural covariates. Then we rely on the findings of Fieler (2011) and assume that the trade elasticity between two large countries is $\theta_{LL} = -0.5$ and that the elasticity for the pairs of two small or one small and one large country is $\theta_{SS} = \theta_{SL} = -1$. Our results, however, do not depend on the source of heterogeneity and in general, we are agnostic about it.[11]

[10] Due to the differences between our DGP and that in SST, the fourth case of heteroskedasticity is omitted.

[11] While GDPs are observed and controlled for by fixed effects, its effect on distance elasticity is of bilateral

We start by applying the DGP when all countries are small, i.e., $N_L = 0$, and increase the number of large countries in the economy by one after each iteration until $N_L = 100$. For each of the $K = 101$ cases, we run $S = 80$ simulations with new draws of errors and bilateral distances. For each simulation $s \in S$, the elasticity of average is calculated as below:

$$\delta_s = \frac{1}{N^2}\sum_i^N \sum_j^N \left(\frac{\frac{\sum_i^N \sum_j^N Y'_{ij} - \sum_i^N \sum_j^N Y_{ij}}{\sum_i^N \sum_j^N Y_{ij}}}{\frac{distance'_{ij} - distance_{ij}}{distance_{ij}}} \right) = \frac{1}{N^2}\sum_i^N \sum_j^N \left(\frac{\frac{\sum_i^N \sum_j^N (Y'_{ij} - Y_{ij})}{\sum_i^N \sum_j^N Y_{ij}}}{\frac{distance'_{ij} - distance_{ij}}{distance_{ij}}} \right) \tag{5}$$

where $distance'_{ij}$ is the counterfactual distance with a 1% increase from $distance_{ij}$ and Y'_{ij} is the corresponding trade flows generated from Equation (4). The numerator of Equation (5) is the percentage change in average trade volumes, and the denominator is the percentage change in bilateral distance. And the average elasticity is calculated as follows:

$$\delta'_s = \frac{1}{N^2}\sum_i^N \sum_j^N \left(\frac{\frac{Y'_{ij} - Y_{ij}}{Y_{ij}}}{\frac{distance'_{ij} - distance_{ij}}{distance_{ij}}} \right) \tag{6}$$

which is the average ratio of the percentage change in trade volumes and the percentage change in distance.

Heterogeneous Treatment

In the previous section we define the elasticity of average as the average percentage response of trade flows to a uniform 1% change in trade costs (distances). Although such a shock represents a particular case of a generic trade shock, our emphasis is placed on it due to its explicit interpretability and its immediate association with a PPML estimator.

Note, however, that in the presence of heterogeneous effects, non-uniform treatment would have different effect on the elasticity of the average depending on which observations are treated and with what intensity.[12]. In this section, we derive the expression for the elasticity of the average in case of an arbitrary treatment and show that it can be estimated by a WPPML estimator with weights proportionate to treatment intensity.

Average elasticity is calculated as

nature and is unobserved. We chose the distance elasticity to depend on GDP to generate a link between the size of trade flows and the elasticity heterogeneity. In the absence of such a link, Jensen's inequality becomes equality, and the average elasticity becomes equal to the elasticity of the average.

[12]It is true even for linear models with heterogenous coefficients.

$$\delta_s' = \frac{1}{N^2}\sum_i^N\sum_j^N\left(\frac{\frac{Y_{ij}'-Y_{ij}}{Y_{ij}}}{\frac{d_{ij}'-d_{ij}}{d_{ij}}}\right) = \frac{1}{N^2}\sum_i^N\sum_j^N \delta_{ij} \tag{7}$$

which is the average ratio of the percentage change in trade volumes and the percentage change in distance. $\delta_{ij} \equiv \frac{\frac{Y_{ij}'-Y_{ij}}{Y_{ij}}}{\frac{d_{ij}'-d_{ij}}{d_{ij}}}$ is the pair-specific elasticity.

The elasticity of the average is the ratio of total percentage change in trade flows and average percentage change in distances:

$$\delta_s = \frac{\frac{\sum_i^N\sum_j^N Y_{ij}'-\sum_i^N\sum_j^N Y_{ij}}{\sum_i^N\sum_j^N Y_{ij}}}{\sum_i^N\sum_j^N \frac{d_{ij}'-d_{ij}}{d_{ij}}/N^2} = \frac{\frac{\sum_i^N\sum_j^N (Y_{ij}'-Y_{ij})}{\sum_i^N\sum_j^N Y_{ij}}}{\sum_i^N\sum_j^N \frac{d_{ij}'-d_{ij}}{d_{ij}}/N^2} = \frac{\sum_i^N\sum_j^N \frac{(Y_{ij}'-Y_{ij})}{Y_{ij}}\frac{Y_{ij}}{\sum_i^N\sum_j^N Y_{ij}}}{\sum_i^N\sum_j^N \frac{d_{ij}'-d_{ij}}{d_{ij}}/N^2} \tag{8}$$

$$= \sum_i^N\sum_j^N \frac{\frac{(Y_{ij}'-Y_{ij})}{Y_{ij}}\frac{Y_{ij}}{\sum_i^N\sum_j^N Y_{ij}}}{\sum_i^N\sum_j^N \frac{d_{ij}'-d_{ij}}{d_{ij}}/N^2} = \sum_i^N\sum_j^N\left(\frac{\frac{(Y_{ij}'-Y_{ij})}{Y_{ij}}}{\frac{d_{ij}'-d_{ij}}{d_{ij}}}\frac{Y_{ij}}{\sum_i^N\sum_j^N Y_{ij}}\frac{\frac{d_{ij}'-d_{ij}}{d_{ij}}}{\sum_i^N\sum_j^N \frac{d_{ij}'-d_{ij}}{d_{ij}}/N^2}\right) \tag{9}$$

$$= \sum_i^N\sum_j^N\left(\delta_{ij}\frac{Y_{ij}}{\sum_i^N\sum_j^N Y_{ij}}s_{ij}\right) = \frac{\sum_i^N\sum_j^N \delta_{ij}Y_{ij}s_{ij}}{N^2}/\frac{\sum_i^N\sum_j^N Y_{ij}}{N^2}, \tag{10}$$

where $s_{ij} \equiv \frac{\frac{d_{ij}'-d_{ij}}{d_{ij}}}{\sum_i^N\sum_j^N \frac{d_{ij}'-d_{ij}}{d_{ij}}/N^2}$ and $\mathbf{s} \equiv \{s_{ij}\}_{i,j\in[1,N]}$ are the weights and the vector of treatment intensity. When all the weights are equal, $(s_{ij} = s_{kl}\forall i,j,k,l \in [1,N])$, it corresponds to the case of a uniform decrease in all the trade barriers $\mathbf{d}$. Note that $\mathbf{s}$ is not random and instead represents the counterfactual of interest. Naturally, in the presence of heterogeneity, the average elasticity (average responsiveness of the dependent variable to treatment) will depend on the distribution of treatment. For example, if the counterfactual of interest is the decrease in trade barriers between countries 1 and 2 with all other trade barriers remaining unchanged ($d_{ij}' \neq d_{ij}$ if $i=1, j=2$ and $d_{ij}' = d_{ij}$ if $i \neq 1, j \neq 2$), the true elasticity of such a treatment will be equal to δ_{12}.

Now, for a given treatment vector of interest $\mathbf{s}$, a policymaker can properly estimate the average elasticity by assigning corresponding weights to the relevant observations. In particular, if only a subset of observations are being treated, untreated observations should be omitted from the estimation (otherwise, values of δ_{ij} irrelevant to a given treatment would contaminate the true elasticity of average δ_s).[13]. More generally, if some observations receive larger treatment than others (say, trade costs between Germany and Italy decreased

[13]Consider, for example, the case when there are two subpopulations with different true values of trade elasticity δ_1 and δ_2. Then, if only the first group was treated, the responsiveness of the group to this treatment will correspond to δ_1, while values of δ_2 are completely irrelevant.

by 99% and trade costs between France and Poland decreased by 0.01%), the weight of their elasticity in calculation of the total elasticity should be higher.

Note that unlike size weights $\frac{Y_{ij}}{\sum_i^N \sum_j^N Y_{ij}}$, which appear because we deal with a non-linear model, treatment weights would appear in any model, including linear, with heterogeneous coefficients. Now, the regular PPML estimator considers the baseline case of a uniform treatment $\mathbf{s} = 1$; if a policymaker is interested in the outcomes of any other treatment vector, a weighted PPML (WPPML) can be applied. Finally, there exists such a treatment vector such that the average elasticity and elasticity of the average are identical; it happens when larger observations (country pairs) receive proportionally lower treatment and thus two weights (size and treatment) cancel out.

Interestingly, average elasticity δ_s' does not depend on the treatment vector, because it is a simple arithmetic average of heterogenous elasticities δ_{ij} and does not account the relative size and importance of different observations or whether these observations were treated or not (which is another advantage of the elasticity of the average).

Finally, the last expression for δ_s is the empirical counterpart of $\frac{\mathbb{E}[Y\theta s|\mathbf{s}]}{\mathbb{E}[Y]}$. In the case of a uniform treatment, $\mathbf{s} = 1$, this expression becomes $\frac{\mathbb{E}[Y\theta]}{\mathbb{E}[Y]}$, corresponding to the moment condition of the regular (unweighted) PPML estimator.

Estimates

In this section, we investigate the behavior of the two distinct parameters of interest, the elasticity of average and the average elasticity, and the performance of different estimators for models with heterogeneous country pairs in the simulated economy and the above-mentioned heteroskedastic errors. The WLS estimator is applied with the share of every country pair in total world trade $Y_{ij}/\mathbf{Y}$ as weights, where $\mathbf{Y}$ is the total trade volumes across all country pairs. For the FWLS estimator, the share of predicted trade flows $\hat{Y_{ij}}/\mathbf{Y}$ is used as weights. The GPML and PPML estimators use Y_{ij} as the dependent variable and the log of GDPs and log of bilateral distance as independent variables, while all other estimators use $\log Y_{ij}$ as the dependent variable.

To provide a clearer picture of the differences between estimators, We express the bias of using different estimators as the percentage difference between the estimates and the elasticity of the average as follows:

$$bias_k = 1 - \frac{\hat{\theta_k}}{\delta_k} \tag{11}$$

where $k \in K$ is the number of large countries in the economy, $\hat{\theta_k}$ is the estimated result when there are k large countries in the economy. The elasticity of the average in the k^{th} case, δ_k, is calculated as the average of all the δ_s (s = 1, ... 80).

We start by focusing on the case where unobserved heterogeneity is assumed to be independent of the country sizes, so each country pair will have an equal opportunity of having a trade elasticity of -1 or -0.5. Figure 1 displays the behaviour of all estimators of the trade elasticity. As a reference, we also include the elasticity of the average of each heterogeneity case, δ_k, and the average elasticity, δ'_k [14], in the figures of estimates.

The stark difference between the values of the elasticity of the average, δ_k, and the average elasticity, δ'_k, is straightforward from the left panel graphs of Figure 1. Consequently, the behaviour of the selected estimators differentiates from each other and centres around the two distinctive concepts. Firstly, PPML, WLS and FWLS have similar patterns as δ_k, while GPML and OLS have similar patterns as δ'_k. This is consistent with our grouping of the estimators' empirical moment conditions in Table 1. In terms of estimated values, PPML and GPML estimates are close to δ_k, while OLS and FWLS estimates are close to δ'_k, and WLS estimates are close to neither.

The distinction between δ_k and δ'_k is due to the existence of country-pair heterogeneity and heteroskedasticity in the error term. SST shows that both PPML and GPML estimators address the issue of heteroskedastic errors, so the difference between these two estimators results from the fact that PPML assigns different weights to country pairs in the sample while GPML does not. When country-pair heterogeneity is randomly determined, the weighting does not seem to be crucial, as the biases from these two estimators are almost negligible when Case 1 and 2 types of heteroskedastic error are present (Figure 1b and 1d). On the other hand, OLS, WLS and FWLS estimators all suffer from having heteroskedastic errors. While WLS and FWLS estimators also attempt to address the country-pair heterogeneity issue by assigning different weights to observations, they all result in notable biases from the parameter of interest, δ_k.

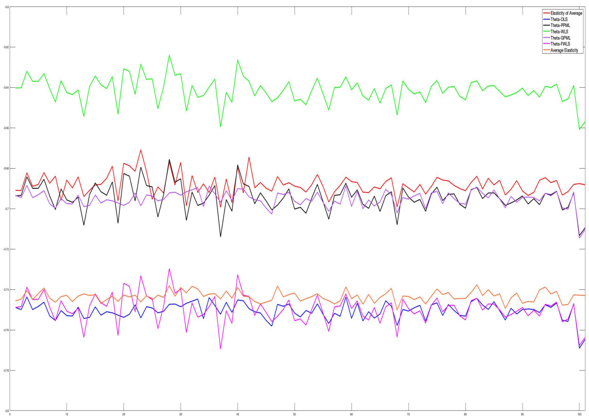
As discussed by SST, Case 3 is a special case that the log linear model corrects the heteroskedasticity in the data and that the GPML is the optimal PML estimator in this case. Therefore, in Figure 1f, the GPML estimator still has a negligible bias, while the bias of the PPML estimator, in this case, is much larger.

Next, we follow Fieler (2011) and introduce the dependence of heterogeneity on country size, where trade elasticities between a pair of large countries will be larger in absolute value than those between country pairs involving a small country. Before analyzing the main estimated results, we present a summary table below (Table 2) for the benchmark results where no country-pair heterogeneity is considered, and all countries are small. Recall that

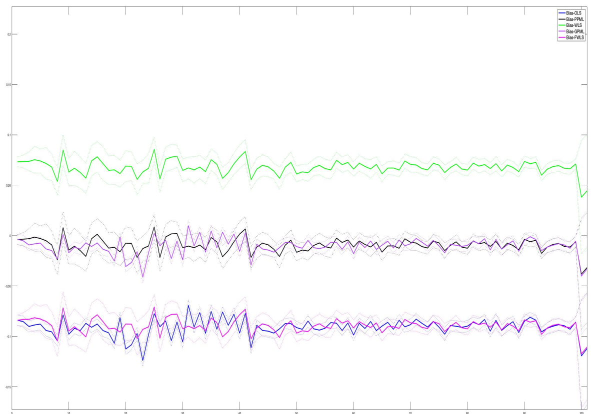
[14] This is calculated as the average of all θ_s over $s \in S$ for each $k \in K$.

Figure 1. Estimates and biases of OLS, WSL, GPML and PPML (Independent Heterogeneity)

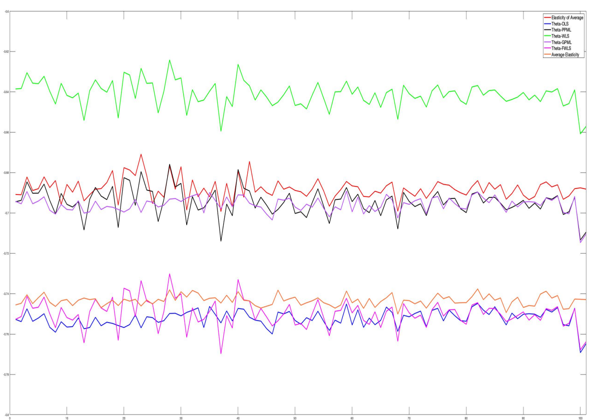
(a) Estimates Case 1



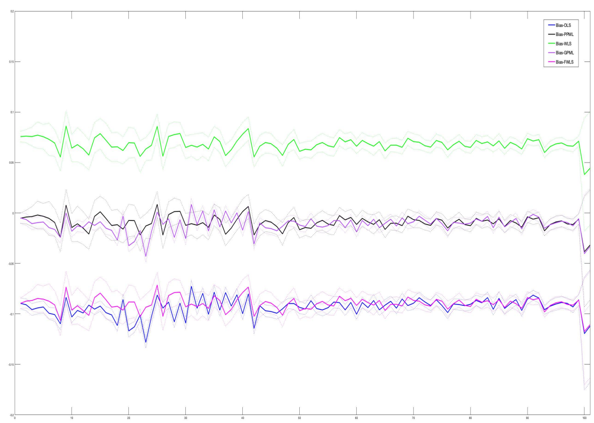
(b) Biases Case 1



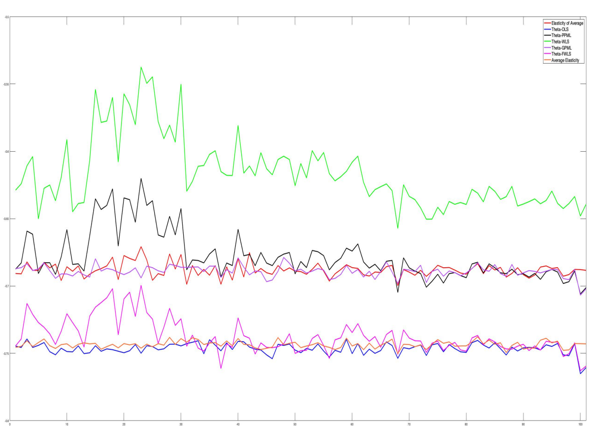
(c) Estimates Case 2



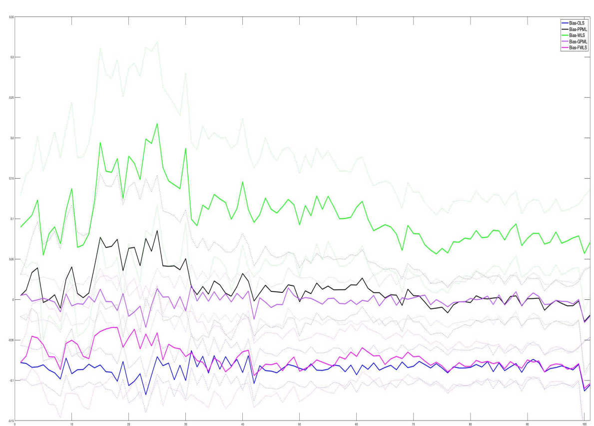
(d) Biases Case 2



(e) Estimates Case 3



(f) Biases Case 3



Note: The horizontal axis shows the number of large countries in the economy, and the vertical axis is the estimated result. Cases 1, 2 and 3 heteroskedastic errors follow the same specification as in Section 3.1. Trade elasticity (elasticity of average) is calculated as the change in level trade volume with a 1% increase in bilateral distance and is shown in red. Average elasticity is calculated as the unweighted average of the true θ's in the DGP and is shown in orange. Biases are calculated using Equation 5, which are the percentage difference between the elasticity of average and the estimated coefficient for different estimators. The 95% confidence intervals of the estimated biases are included as dashed lines.

Table 2
Sample estimates and biases

$N_L = 0$	Case 1		Case 2		Case 3	
	Estimate	Bias	Estimate	Bias	Estimate	Bias
OLS	-0.99997 (-1.00010, -0.99980)	-0.000067 (-0.000230, 0.000097)	-1.00340 (-1.00500, -1.00170)	-0.003456 (-0.005137, -0.001775)	-0.99043 (-1.00500, -0.97584)	0.009467 (-0.005128, 0.024062)
GPML	-0.99988 (-1.00000, -0.99972)	0.000016 (-0.000148, 0.000181)	-0.99883 (-1.00050, -0.99715)	0.001069 (-0.000610, 0.002479)	-0.98833 (-1.00570, -0.97094)	0.011568 (-0.005824, 0.028959)
PPML	-0.99989 (-1.00000, -0.99973)	0.000012 (-0.000145, 0.000169)	-0.99886 (-1.00050, -0.99724)	0.001038 (-0.000586, 0.002661)	-0.98788 (1.00470, -0.97110)	0.012024 (-0.004754, 0.028802)
WLS	-0.99981 (-0.99996, -0.99965)	0.000095 (-0.000062, 0.000252)	-0.99436 (-0.99599, -0.99273)	0.005541 (0.003913, 0.007170)	-0.98379 (-1.01020, -0.95735)	0.016107 (-0.010338, 0.042553)
FWLS	-0.99997 (-1.00010, -0.99981)	-0.000071 (-0.000227, 0.000086)	-1.00340 (-1.00500, -1.00170)	-0.003466 (-0.005093, -0.001839)	-0.99047 (-1.00480, -0.97617)	0.009429 (-0.0004876, 0.023735)
$N_L = 50$	Case 1		Case 2		Case 3	
	Estimate	Bias	Estimate	Bias	Estimate	Bias
OLS	-0.88112 (-0.88294, -0.87931)	-0.567010 (-0.570290, -0.563730)	-0.88177 (-0.88369, -0.87985)	-0.568160 (-0.571610, -0.564710)	-0.87162 (-0.88598, -0.85726)	-0.549710 (-0.575180, -0.524240)
GPML	-0.86736 (-0.86918, -0.86554)	-0.542540 (-0.545820, -0.539250)	-0.86653 (-0.86852, -0.86455)	-0.541060 (-0.544630, -0.537490)	-0.85408 (-0.87129, -0.83686)	-0.518510 (-0.549030, -0.487980)
PPML	-0.56238 (-0.56264, -0.56212)	-0.000141 (-0.000616, 0.000333)	-0.56220 (-0.56254, -0.56186)	0.000176 (-0.000433, 0.000784)	-0.53943 (-0.56987, -0.50889)	0.040914 (-0.013219, 0.095047)
WLS	-0.56130 (-0.56157, -0.56104)	0.001769 (0.001284,0.002253)	-0.56101 (-0.56135, -0.56067)	0.002284 (0.001670, 0.002897)	-0.53531 (-0.58083, -0.48978)	-0.032658 (-0.032658,0.129260)
FWLS	-0.57515 (-0.57545, -0.57485)	-0.022858 (-0.023411, -0.22305)	-0.57494 (-0.57531, -0.57457)	-0.002483 (-0.023143, -0.021824)	-0.55603 (-0.58076, -0.53130)	0.011342 (-0.032656, -0.055340)

Note: This table contains both the case with no country-pair heterogeneity and the heterogeneous case with the number of large and small countries both equal to 50. N_L is the number of large countries in the economy. Case 1, 2, and 3 heteroskedasticity follow the corresponding specifications in Section 3.1. Biases are calculated using equation 11. The estimated 95% confidence intervals are included in brackets.

we have $\theta_{LL} = -0.5$ and $\theta_{SS} = \theta_{SL} = -1$. Under these conditions, according to SST, we could expect good performance from the PPML estimator across all three specifications of heteroskedasticity, while the OLS estimator only performs well under Case 3. As shown in Table 2, under Case 1 and 2, PPML gives an estimate closer to -1, the universal trade elasticity, than the OLS estimator with smaller standard errors. The reverse is true under Case 3, consistent with SST's findings under their DGP.

Then we introduce country-pair heterogeneity into the benchmark by making half of the countries large. As shown in the second half of Table 2, the OLS estimates are drastically different from the PPML ones. We also include WLS estimates to show the similarity between the estimated results using WLS and PPML, which is consistent with the discussion in Section 3. The estimated results from FWLS are reasonably close to those of WLS, indicating that endogenous weights do not significantly impact our results. Furthermore, while GPML performs well under both specifications of heteroskedasticity when there are no large countries in the economy, it fails to address the issue of heterogeneity, resulting in similar results as OLS and much larger biases than PPML in estimating the elasticity of the average.

Figure 2 displays the behaviour of all estimators of the trade elasticity under this circumstance together with the aforementioned heteroskedasticity. With the increasing number of large countries in the economy, we should expect an increase in the average of estimated elasticities. As before, we also include the elasticity of the average of each heterogeneity case, δ_k, and the average elasticity, δ'_k, in the figures of estimates.

It is, again, clear from the left panel graphs that OLS and GPML estimators behave differently than other estimators and the elasticity of the average, δ_k, but are in close alignment with the average elasticity, δ'_k. In general, the PPML, WLS, and FWLS estimators outperform OLS and GPML, with both smaller deviations from δ_k and similar behaviour as the number of large countries in the economy k increases. The differences between the two categories of estimators are significant when there is more heterogeneity in the economy, i.e., the number of large and small countries are close to each other. In contrast, in the case of little heterogeneity in the economy, the estimates are not drastically different. These two points indicate that in our simulations, country-pair heterogeneity dependent on country size generates a larger share of bias than heteroskedastic errors, which confirms our point that the PPML, WLS, and FWLS estimators are capable of addressing heterogeneity-related issues better than OLS and GPML.

Figures 2b, 2d, and 2f represent biases for Cases 1, 2 and 3, respectively. We can see that despite the distinct patterns of heteroskedasticity in the error term, the effect of this distinction is trivial on both the shape of the graph and the scale of the bias. Here, the bias caused by country-pair heterogeneity dependent on country size outweighs the bias caused by heteroskedastic errors, so unlike the results in Figure 1, it is crucial to assign proper weights to observations if we want to estimate the elasticity of average, δ_k, correctly. [15]

The results in Table 2 are also presented in these graphs. In Figure 2a, when the number of large countries $k = 0$ (no large countries) and $k = 100$ (no small countries), i.e., no heterogeneity, the PPML estimates are the closest to the elasticity of the average δ. Conversely, the OLS estimator performs best without heterogeneity in Case 3 (Figure 2e).

The dashed lines show the narrow 95% confidence interval of the biases under each case, indicating that the differences across estimators are statistically significant.

4 Gravity equation

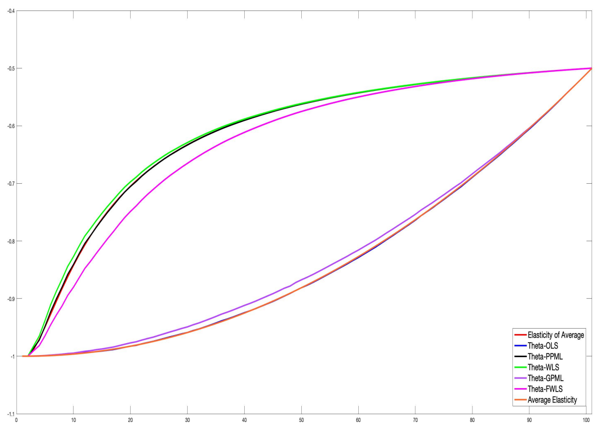
In this section, we apply different estimation techniques to gravity data. We document not only the existence of country-pair heterogeneity but also decompose the bias of OLS estimates by heterogeneity and heteroskedasticity channels.

To make our results comparable with those in SST, we employ the same data and follow their specification by including distance and indicator variables for remoteness, common

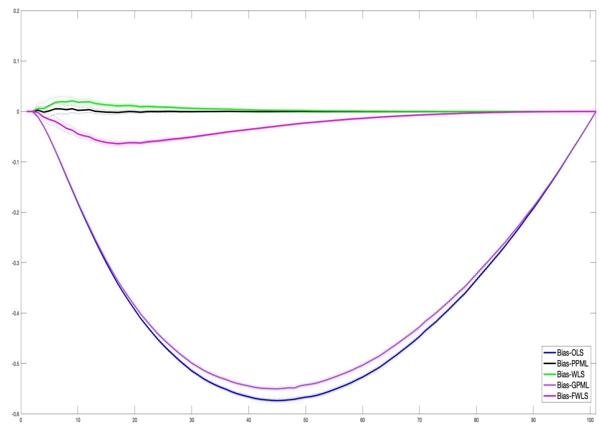
[15]Note that the estimates and biases are similar for the cases of $N_L = 0$ and $N_L = 1$ because in both cases, there are no country pairs where both countries are large, hence in both cases the country-pair elasticities are uniform.

Figure 2. Estimates and biases of OLS, WSL, GPML and PPML

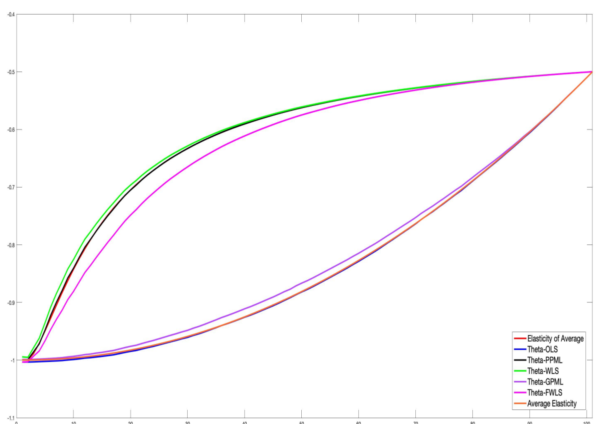
(a) Estimates Case 1



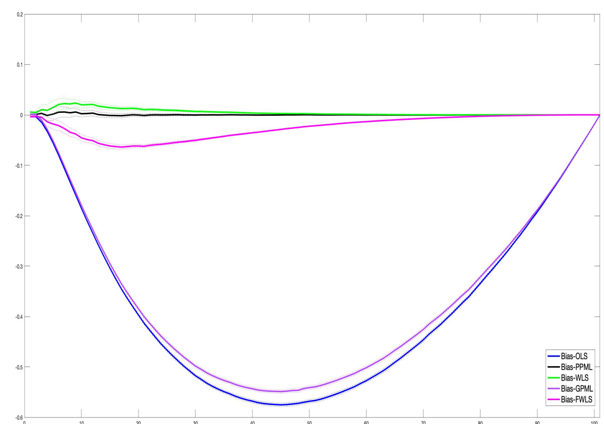
(b) Biases Case 1



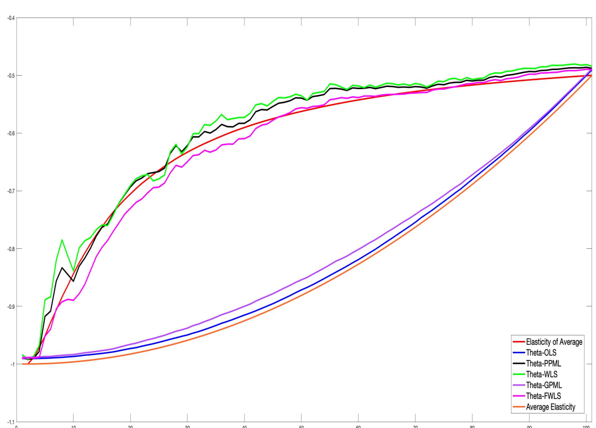
(c) Estimates Case 2



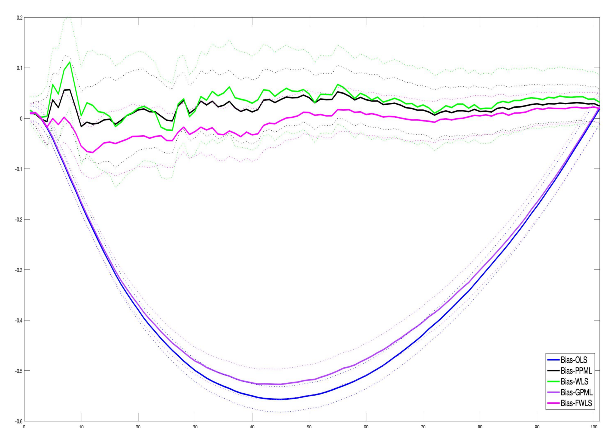
(d) Biases Case 2



(e) Estimates Case 3



(f) Biases Case 3



Note: The horizontal axis shows the number of large countries in the economy, and the vertical axis is the estimated result. Cases 1, 2 and 3 heteroskedastic errors follow the same specification as in Section 3.1. Trade elasticity (elasticity of average) is calculated as the change in level trade volume with a 1% increase in bilateral distance and is shown in red. Average elasticity is calculated as the unweighted average of the true θ's in the DGP and is shown in orange. Biases are calculated using Equation 5, which are the percentage difference between the elasticity of average and the estimated coefficient for different estimators. The 95% confidence intervals of the estimated biases are included as dashed lines.

language, colonial heritage, and preferential trade agreements in the model. We replicate the findings of SST's Table 5 of the gravity equation regression on OLS, PPML and GPML estimators, controlling for multilateral resistance terms with importer and exporter fixed effects, and provide the estimated results of WLS and FWLS estimators. Table 3 reports the results of all five estimators and corresponding standard errors.

TABLE 3

Gravity equation

	OLS	*WLS*	*FWLS*	*PPML (With ZTF)*	*PPML (No ZTF)*	*GPML (With ZTF)*	*GPML (No ZTF)*
Distance	-1.347***	-0.722***	-0.685***	-0.750***	-0.770***	-1.933***	-1.173***
	(0.031)	(0.049)	(0.038)	(0.041)	(0.049)	(0.055)	(0.029)
Common	0.174	0.395***	0.141**	0.369***	0.352***	-0.457**	0.326***
Border	(0.130)	(0.080)	(0.057)	(0.091)	(0.090)	(0.230)	(0.117)
Common	0.406***	0.419***	0.558***	0.383***	0.418***	0.681	0.416***
Language	(0.068)	(0.076)	(0.052)	(0.093)	(0.094)	(0.097)	(0.065)
Colonial	0.666***	-0.047	0.019	0.079	0.038	0.807***	0.510***
Ties	(0.070)	(0.103)	(0.098)	(0.134)	(0.134)	(0.100)	(0.067)
FTA	0.310***	0.424***	0.585***	0.376***	0.374***	1.472***	0.579***
	(0.098)	(0.080)	(0.085)	(0.077)	(0.077)	(0.257)	(0.085)
Exporter FE	Y	Y	Y	Y	Y	Y	Y
Importer FE	Y	Y	Y	Y	Y	Y	Y
N	9,613	9,613	9,613	18,360	9,613	18,360	9,613

Note: The OLS, WLS, FWLS, PPML, and GPML columns represent the estimates of the main regression estimated by the ordinary least squares, weighted least squares, fitted weighted least squares, Poisson pseudo-maximum likelihood, and gamma pseudo-maximum likelihood, respectively. Zero trade flows are excluded from the specifications of OLS, WLS, and FWLS. For PPML and GPML, the results from both cases of excluding and including zero trade flows in the main regression are reported. Standard errors are in parentheses. N stands for the number of observations. Details are in the main text.

First, we present the evidence obtained from the gravity dataset that country-pair elasticities are not homogeneous. Naturally, we cannot establish a link between the unobserved variable and distance elasticity; trade volumes, conversely, are observed. If the OLS estimates of the distance elasticity do not vary much depending on how large trade flows are, then there will not be a significant difference between the average elasticity and the elasticity of the average. The reason is that trade shares serve as weights while computing the elasticity of the average.

After sorting the bilateral trade flows into ascending order, we run OLS regression multiple times, dropping the 10 smallest trade flows in the sample at each iteration. The estimated $\hat{\theta}$ and corresponding 95% confidence interval are saved and plotted in Figure 3a. This is equivalent to running regressions on different subsamples containing various levels of trade flows. As the number of iterations increases, the weight of large trade volumes increases as well. The graph demonstrates an unambiguously positive relationship between estimated $\hat{\theta}$ and larger average trade flows in the subsamples. Compared to the results generated from the same procedure but with 10 trade flows being randomly dropped at every iteration on Figure 3b, it is clear that a positive relationship between the estimated $\hat{\theta}$ and larger trade volumes in the subsamples provides evidence of country-pair heterogeneity.[16] Note that we included exporter and importer fixed effects and standard controls for country pairs, so the results in Figures 3a and 3b are driven by unobserved heterogeneity.[17]

As shown in the previous section, the WLS, FWLS, and PPML estimators are robust to the presence of country-pair heterogeneity when estimating the "elasticity of the average" in the simulated economy, and the GPML estimator yields similar results as the OLS estimator, which estimates the average elasticity. Now we show that these methods lead to similar results with the same gravity dataset. We still drop the 10 smallest trade values on each iteration but now apply the WLS, FWLS, PPML, and GPML estimators to the regression equation. For WLS, the weights are calculated as the share of each observation's bilateral trade flows on the global total trade flows. For FWLS, the weights are calculated as the share of each observation's predicted bilateral trade flows on the sum of total predicted trade flows.

Compared to Figure 3a, the estimated $\hat{\theta}$ from WLS, FWLS, and PPML illustrated in Figure 3c do not depend on the choice of subsamples except for the case when the sample size becomes very small. The overall constant estimates show that these estimators are robust to country-pair heterogeneity. The three estimators exhibit a similar pattern, indicating that all three of them are capable of addressing the heterogeneity in the country-pair elasticity problem, consistent with our findings in Section 3. Moreover, the estimated $\hat{\theta}$ from GPML exhibits an increasing trend as the average trade flows in the subsamples grow larger. This result is also consistent with our findings in the previous section that GPML and OLS would yield similar results.

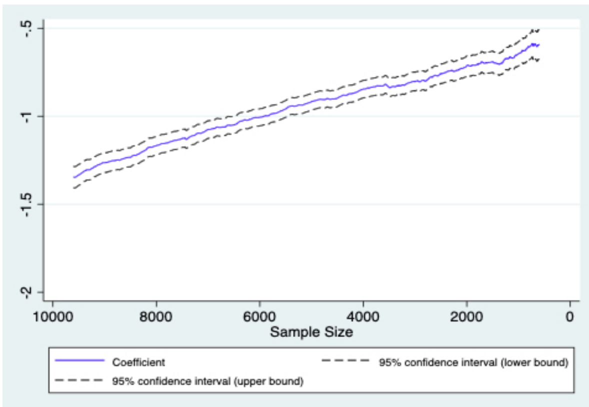
With PPML as the benchmark estimator, we can perform a decomposition to show the proportion of bias caused by heterogeneity and heteroskedasticity. The total bias is calculated as the difference between the OLS estimates and PPML estimates. According to our previous results, however, there are two alternative ways to decompose the total bias. With one method, we calculate the bias caused by heterogeneity as the difference between

[16] The standard errors for the estimates with smaller subsamples in Figure 3b are much larger because in small subsamples the number of countries increases relative to the number of observations, and consequently more fixed effects are included leading to higher standard errors of $\hat{\theta}$.

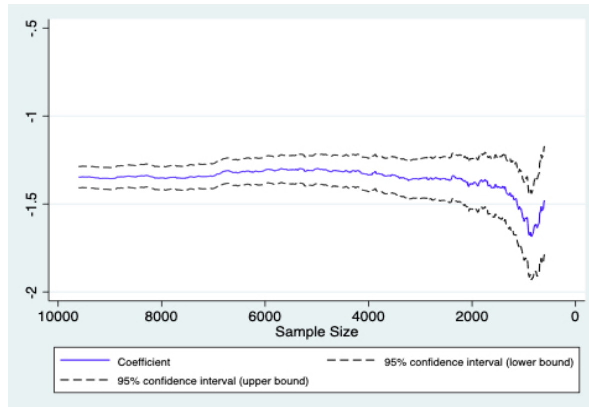
[17] Our findings are consistent with the study Yotov (2012) and Larch et al. (2019).

Figure 3. Evidence of heterogeneity

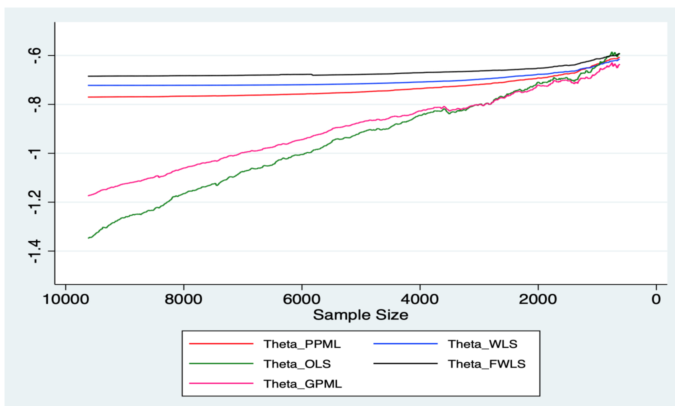
(a) Increasing trade volumes in subsamples - OLS



(b) Random subsamples - OLS



(c) Comparison across all estimators



Note: This figure presents evidence of unobserved heterogeneity in country-pair elasticity. In panel (a), at each iteration, we drop 10 observations with the smallest corresponding volume of trade, while in panel (b), we drop 10 random observations. In panel (c), we sort the data into ascending order and show the OLS, WLS, FWLS, GPML, and PPML estimates on different subsamples. There are 9,613 observations in the full sample. The last iteration has 623 observations.

the OLS estimates and WLS estimates, and the bias caused by heteroskedasticity is the difference between the WLS estimates and PPML estimates. With another method, we calculate the bias caused by heterogeneity as the difference between the GPML estimates and PPML estimates, and the bias caused by heteroskedasticity is the difference between the OLS estimates and GPML estimates. Table 4 summarizes the estimated $\hat{\theta}$, the corresponding standard errors, and the bias decomposition results.

The difference between the two methods is that the former evaluates the size of the heteroskedasticity bias when estimating the average elasticity, while the latter evaluates it for the estimation of the elasticity of the average. Not surprisingly, the sizes of biases obtained by these two methods differ — as evident from our Monte Carlo simulations, the size and the sign of the heteroskedasticity bias depend on the pattern of heteroskedasticity and the degree of heterogeneity.

Table 4 indicates that bias caused by heteroskedasticity accounts for a smaller share of the total bias, while heterogeneity is the primary driver of the differences between the estimators.[18] These findings echo the conclusion of SST that the PML estimators are the best option for estimating the trade elasticity and provide a fresh perspective on the interpretation of the estimates obtained with different methods.

TABLE 4

Bias decomposition

(A) *Decomposition through WLS*

	OLS	*WLS*	*PPML*
Estimates	-1.347 (0.031)	-0.722 (0.049)	-0.770 (0.042)
Heterogeneous Bias Percentage	-0.625 108.32%	0	0
Heteroskedastic Bias Percentage	0.048 8.32%	0.048 100%	0
Total Bias	-0.577	0.048	0

(B) *Decomposition through GPML*

	OLS	*GPML*	*PPML*
Estimates	-1.347 (0.031)	-1.173 (0.029)	-0.770 (0.042)
Heterogeneous Bias Percentage	-0.403 69.84%	-0.403 100%	0
Heteroskedastic Bias Percentage	-0.174 30.16%	0	0
Total Bias	-0.577	0.403	0

Note: Total bias is computed as the difference between OLS and PPML estimates. Heteroskedastic bias is computed as the difference between WLS and PPML estimates or the difference between OLS and GPML estimates. Heterogeneous bias is computed as the difference between OLS and WLS estimates or the difference between GPML and PPML estimates. Details are in the main text.

5 Conclusion

In this paper, we argue that the interpretation of PPML regression, the most popular method for estimating the gravity equation, is different from the results generated by the previously dominant OLS regression. We show that the former can be interpreted as the elasticity of the average and the latter as the average elasticity.

We use Monte Carlo simulations to show that when distance elasticity is systematically heterogeneous between country pairs, OLS and GPML cannot be used to estimate the elasticity of the average, while WLS and PPML are appropriate methods.

We employ trade data and find evidence of the existence of heterogeneity across different country pairs. Furthermore, a comparison of OLS, WLS, GPML, and PPML allows us to decompose the difference between OLS and PPML estimates into two channels: pre-

[18]French (2017) performs an alternative bias decomposition between OLS and PPML, focusing on the aggregation properties of both estimators. In this paper, we are agnostic about the source of the heterogeneity elasticity, but if it is industrial heterogeneity, our findings are close to those in French (2017).

viously extensively studied, heteroskedasticity bias and different interpretations of the two estimators, which we call bias in the estimation of the elasticity of the average.

We find that while the bias associated with heteroskedasticity exists, it is 7 to 12 times smaller and has the opposite sign compared to the previous findings; it provides a choice between PPML and OLS another perspective: when choosing between these two estimation techniques, we recommend making a decision based on the desired interpretation of the results or providing the results obtained with both methods. As the presence of heteroskedasticity bias may still be a problem, to obtain the estimates of the average elasticity, we recommend applying GPML.

While in this paper, we focused on properties of OLS, GPML, and PPML in estimating the gravity equation, our findings are much more general: any standard log-log regression estimated by PPML will have a different interpretation, which, depending on the research question, may be preferable to the interpretation of a standard OLS regression.

References

Arvis, J. F.and Shepherd, B. (2013). ‘The poisson quasi-maximum likelihood estimator: A solution to the ‘adding up’ problem in gravity models’, *Applied Economics Letters*, Vol. 20, pp. 515–519.

Bas, M., Mayer, T.and Thoenig, M. (2017). ‘From micro to macro: Demand, supply, and heterogeneity in the trade elasticity’, *Journal of International Economics*, Vol. 108, pp. 1–19.

Ciani, E.and Fisher, P. (2019). ‘Dif-in-dif estimators of multiplicative treatment effects’, *Journal of Econometric Methods*, Vol. 8(1).

Costinot, A.and Rodríguez-Clare, A. (2014). ‘Trade theory with numbers: Quantifying the consequences of globalization’, *Handbook of international economics*, Vol. 4, pp. 197–261.

Coughlin, C. C.and Novy, D. (2016). ‘Estimating border effects: The impact of spatial aggregation’, *FRB St. Louis Working Paper*.

Eaton, J., Kortum, S. S.and Sotelo, S. (2012). *International trade: Linking micro and macro* (tech. rep.). National bureau of economic research.

Fieler, A. C. (2011). ‘Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation’, *Econometrica*, Vol. 79, pp. 1069–1101.

Fratianni, M.and Kang, H. (2006). ‘Heterogeneous distance–elasticities in trade gravity models’, *Economics Letters*, Vol. 90, pp. 68–71.

French, S. (2017). ‘Comparative advantage and biased gravity’, *UNSW Business School Research Paper*.

Hausman, J. A.and Wise, D. A. (1981). ‘Stratification on endogenous variables and estimation: The gary income maintenance experiment’, *Structural analysis of discrete data with econometric applications*, pp. 365–391.

Head, K.and Mayer, T. (2014). ‘Gravity equations: Workhorse, toolkit, and cookbook’, *Handbook of International Economics*, Vol. 4, pp. 131–195.

Imbs, J.and Mejean, I. (2015). ‘Elasticity optimism’, *American Economic Journal: Macroeconomics*, Vol. 7, pp. 43–83.

Imbs, J.and Mejean, I. (2017). ‘Trade elasticities’, *Review of International Economics*, Vol. 25, pp. 383–402.

Kehoe, T. J., Pujolas, P. S.and Rossbach, J. (2017). ‘Quantitative trade models: Developments and challenges’, *Annual Review of Economics*, Vol. 9, pp. 295–325.

Larch, M., Wanner, J., Yotov, Y. V.and Zylkin, T. (2019). 'Currency unions and trade: A ppml re-assessment with high-dimensional fixed effects', *Oxford Bulletin of Economics and Statistics*, Vol. 81, pp. 487–510.

Magerman, G., Studnicka, Z.and Van Hove, J. (2016). 'Distance and border effects in international trade: A comparison of estimation methods', *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 10, pp. 1–31.

Mayer, T., Vicard, V.and Zignago, S. (2019). 'The cost of non-europe, revisited', *Economic Policy*, Vol. 34(98), pp. 145–199.

Melitz, M. J. (2003). 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica*, Vol. 71, pp. 1695–1725.

Novy, D. (2013). 'International trade without ces: Estimating translog gravity', *Journal of International Economics*, Vol. 89, pp. 271–282.

Redding, S. J.and Weinstein, D. E. (2019). 'Aggregation and the gravity equation', *AEA Papers and Proceedings*, Vol. 109, pp. 450–55.

Santos Silva, J. M. C.and Tenreyro, S. (2006). 'The log of gravity', *The Review of Economics and statistics*, Vol. 88, pp. 641–658.

Wooldridge, J. M. (2005). 'Unobserved heterogeneity and estimation of average partial effects', *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg*, pp. 27–55.

Yotov, Y. V. (2012). 'A simple solution to the distance puzzle in international trade', *Economics Letters*, Vol. 117, pp. 794–798.

Yotov, Y. V., Piermartini, R., Monteiro, J.-A.and Larch, M. (2016). *An advanced guide to trade policy analysis: The structural gravity model*. World Trade Organization Geneva.