Comparative Advantage and Biased Gravity

Scott French*

July, 2019

Abstract

This paper considers the consequences of aggregate estimation of nonlinear empirical models with two-way heterogeneous, multiplicative fixed effects. Aggregate estimators cannot control for micro-level interactions of these effects and are thus misspecified. I characterize the bias in aggregate estimates and propose a set of disaggregate pseudo-maximum likelihood (PML) estimators that control for the unobserved effects using a structural gravity equation. I apply these estimators to bilateral trade data, where the micro-level heterogeneity has the interpretation of product-level comparative advantage (PLCA), and find significant bias due to PLCA in more aggregated estimates. After controlling for PLCA, remaining biases due to heteroskedasticity, sample selection, and heterogeneity in the common parameters are relatively small. I also show that the pooled product-level Poisson PML estimator has a number of desirable properties, including that it estimates an ideal index of heterogeneous coefficients and outperforms a product-by-product estimator out-of-sample. Applied to panel data, I find that controlling for PLCA reveals a significant decline in the distance elasticity over time.

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

Keywords: nonlinear, fixed effects, aggregation, misspecification, heteroskedasticity, product-level, multi-sector

*School of Economics, University of New South Wales. Contact: scott.french@unsw.edu.au. Some of the content of this paper has been adapted from an unpublished paper previously circulated under the title "The Composition of Exports and Gravity".

1 Introduction

In the presence of micro-level heterogeneity, estimation based on aggregated data generally leads to biased estimates of model parameters (van Garderen et al., 2000). In this paper, I consider the estimation of nonlinear empirical models with two-way heterogeneous, multiplicative fixed effects and the bias in estimates based on aggregated data. When the micro-level effects are considered incidental parameters, it is tempting and convenient to ignore the possibility of heterogeneity and aggregate to the level of interest in the outcome variable. I consider a popular class of nonlinear estimators, including fixed-effects pseudo-maximum likelihood (PML) and structural gravity estimators, that consistently estimate the common parameters using disaggregated data. I show that aggregate estimators in this class must control for an additional unobservable component that accounts for micro-level interactions between the heterogeneous fixed effects. Aggregate estimation necessarily omits this component, which is generally correlated with the explanatory variables and cannot be controlled for using aggregate fixed effects, and thus yields biased estimates.

My primary application is to bilateral trade flows, where the interaction between unobserved importer-product and exporter-product effects embodies product-level comparative advantage (PLCA), which is correlated with trade costs and influences sector-level trade flows. However, this framework can be applied in many contexts in which the outcome varies along multiple dimensions, such as panel data with individual and time effects that vary by product, location, employer, academic subject, or type of health condition, or in dyadic data in which agents interact multiple times or in multiple ways, as with product-level trade flows. Often, interest lies in the effect of an explanatory variable on a unit of observation (such as a firm, worker, or country) at a higher level of aggregation than the heterogeneity in the fixed effects (e.g., the worker-employer or country-product level). The focus of this paper is on the estimation of such an effect and the bias in estimates using aggregated data. This is a common practice in many fields, including the international trade literature, where gravity estimation almost universally uses aggregate data.[1]

I propose a set of disaggregate PML estimators that control for the heterogeneous unobserved effects using a set of structural gravity equations. These estimators consistently identify the common parameters, are straightforward to implement, and can be applied outside of an international trade context. I consider pooled estimators, which maintain the assumption of homogeneous common parameters implicit in aggregate estimators, as well

[1]In the set of papers compiled by Head and Mayer (2014) for a meta-analysis of gravity estimates, more than 80% used only aggregate data. Of the papers that used disaggregated data, the median number of industries was 16. Only two used product-level data, and both used single-country datasets that did not allow them to control for multilateral effects.

as estimators that allow for heterogeneity in the common parameters. The latter can yield an overwhelming number of parameter estimates, hindering exposition and interpretation of the results. Statistical tests may also lack power because the heterogeneous parameters are less precisely estimated and require testing multiple hypotheses.[2] As a solution to both problems, I propose an ideal coefficient index whose interpretation is analogous to estimates based on aggregate data. I show that pooled Poisson PML (PPML) automatically estimates this index, making it a particularly useful estimator.

I apply these estimators to data on bilateral trade of manufactures at the 6-digit HS product level. To assess the severity of the bias in aggregate estimates, I compare the results to standard gravity estimators using data aggregated to the sector level, defined either as all manufactures or 2-digit ISIC industries. I find large and statistically significant bias in sector-level estimates. For example, the distance elasticity is biased downward by 17% for PPML and upward by 70% for fixed-effects OLS in aggregate estimations. Substantial differences across sector-level estimators – attributed to biases in OLS due to heteroskedasticity and sample selection and to finite sample biases in PML estimators by Santos Silva and Tenreyro (2006) and others – become small and statistically insignificant in product-level estimates, indicating that such biases become much less important once the estimation controls for PLCA. I also find that industry-level estimates are nearly identical to aggregate estimates for all manufactures, indicating the importance controlling for heterogeneity in unobserved effects at the lowest level of aggregation possible.

These findings make a strong case that gravity estimation should always make use of the most disaggregated data available. A remaining question is whether pooled or heterogenous estimators should generally be the preferred replacement for aggregate estimators. I address this question in two ways. First, I test for heterogeneity in coefficient estimates and find that homogeneity can be rejected for many, though far from all, products. Second, I develop and implement a split-sample bootstrap procedure to examine the out-of-sample performance of the estimators and find that the pooled estimator outperforms the heterogenous estimator, indicating that the latter overfits the product-level data in-sample. This is consistent with the conclusion of Baltagi (2008) that homogeneous panel estimators tend to perform well out of sample. Based on these results, I propose a rule of thumb: Pool to the level of aggregation of primary interest in the outcome. Pooled PPML is particularly useful in this regard, as it is a pooled estimator that also delivers an ideal index of heterogeneous coefficients, making it a strong candidate to be the workhorse estimator in applications where any form of micro-level heterogeneity is suspected.

This is not the first paper to recommend disaggregate gravity estimation. Most notably,

[2]See, e.g., Anderson (2008) for a discussion of the multiple inference problem.

Anderson and van Wincoop (2004) demonstrate that estimates suffer from aggregation bias when trade costs vary across products, and Anderson and Yotov (2010a,b) argue that parameter heterogeneity across industries is significant. By contrast, I show that aggregate estimates are biased due to PLCA even when parameters are homogeneous. In fact, because pooled PPML estimates the ideal coefficient index, *all* of the bias in aggregate PPML estimates is due to PLCA. Even for other estimators that do not share this property, the bias due to PLCA swamps the bias due to heterogeneity in my empirical application. The empirical tools that I implement also obviate many possible reasons why researchers have largely ignored the guidance of Anderson and van Wincoop (2004) that the "obvious recommendation is to disaggregate". Specifically, the pooled estimators and coefficient index alleviate the reporting, efficiency, and multiple inference concerns, and I show that the common parameters as well as the coefficient index for heterogeneous border costs can be identified without product-level data on domestic trade flows, which are often unavailable. In addition, the product-level estimators are not overly computationally burdensome.[3] Thus, I argue that the product-level estimators are unambiguously superior to sector-level estimators.

The predicted effects of changes in trade barriers based on sector-level estimates can be misleading both due to bias in estimates of the common parameters and as a direct result of ignoring patterns of PLCA. To demonstrate this, I evaluate the Modular Trade Impact (MTI) of eliminating border costs. The MTI is a very useful impact measure because it can be calculated using the common parameter estimates from a disaggregate structural gravity estimator and requires no additional data nor any restrictions on the form of demand across products, details of factor markets, or sources of comparative advantage. The product-level models predict much more consistent and substantially smaller changes in trade flows than the sector-level models. The former distinction is primarily due to bias in sector-level parameter estimates, while the latter is largely a result of sector-level MTI calculations ignoring PLCA.

I also consider the bias in sector-level panel estimates that include country-pair fixed effects to control for all static determinants of bilateral trade flows. I find that, when estimating the effect of free trade agreements (FTAs) in isolation, sector-level estimates overstate the effect by 30% for PPML and nearly 90% for OLS. When controlling for time-variation in the effects of other common gravity variables, I find less bias in the effect of FTAs, but sector-level estimates fail to identify a 4-6% per-decade fall in the distance elasticity and a 6-8% per-decade fall in the effect of sharing a common language. These results are consistent

[3]Matlab routines for the structural gravity PML estimators and fixed-effects OLS are available on the author's website. Product-level PPML and fixed-effects OLS can be implemented in Stata using the PPML_PANEL_SG and REG_HDFE commands, respectively. The baseline estimates can be computed in about 15 minutes on a standard desktop computer.

with increasing specialization over time for country pairs that are geographically close and who tend to sign FTAs, and they help explain the "puzzling" persistence of distance elasticity estimates (Disdier and Head, 2008).

Gravity estimation has a long tradition in the international trade literature (Head and Mayer, 2014). Following Eaton and Kortum (2002) and Anderson and van Wincoop (2003), the gravity literature has adopted estimation methods consistent with trade theory, using the structure of trade models to control for unobserved country-specific effects. I contribute to this literature in two ways. First, I cast the structural gravity equation as a method-of-moments estimator for a general nonlinear empirical model with two-way fixed effects. This demonstrates that structural gravity estimators are valid even when output and expenditure are observed with error, which is not true under standard derivations in which an error term is appended after imposing the equilibrium conditions of an underlying theoretical model. It also implies that structural gravity estimators are generally applicable in fields beyond international trade. Second, I show that unbiased estimation of the effects of explanatory variables on sector-level trade flows requires controlling for PLCA, and I develop a set of estimators that use product-level data to do so.

This paper is related to the literature devoted to the estimation of nonlinear models with two-way fixed effects. Fernández-Val and Weidner (2016) review this literature and derive unbiased maximum likelihood estimators that treat the unobserved effects as parameters to be estimated. Charbonneau (2012) and Jochmans (2017) derive alternative estimators that eliminate two-way fixed effects using differences of pairwise interactions of residuals. I show that the well-known structural gravity equation also constitutes a method-of-moments estimator based on a conditional mean assumption. The split sample bootstrap procedure that I implement adapts model evaluation techniques used in time series (White, 2003), panel data (Baltagi, 2008), and cross-section (Anderson, 2008; Fafchamps and Labonne, 2017) applications to a setting with dyadic data.

There is a vast literature devoted to aggregation issues in estimation.[4] This paper contributes most directly to the literature on estimation of non-linear micro-level models using aggregated data, including Lewbel (1992) and van Garderen et al. (2000). In the international trade literature, a number of papers, including Hillberry and Hummels (2002), Hillberry (2002), and Yi (2003), develop models in which the production and demand structure depends on trade costs and show that this leads to bias in aggregate estimates. While they rely on the structure of their models to infer trade costs, I take the alternative approach of employing estimators that control for all possible patterns of unobserved micro-level heterogeneity using product-level data. Another related paper is Hallak (2010), which finds that

[4] See, for example, the review by Stoker (1993).

tests of the Linder (1961) hypothesis based on aggregate data are biased because income per capita is correlated with patterns of comparative advantage.

My paper is also related to a fast-growing strand of the international trade literature featuring models with rich sectoral heterogeneity. Costinot and Rodríguez-Clare (2014) and Kehoe et al. (2017) review recent advances in this area. Because my empirical framework is consistent with the structure of the vast majority of these models, the proposed estimators can be used to parameterize them. A case in point is French (2016), which uses the pooled PPML estimator, developed in an earlier version of this paper, to calibrate a many-product Eaton and Kortum (2002) model and finds that the welfare gains from trade depend on the patterns of PLCA in the data.

In the next section, I specify the empirical model and characterize the bias in aggregate estimates in the presence of heterogeneous fixed effects. Section 3 introduces the application to trade flows and refines the bias characterization in this setting. Section 4 develops a set of product-level structural gravity estimators and discusses practical estimation issues. Sections 5 and 6 present the empirical results and extension to panel data. Section 7 concludes.

2 Empirical Model

Given observational units indexed by (n, i), where $n = 1, .., N$ and $i = 1, ..., I$, suppose that a scalar outcome variable, X_{ni}^{jk}, is observed across micro-level categories indexed by $k = 1, .., K^j$, where $j = 1, .., J$ indexes categories at a higher level of aggregation. Further suppose that interest lies in the effect on the aggregate variable, $X_{ni}^{j} = \sum_{k=1}^{K^j} X_{ni}^{jk}$, of a set of explanatory variables, $\boldsymbol{z}_{ni}^{j}$.[5] I consider an empirical model in which the conditional mean of X_{ni}^{jk} is given by a nonlinear function of $\boldsymbol{z}_{ni}^{j}$ together with unobserved multiplicative effects that are heterogeneous across k:

Assumption 1.

$$\mathrm{E}\left[X_{ni}^{jk}\middle|\boldsymbol{Z}^{j}, \boldsymbol{\Gamma}^{jk}\right] = \omega_{ni}^{jk} = \phi_{n}^{jk}\psi_{i}^{jk} f(\boldsymbol{z}_{ni}^{j}, \boldsymbol{\beta}^{jk}), \tag{1}$$

where ϕ_{n}^{jk} and ψ_{i}^{jk} are unobserved effects, $\boldsymbol{z}_{ni}^{j}$ is an $L \times 1$ vector of explanatory variables, $\boldsymbol{\beta}^{jk}$ is an $L \times 1$ vector of (potentially k-specific) common parameters, $\boldsymbol{Z}^{j} = (\boldsymbol{z}_{11}^{j}, ..., \boldsymbol{z}_{NI}^{j})'$, and $\boldsymbol{\Gamma}^{jk} = (\boldsymbol{\beta}^{jk\prime}, \boldsymbol{\phi}_{1}^{jk}, ..., \boldsymbol{\phi}_{N}^{jk}, \boldsymbol{\psi}_{1}^{jk}, ..., \boldsymbol{\psi}_{I}^{jk})'$.[6]

It is also useful to define the disturbance term $\eta_{ni}^{jk} = X_{ni}^{jk}/\omega_{ni}^{jk}$. The model is semiparametric in that the joint distribution of ϕ_{n}^{jk}, ψ_{i}^{jk}, and η_{ni}^{jk} is not specified. Because my focus

[5] The analysis that follows holds analogously for the weighted average outcome $\bar{X}_{ni}^{j} = \sum_{k=1}^{K^j} w^k X_{ni}^{jk}$, as long as the weights are exogenously given.

[6] Assuming strict exogeneity of the explanatory variables allows me to consider a broader class of estimators. However, this assumption can be relaxed in many settings.

is on situations in which the researcher may be tempted to estimate an aggregate version of (1), the explanatory variables are assumed to be constant across k. All of the estimation methods that I employ can accommodate explanatory variables that vary by k, but I leave this case for future work for brevity and notational convenience. In what follows, every object is allowed to vary across j, so to avoid excessive notation I omit the j superscript wherever there is no ambiguity.

I apply this model to bilateral trade data, where n and i are countries, k is a product category, and j is the sector of interest. However, this setup is applicable in numerous settings where unobserved heterogenous effects interact. For example, one may wish to identify (a) the effect of R&D expenditure using a panel of firms (n indexing firms and i indexing time), where unobservable firm-product effects may interact with unobservable product-time spillovers to influence patenting activity or (b) the effect of class size on student outcomes (n indexing student and i indexing semester), where academic achievement depends on the interaction between student proclivity in a subject, ϕ_n^k, and peer effects that vary by subject and semester, ψ_i^k.

The most common approach to controlling for the unobserved effects in specifications like (1) is to treat them as parameters to be estimated.[7] While it is common to estimate a log-linear version of (1) by OLS, Assumption 1 does not imply consistency of OLS, as Santos Silva and Tenreyro (2006) show. Instead, the common parameters and unobserved effects can be consistently estimated by a fixed-effects (FE) pseudo-maximum likelihood (PML) estimator based on (1) in its multiplicative form.[8] Another approach is to structurally control for the unobserved effects. This is exemplified by structural gravity estimators of international trade models, introduced by Anderson and van Wincoop (2003), which impose the model's market-clearing conditions to express ϕ_n and ψ_i as functions of data and parameters.

To consider both approaches in a common framework, I define the following class of method-of-moments estimators. Given scalar residual functions, $r(X_{ni}^k, \hat{\omega}_{ni}^k)$ and $\tilde{r}(X_{ni}^k, \hat{\omega}_{ni}^k)$, where $\hat{\omega}_{ni}^k = \hat{\phi}_n^k \hat{\psi}_i^k f(\boldsymbol{z}_{ni}; \hat{\boldsymbol{\beta}}^k)$, that satisfy $\mathrm{E}[r(X_{ni}^k, \omega_{ni}^k)|\boldsymbol{Z}, \boldsymbol{\Gamma}^k] = 0$ and $\mathrm{E}[\tilde{r}(X_{ni}^k, \omega_{ni}^k)|\boldsymbol{Z}, \boldsymbol{\Gamma}^k] =$

[7] See Head and Mayer (2014) and Fernández-Val and Weidner (2016) for discussions of the international trade and panel estimation literatures.

[8] Fernández-Val and Weidner (2016) prove the consistency of maximum likelihood estimators of nonlinear models with two-way fixed effects under only a conditional mean assumption.

0 under Assumption 1, these estimators solve empirical moment conditions of the form

$$\sum_{n=1}^{N}\sum_{i=1}^{I} r(X_{ni}^k, \hat{\omega}_{ni}^k)\boldsymbol{Z}_{ni} = 0 \tag{2}$$

$$\sum_{i=1}^{I} \tilde{r}(X_{ni}^k, \hat{\omega}_{ni}^k) = 0, \quad \forall n, k, \tag{3}$$

$$\sum_{n=1}^{N} \tilde{r}(X_{ni}^k, \hat{\omega}_{ni}^k) = 0, \quad \forall i, k. \tag{4}$$

where $\boldsymbol{Z}_{ni} = \boldsymbol{h}_{ni}(\boldsymbol{Z}, \hat{\boldsymbol{\Gamma}}^k)$ is an $L \times 1$ vector that depends on the explanatory variables and parameters.[9]

Fally (2015) shows that FE OLS and FE PML estimators imply "adding-up" constraints of the form of (3) and (4). For these estimators, $r(\cdot) = \tilde{r}(\cdot)$, and the functional form depends on the chosen objective function. Structural gravity estimators are equivalent to specifying $\tilde{r}(X_{ni}^k, \hat{\omega}_{ni}^k) = X_{ni}^k - \hat{\omega}_{ni}^k$. Thus, they solve concentrated empirical moment conditions of the form

$$\sum_{n=1}^{N}\sum_{i=1}^{I} r(X_{ni}^k, \hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}^k))\boldsymbol{Z}_{ni} = 0, \quad \forall k, \tag{5}$$

where

$$\hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}^k) = \frac{X_n^k}{\Phi_n^k}\frac{Y_i^k}{\Psi_i^k} f(\boldsymbol{z}_{ni}; \hat{\boldsymbol{\beta}}^k), \tag{6}$$

and where $X_n^k \equiv \sum_i X_{ni}^k$, $Y_i^k \equiv \sum_n X_{ni}^k$, and Φ_n^k and Ψ_i^k solve the system of equations

$$\Phi_n^k = \sum_i \frac{Y_i^k}{\Psi_i^k} f(\boldsymbol{z}_{ni}; \hat{\boldsymbol{\beta}}^k) \quad \text{and} \quad \Psi_i^k = \sum_n \frac{X_n^k}{\Phi_n^k} f(\boldsymbol{z}_{ni}; \hat{\boldsymbol{\beta}}^k). \tag{7}$$

Poisson PML (PPML) is a special case for which the FE and structural estimators coincide because the Poisson likelihood function implies the same form for $\tilde{r}(\cdot)$ as is imposed by the structural gravity estimators.[10]

Equation (6) is a disaggregated structural gravity equation in its most common form, with Φ_n^k and Ψ_i^k being the "multilateral resistance" (MR) terms defined by Anderson and van

[9] In many settings $\boldsymbol{Z}_{ni} = \boldsymbol{h}(\boldsymbol{z}_{ni}, \hat{\boldsymbol{\beta}}^k, \hat{\boldsymbol{\phi}}_n^k, \hat{\boldsymbol{\psi}}_i^k)$ or even $\boldsymbol{Z}_{ni} = \boldsymbol{z}_{ni}$. The more general specification includes structural m-estimators whose first-order conditions include functions of explanatory variables and unobserved effects across observations. Note that I do not claim that all such estimators are consistent for $\boldsymbol{\beta}$. I define this class to characterize the bias in estimates of $\boldsymbol{\beta}$ based on aggregated data, given a consistent micro-level estimator in the class.

[10] Fally (2015) and Arvis and Shepherd (2013) provide formal proofs of this equivalence.

Wincoop (2003).[11] Treating structural gravity estimators as method-of-moments estimators based on Assumption 1 is useful for two reasons. First, it makes clear that structural gravity estimators are applicable outside of an international trade context. Second, structural gravity models are typically derived under the assumption that $\eta_{ni}^k = 1$ and an error term appended to (6) ex post, which is inconsistent with the fact that shocks to X_{ni}^k are also shocks to X_n^k and Y_i^k. My derivation justifies the specification of the MR terms as functions of observed, rather than expected, values of X_n^k and Y_i^k.

If not for heterogeneity in the unobserved effects and common parameters, (1) and (6) would hold at the aggregate level. Thus, a common practice is to estimate an aggregate version of (1). Such an exercise implicitly assumes homogeneity of the unobserved effects and either likewise assumes homogeneity of the common parameters or seeks to estimate an index of $\boldsymbol{\beta}$ or an average partial effect of an explanatory variable. I use the following proposition to characterize the potential bias in estimates based on such practice.

Proposition 1. *Let $\bar{\alpha}$ denote a sector-level index of a given set of product-level parameters, $\boldsymbol{\alpha} = \{\alpha^k : k = 1, .., K\}$, which is homogeneous of degree 1. Assumption 1 implies that* $\mathrm{E}[r(X_{ni}, \bar{\omega}_{ni} T_{ni}) | \boldsymbol{Z}, \bar{\boldsymbol{\Gamma}}, T_{ni}] = 0$ *and* $\mathrm{E}[\tilde{r}(X_{ni}, \bar{\omega}_{ni} T_{ni}) | \boldsymbol{Z}, \bar{\boldsymbol{\Gamma}}, T_{ni}] = 0$*, where* $\bar{\omega}_{ni} = \bar{\phi}_n \bar{\psi}_i f(\boldsymbol{z}_{ni}; \bar{\boldsymbol{\beta}})$*,* $\bar{\boldsymbol{\Gamma}} = (\bar{\boldsymbol{\beta}}', \bar{\boldsymbol{\phi}}_1, ..., \bar{\boldsymbol{\phi}}_N, \bar{\boldsymbol{\psi}}_1, ..., \bar{\boldsymbol{\psi}}_I)'$*, and T_{ni} is given by*

$$T_{ni} = \sum_k \frac{\phi_n^k}{\bar{\phi}_n} \frac{\psi_i^k}{\bar{\psi}_i} \frac{f(\boldsymbol{z}_{ni}; \boldsymbol{\beta}^k)}{f(\boldsymbol{z}_{ni}; \bar{\boldsymbol{\beta}})}. \tag{8}$$

Proofs of all propositions are provided in Appendix B. The T_{ni} term summarizes the effect of the interaction of the unobserved effects on X_{ni}. In the context of bilateral trade flows, T_{ni} embodies the effect of countries' product-level comparative advantage on sector-level trade flows. The magnitude of T_{ni} is also affected by heterogeneity in $\boldsymbol{\beta}$ as well as the choice of index $\bar{\boldsymbol{\beta}}$. Regarding the former, T_{ni} will tend to be greater (smaller) if $f(\boldsymbol{z}_{ni}; \boldsymbol{\beta}^k)$ tends to be greater (smaller) for observations where the interaction effect is greater.

In general, Assumption 1 implies neither $\mathrm{E}[r(X_{ni}^k, \bar{\omega}_{ni}) | \boldsymbol{Z}, \bar{\boldsymbol{\Gamma}}] = 0$ nor $\mathrm{E}[\tilde{r}(X_{ni}^k, \bar{\omega}_{ni}) | \boldsymbol{Z}, \bar{\boldsymbol{\Gamma}}] = 0$. Instead, Proposition 1 implies aggregate analogues of the method-of-moments estimators

[11] Some authors refer to FE estimators as "structural" because they control for model-implied fixed effects. To be clear, I reserve the term "structural" to refer to estimators that solve for the MR terms using (7).

given by (2)-(4), which take the form

$$\sum_{n=1}^{N}\sum_{i=1}^{I} r(X_{ni}, \hat{\bar{\omega}}_{ni} T_{ni}) \boldsymbol{Z}_{ni} = 0$$
$$\sum_{i=1}^{I} \tilde{r}(X_{ni}, \hat{\bar{\omega}}_{ni} T_{ni}) = 0, \quad \forall n,$$
$$\sum_{n=1}^{N} \tilde{r}(X_{ni}, \hat{\bar{\omega}}_{ni} T_{ni}) = 0, \quad \forall i.$$

where $\hat{\bar{\omega}}_{ni} = \hat{\bar{\phi}}_n \hat{\bar{\psi}}_i f(\boldsymbol{z}_{ni}; \hat{\bar{\boldsymbol{\beta}}})$. Typical aggregate estimators omit T_{ni}, which is unobservable, meaning that they are generally misspecified, yielding biased estimates of $\bar{\boldsymbol{\beta}}$. Further, T_{ni} cannot be estimated using aggregate data because it varies by both n and i.

The severity and direction of the bias depends on the relationship between T_{ni} and $\boldsymbol{z}_{ni}$ in the data and the form of $f(\cdot)$. Given that the empirical model imposes no restrictions on the distribution of the micro-level unobserved effects or their relationship with $\boldsymbol{z}_{ni}$, T_{ni} and $\boldsymbol{z}_{ni}$ will generally be related and estimates of $\bar{\boldsymbol{\beta}}$ biased. To illustrate, suppose that $f(\cdot)$ takes the commonly assumed exponential form:

Assumption 2. $f(\boldsymbol{z}_{ni}; \boldsymbol{\beta}^k) = e^{\boldsymbol{z}'_{ni}\boldsymbol{\beta}^k}$.

For OLS, the bias can be expressed using the well-known omitted variable bias formula:

$$\text{plim}_{N\to\infty}\, \hat{\bar{\boldsymbol{\beta}}}_{\text{OLS}} = \bar{\boldsymbol{\beta}} + \text{plim}_{N\to\infty}\, (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}' \ln(\boldsymbol{T}), \tag{9}$$

where $\ln(\boldsymbol{T}) \equiv (\ln(T_{11}), ..., \ln(T_{N1}), ..., \ln(T_{NN}))'$.[12] Therefore, if a regressor is positively correlated with T_{ni}, omitting T_{ni} will tend to bias its coefficient upward. Nonlinear estimators do not generally yield analogous closed-form expressions. However, Neuhaus and Jewell (1993) derive an approximation for PML estimators which takes a very similar form to (9), in which the final term is replaced by the expected change in $\ln(T_{ni})$ conditional upon a change in $\boldsymbol{z}_{ni}$. Thus, the intuition regarding the direction of the bias is the same.

3 Application to Trade Flows

To assess the severity of the bias in estimates that ignore unobserved micro-level effects and to evaluate the relative merits of different approaches to controlling for micro-level

[12] To highlight the bias due to omitting T_{ni}, this expression implicitly assumes that the aggregate errors are homoskedastic so that there is no bias due to the log transformation of the data.

heterogeneity, I consider the estimation of the determinants of bilateral international trade flows. A wide range of models imply that trade flows are characterized by a specification which takes the form of (1).[13] For concreteness, consider a model economy consisting of N countries, each of which contains buyers who demand goods from $j = 1, ..., J$ sectors. Each sector is made up of a finite number of product categories, $k = 1, ..., K^j$, and each category contains a continuum of product varieties $u \in \boldsymbol{U}^{jk}$, of which a weak subset $\boldsymbol{U}_n^{jk} \subseteq \boldsymbol{U}^{jk}$ are available in n. Given this basic setup, a wide variety of assumptions regarding buyers' objectives, technologies, and market structure imply that trade flows from i to n of product k are consistent with (1).[14]

In this setting $f(\boldsymbol{z}_{ni}^j; \boldsymbol{\beta}^{jk})$ represents the effect of trade barriers on trade flows of product k, and η_{ni}^{jk} represents exogenous shocks to observed product-level trade flows that do not affect the conditional expectation of X_{ni}^{jk}.[15] This specification allows for substantial heterogeneity in trade costs across products from variation in $\boldsymbol{\beta}^{jk}$, which subsumes heterogeneous effects of regressors on trade costs, heterogenous trade cost elasticities, and heterogenous country-specific border costs. This specification also allows for non-zero domestic trade costs and asymmetric international trade costs, both of which have been found to be important dimensions of variation in the literature.

In the context of trade flows, T_{ni} has a clear interpretation as the effect of countries' product-level comparative advantage (PLCA) on sector-level trade flows. To see this, note that $\phi_i^k/\bar{\phi}_i$ is a measure of i's ability to produce k, relative to its own ability to produce a bundle of all products, while $\psi_n^k/\bar{\psi}_n$ measures the relative strength of demand for k in n. The latter can include both factors outside the model that determine relative expenditure on k, X_n^k/X_n, and factors that endogenously determine Φ_n^k through (7). Their interaction embodies the classic notion of comparative advantage; i will export more of k to n if i is better able to supply k, relative to both i's ability to supply other products and other countries' ability to supply k to n. If this is the case for many products, then i will export relatively more to n overall. In addition, parameter heterogeneity will amplify (weaken) the effect of PLCA if trade costs tend to be lower (higher) for i's comparative advantage products in n.

While Assumption 1 is sufficient for identification of $\boldsymbol{\beta}$, to gain further insight into the bias in sector-level estimates, it is useful to impose slightly more structure in the form of the

[13]See Anderson (2011) and Head and Mayer (2014) for discussions of this class of theoretical models.

[14]These include those delineated by Arkolakis et al. (2012) as well as Arkolakis et al. (2015) and Melitz and Ottaviano (2008).

[15]Such shocks may include measurement error or any phenomenon that breaks the link between observed shipments and the forces that determine goods and factor prices in general equilibrium, such as lumpiness in trade flows, shipping lags, inventories, or price stickiness. Egger and Nigai (2015) show that it is problematic to relegate unobserved trade costs to the error term, in contrast to the argument of Anderson and van Wincoop (2003) that resulting biases are likely to be small.

following assumption:

Assumption 1′. *Observed trade flows from i to n of product k are given by*

$$X_{ni}^{k} = \frac{c_{i}^{k} f(\boldsymbol{z}_{ni}; \boldsymbol{\beta}^{k})}{\Phi_{n}^{k}} X_{n}^{k} \eta_{ni}^{k},$$

where $\Phi_n^k = \sum_i c_i^k f(\boldsymbol{z}_{ni}; \boldsymbol{\beta}^k)$.

Most models that are consistent with Assumption 1 are also consistent with Assumption 1′.[16] The term c_i^k includes all exogenous and endogenous factors that affect country i's overall ability to supply product k, such as productivity, factor prices, and the mass of firms that produce varieties of k. The term Φ_n^k is an index of all exporters' abilities to deliver product k to destination n, which also serves to ensure that product-level trade flows sum across exporters to product-level expenditure. Assumption 1′ does not necessarily require further restrictions on the form of buyers' objective functions.[17] In particular, I do not impose any restrictions on factor markets or the demand system that determines buyers' allocation of expenditure across sectors or products. The latter implies that I also impose no restrictions on the input-output structure of the economy.

Assumption 1′ implies that T_{ni} is a function of trade costs. Thus, significant PLCA will cause T_{ni} and $\boldsymbol{z}_{ni}$ to be strongly related and the bias in sector-level estimates severe. This can be seen in the following result:

Proposition 2. *Given Assumptions 1′ and 2, changes in $\ln(T_{ni})$ associated with changes in $\boldsymbol{Z}$ are given by*

$$\begin{aligned} d\ln(T_{ni}) = \sum_k \frac{\hat{X}_{ni}^k}{\hat{X}_{ni}} \Bigg[& d\boldsymbol{z}'_{ni}(\boldsymbol{\beta}^k - \bar{\boldsymbol{\beta}}) - \sum_m d\boldsymbol{z}'_{nm}\bar{\boldsymbol{\beta}} \left(\frac{\hat{X}_{nm}^k}{X_n^k} - \frac{\hat{X}_{nm}}{X_n} \right) \\ & - \sum_m d\boldsymbol{z}'_{nm} \left(\boldsymbol{\beta}^k - \bar{\boldsymbol{\beta}}\right) \frac{\hat{X}_{nm}^k}{X_n^k} + d\ln(X_n^k) \Bigg], \end{aligned} \tag{10}$$

where $\hat{X}_{ni} = \sum_k \hat{X}_{ni}^k$, holding constant all values of X_n and c_i^k and normalizing T_{ni} such that

[16] See Head and Mayer (2014), who refer to the aggregate analogue of Assumption 1 as "general gravity" and that of Assumption 1′ as "structural gravity". However, note that the "general gravity" assumption is sufficient for estimation using a structural gravity estimator.

[17] Different micro structures require different functional form restrictions. The Armington model (Anderson and van Wincoop, 2003) requires that buyers have CES preferences. French (2015) and Arkolakis et al. (2015) provide models with perfect and monopolistic competition that relax this assumption. Assumption 1′ does require that buyers maximize an objective function that is separable across varieties – i.e., expenditure on (j, k, u) is given by $x_n^{jk}(u) = f_n^{jk}(u, \mathbf{p}_n^{jk}, X_n^{jk})$, where $\mathbf{p}_n^{jk}$ is the set of prices of varieties of (j, k). This implies the "trade separability" property assumed by Anderson and van Wincoop (2004).

$\sum_i \hat{X}_{ni} d\ln T_{ni} = 0$.[18]

The first term in brackets represents the direct effect of parameter heterogeneity on T_{ni}. The second term represents the effect of PLCA. Note that the values in parentheses can be thought of as measures of revealed comparative advantage for country m in market n.[19] Thus, an increase in trade barriers between m and n – i.e., a change in explanatory variable l such that $\sum_k X_n^k \beta^{k(l)} dz_{nm}^{(l)} < 0$ – will disproportionately shift n's sector-level expenditure toward i if m's PLCA is positively correlated with n's product-level imports from i, $\hat{X}_{ni}^k$. The third term represents the effect of the interaction of parameter heterogeneity with PLCA.

This expression simplifies considerably if we assume that $\boldsymbol{\beta}$ is homogeneous across products and consider only the partial effect of $\boldsymbol{z}_{ni}$ on T_{ni}:

$$\frac{\partial \ln(T_{ni})}{\partial \boldsymbol{z}_{ni}} = -\bar{\boldsymbol{\beta}} \sum_k \frac{X_n^k}{X_{ni}} \left(\frac{X_{ni}^k}{X_n^k} - \frac{X_{ni}}{X_n} \right)^2 + \sum_k \frac{X_{ni}^k}{X_{ni}} \frac{\partial \ln(X_n^k)}{\partial \boldsymbol{z}_{ni}}. \tag{11}$$

Consider that $\bar{\boldsymbol{\beta}}$ is the direct partial effect of $\boldsymbol{z}_{ni}$ on $\ln(X_{ni})$. Equation (11) demonstrates that this effect is weakened in proportion to the strength of i's PLCA in n, measured by the weighted variance of its product-level revealed comparative advantage. Intuitively, stronger comparative advantage ameliorates the effects of trade barriers.[20]

What does this imply for the bias in sector-level estimates of $\bar{\boldsymbol{\beta}}$? Clearly, the bias will be more severe the stronger are countries' patterns of PLCA. Further, if the bilateral effect of trade costs on T_{ni}, given by (11), dominates the multilateral effects, given by (10), or if the latter are uncorrelated with $\boldsymbol{z}_{ni}$, then T_{ni} will be negatively correlated with $\boldsymbol{z}_{ni}'\bar{\boldsymbol{\beta}}$. In this case, omitting T_{ni} will bias estimates of $\boldsymbol{\beta}$ toward zero, particularly for variables whose effects are large in magnitude.[21]

Estimates of $\bar{\boldsymbol{\beta}}$ may be biased further when there is heterogeneity in parameters across products. Anderson and van Wincoop (2004) consider this "aggregation bias" and make clear that the value of $\bar{\boldsymbol{\beta}}$ estimated from aggregated data may not be an ideal index of $\boldsymbol{\beta}$.[22] While (11) shows that the bias due to PLCA is distinct from aggregation bias due to parameter heterogeneity, the latter can also be viewed as a consequence of omitting T_{ni}, which captures the aggregate effects of deviations of $\boldsymbol{\beta}^k$ from $\bar{\boldsymbol{\beta}}$. Parameter heterogeneity may amplify or

[18] I suppress the argument $\hat{\boldsymbol{\beta}}^k$ when $\hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}^k)$ is evaluated at the true parameter value $\boldsymbol{\beta}^k$. The normalization of T_{ni} is isomorphic to a choice of index $\bar{\Phi}_n$

[19] In fact, this is the Bilateral Additive Index, defined by French (2017), and equation (10) is closely related to the Trade Elasticity Index.

[20] The sign of $\partial \ln(T_{ni})/\partial z_{ni}^{(l)}$ is always the opposite of the sign of $\bar{\beta}^{(l)}$ under the natural assumption that the elasticity of substitution across sources of a product is greater than the elasticity of substitution across products, which ensures that the first expression in (11) is greater in absolute value than the second.

[21] I.e., those for which the size of $\bar{\beta}^{(l)} z_{nm}^{(l)} / \boldsymbol{z}_{ni}' \bar{\boldsymbol{\beta}}$ is significant.

[22] See also Hillberry (2002) and Anderson and Neary (2003) for systematic treatments of aggregation bias.

offset comparative advantage bias, depending on the correlation between $\boldsymbol{\beta}^k$ and PLCA, as (10) shows. Which of these biases is of greater concern is an empirical question, which I explore in Section 5. However, I show in Section 4.1 that aggregation bias is not a concern for Poisson PML.

Further note that any bias that exists for an estimator at the product level may interact with PLCA to further bias sector-level estimates. Assumption 1 ensures that $\mathrm{E}\left[\sum_k \phi_n^k \psi_i^k \eta_{ni}^k\right] = \sum_k \phi_n^k \psi_i^k$ and thus that η_{ni}^k does not appear in (8). However, if an estimator is biased under Assumption 1, as Santos Silva and Tenreyro (2006) show is the case for log-linear OLS, or is biased in finite samples, then T_{ni} will depend not only on patterns of PLCA but also the interaction of these patterns with η_{ni}^k.

Some recent gravity estimations have taken advantage of the panel dimension of trade data, estimating a trade cost function of the form

$$f(\boldsymbol{z}_{ni,t}; \boldsymbol{\beta}) = \xi_{ni} e^{\boldsymbol{z}'_{ni,t}\boldsymbol{\beta}},$$

which allows for the inclusion of country-pair fixed effects to control for all time-invariant trade barriers.[23] Sector-level panel estimation also suffers from bias in the presence of PLCA. Because T_{ni} is a general equilibrium object that is a function of all trade costs and patterns of comparative advantage, both of which are time-varying, T_{ni} is also time-varying and cannot be controlled for using static fixed effects.[24] For simplicity and because it is still the most common form of gravity estimation, I focus on cross-sectional estimation in the baseline analysis. However, it is conceptually straightforward to extend the results to panel data, and I return to panel estimation in Section 6.

4 Gravity Estimation with PLCA

My model shows that sector-level gravity estimation that ignores PLCA is generally biased. To assess the severity of this bias in practice, I propose and implement a set of estimators that are robust to any form of PLCA. These estimators make use of product-level trade data, which are now widely available for most countries. They also easily accommodate two common complications for product-level estimation: the overwhelming number of product-level coefficients estimated and the lack of disaggregated data on domestic trade flows.

[23] Baier and Bergstrand (2007) is a seminal article that employs this strategy, followed by many others.

[24] Similarly, the two-stage estimation approach proposed by Egger and Nigai (2015) also suffers from the same form of bias. Their first stage – decomposing trade flows into country-specific and bilateral components – yields the product of the fitted values $e^{\boldsymbol{z}'_{ni}\hat{\boldsymbol{\beta}}}\hat{T}_{ni}$. Without controlling for T_{ni}, estimates of $\boldsymbol{\beta}$ based on these values will be biased in the same way as traditional sector-level gravity estimations.

4.1 Estimation Methods

There are several options for estimating $\boldsymbol{\beta}$ using an estimator in the class defined by (2)-(4). The key choices lie along four dimensions: (a) modelling the error term, (b) controlling for the unobserved effects, (c) whether to pool across products, and (d) the level of aggregation at which to define a sector.

4.1.1 Modelling the Error Term

The optimal moment conditions depend on assumptions about the error term, η_{ni}^k. In the trade literature, the first generation of theory-consistent gravity estimations used nonlinear least squares or OLS based on the logged form of a gravity equation.[25] Since Santos Silva and Tenreyro (2006) demonstrated that the log transformation leads to inconsistent estimates in the presence of heteroskedasticity, PML estimators have become more widely used.

PML estimators maximize a likelihood function based on a linear-exponential probability distribution, such as the Poisson, gamma, and Gaussian distributions. Gourieroux et al. (1984) show that each of these is consistent as long as the conditional mean is correctly specified. They differ only in how they weight observations based on the assumed form of heteroskedasticity. Specifically, gamma, Poisson, and Gaussian (NLS in levels) PML imply the following specifications of $r(\cdot)$:[26]

Distribution	$r(X_{ni}^k, \hat{\omega}_{ni}^k)$
Gamma PML (GPML)	$(X_{ni}^k - \hat{\omega}_{ni}^k)(\hat{\omega}_{ni}^k)^{-1}$
Poisson PML (PPML)	$(X_{ni}^k - \hat{\omega}_{ni}^k)$
Gaussian PML (NLS)	$(X_{ni}^k - \hat{\omega}_{ni}^k)\hat{\omega}_{ni}^k$

Log-linear least squares implies $r(X_{ni}^k, \hat{\omega}_{ni}^k) = \ln(X_{ni}^k) - \ln(\hat{\omega}_{ni}^k)$. While Santos Silva and Tenreyro (2006) prefer PPML due to its performance in Monte Carlo experiments, I also employ GPML, NLS, and log-LS for robustness and due the popularity of the latter in the literature.[27]

[25] Anderson and van Wincoop (2003) and Eaton and Kortum (2002) are seminal examples, respectively.

[26] For the FE estimators, $\boldsymbol{Z}_{ni} = \partial \ln f(\boldsymbol{z}_{ni}, \boldsymbol{\beta}^k)/\partial \boldsymbol{\beta}^k$. For the structural estimators, $\boldsymbol{Z}_{ni} = \partial \ln \hat{X}(\boldsymbol{\beta}^k)/\partial \boldsymbol{\beta}^k = \partial \ln f(\boldsymbol{z}_{ni}, \boldsymbol{\beta}^k)/\partial \boldsymbol{\beta}^k - \partial \ln(\Phi_n^k)/\partial \boldsymbol{\beta}^k - \partial \ln(\Psi_i^k)/\partial \boldsymbol{\beta}^k$. The latter terms drop out of the moment conditions for structural PPML because (7) implies that $\sum_i (X_{ni}^k - \hat{X}_{ni}^k)\partial \ln(\Phi_n^k)/\partial \boldsymbol{\beta}^k = 0$ and $\sum_n (X_{ni}^k - \hat{X}_{ni}^k)\partial \ln(\Psi_i^k)/\partial \boldsymbol{\beta}^k = 0$, which constitutes an alternative proof of the main results of Fally (2015) and Arvis and Shepherd (2013).

[27] Another popular PML estimator is the negative binomial, which nests Poisson PML as a special case. However, Head and Mayer (2014) summarize several compelling arguments against using this estimator, including that estimates depend on the unit of measurement of the dependent variable.

4.1.2 Controlling for Fixed Effects

Both FE and structural estimators have been frequently employed in aggregate estimations. In the class of log-LS estimators, the FE approach corresponds to log-linear OLS (under Assumption 2), as in Eaton and Kortum (2002) and many subsequent papers, while the structural approach corresponds to the NLS (in logs) estimator of Anderson and van Wincoop (2003). PPML, which is both a FE and structural gravity estimator, has recently risen in popularity based on the findings of Santos Silva and Tenreyro (2006).

The relative merits of FE and structural gravity estimators depend on several practical considerations. It is tempting to favor FE estimators based on the argument that they impose less structure and thus are more robust to model misspecification. However, as Anderson and van Wincoop (2003) point out, the structural approach actually imposes no more structure on the estimation. While the structural approach imposes (7), Fally (2015) shows that the FE estimators implicitly impose similar adding-up constraints, meaning that the only difference between the two approaches is in the functional form of $\tilde{r}(\cdot)$.

Fernández-Val and Weidner (2016) show that fixed effect (FE) estimators have asymptotic bias on the order of $1/N$, with two notable exceptions. If (1) is estimated by log-linear OLS, then it is equivalent to the classic two-way within estimator, which is unbiased. However, as Santos Silva and Tenreyro (2006) show, Assumption 1 does not imply consistency of log-linear OLS. Taking logs also drops zero-valued observations, leading to potential sample selection bias.[28] The other exception is the FE PPML estimator, which Fernández-Val and Weidner (2016) show has no asymptotic bias. The structural gravity estimators do not suffer from asymptotic bias under the assumption that (7) holds non-stochastically because they solve the concentrated moment conditions (5), which do not depend on the unobserved effects. The equivalence between structural and FE PPML implies that the former is also generally unbiased under Assumption 1 alone.[29]

An additional practical disadvantage of the FE estimators is that estimation may be computationally intensive for large datasets because the total number of parameters to estimate grows with $(N+I)\times K$.[30] By contrast, the structural gravity estimators are straightforward

[28] Eaton and Tamura (1994), Hallak (2006), and Helpman et al. (2008) propose estimators that deal with the issue of sample selection bias while maintaining the log-linear regression approach. However, in Monte Carlo experiments, Santos Silva and Tenreyro (2006) find that such methods perform comparatively poorly.

[29] It is plausible that the other structural estimators are also generally unbiased because they control for the unobserved effects by imposing the same empirical moment conditions as FE PPML, but I leave this question for future work.

[30] In my empirical application, the dataset contains trade flows among 130 countries in 4,208 product categories. FE estimation would require 1,188,864 dummy variables, meaning that forming the matrix of independent variables would require nearly 80 terabytes of computer memory.

to compute because (7) has a unique solution, and efficient algorithms exist to find it.[31] The FE OLS and PPML estimators are also exceptions to this consideration. Efficient computational methods exist to compute the OLS within estimator.[32] The FE PPML estimator can be computed in the same way as the structural estimator.

Thus, my preferred estimators are the disaggregate structural gravity estimators. For comparison with the aggregate FE estimators common in the literature, I also estimate FE specifications for all aggregate estimators. However, due to computational difficulty and their asymptotic bias problem, I do not estimate disaggregate FE GPML or FE NLS specifications.

4.1.3 To Pool or Not to Pool?

Given an objective function and a strategy for controlling for the FEs, I consider two methods for estimating $\boldsymbol{\beta}$ using product-level trade data: product-by-product and pooled product-level estimation. Sector-level gravity estimation implicity imposes the restriction that $\boldsymbol{\beta}^k = \bar{\boldsymbol{\beta}}$, for all k. If we maintain this homogeneity assumption, then it is efficient to apply the chosen product-level estimator to the full set of product-level data, pooled across products. Product-by-product estimation relaxes the homogeneity assumption but may be less efficient.

In addition to controlling for PLCA, both estimators control for aggregation bias due to parameter heterogeneity. Product-by-product estimators do so by directly estimating heterogeneous coefficients. Pooled estimators estimate an average, or index, of the heterogeneous coefficients across products. In this regard, product-by-product estimation is the natural extension of the industry-by-industry estimations of, e.g., Anderson and Yotov (2010a,b) and Bergstrand et al. (2015), which sought to address this issue. Pooled estimation is analogous to the estimation of "average treatment effects" – common in the labor econometrics literature and adapted to aggregate gravity estimation by Baier and Bergstrand (2007) – across both countries and products.

Though one may be inclined to prefer the flexibility of product-by-product estimators, it is important to consider the tradeoffs between pooled and product-by-product estimation. In addition to efficiency concerns, the high degree of flexibility of product-by-product estimators may cause them to overfit the sample data, reducing their predictive power. For example, Baltagi (2008) finds that homogeneous estimators consistently outperform heterogeneous estimators across a wide range of panel estimations. I evaluate this in the empirical application (see Section 5.3).

[31] Poissonnier (2019) shows that iterating on (7) from any positive starting values converges to the unique solution, up to a scalar normalization.

[32] See, for example, the algorithms of Guimarães and Portugal (2010) and Gaure (2013).

4.1.4 Choice of Sector

The final choice is the level of aggregation at which to define a sector. Though the vast majority of gravity estimations treat countries as one-sector economies, recent papers have considered multiple sectors, typically defined at roughly the 2-digit ISIC level, while still largely focusing on aggregate outcomes.[33] In keeping with the literature, I consider single-sector specifications that treat manufacturing as one sector and multi-sector specifications that define sectors as 2-digit ISIC industries. Multi-sector pooled estimation can be seen as an intermediate step between single-sector pooled estimation and product-by-product estimation, which implicitly treats each product as a different sector.

4.2 An Ideal Coefficient Index

One complication of product-level estimation is that, if we allow for heterogeneity across products, the large number of coefficients makes reporting and inference based on the estimates impractical. A solution to both problems is using a summary index of the estimates. Anderson and Neary (2003) and Anderson and van Wincoop (2004) discuss the reporting issue and construction of an ideal trade cost index. Anderson (2008) discusses the multiple inference problem, which is well-understood in theory but only recently garnering significant attention in applied fields, and advocates for inference using summary indexes to reduce the number of hypotheses tested.

Simple or weighted average trade costs have often been reported in the gravity literature but do not accurately reflect the overall level of trade restrictiveness. Anderson and Neary (2003) propose a summary trade cost index defined as the uniform trade cost that preserves the level of aggregate trade flows. While useful for summarizing overall trade barriers, their index does not directly measure the overall, or average, effect of an explanatory variable, as a coefficient estimated by a gravity model is meant to do. Therefore, I propose a *coefficient index* that implies the same expected aggregate trade flows, given the explanatory variables, as a given set of heterogeneous coefficient estimates, $\hat{\boldsymbol{\beta}}$.

Specifically, I define $\bar{\boldsymbol{\beta}}$ to be the uniform coefficient vector that satisfies the conditions

$$\sum_n \sum_i \left[\hat{X}_{ni}(\hat{\boldsymbol{\beta}}) - \hat{X}_{ni}(\bar{\boldsymbol{\beta}})\right] \boldsymbol{z}_{ni} = 0, \tag{12}$$

where $\hat{X}_{ni}(\hat{\boldsymbol{\beta}}) = \sum_k \hat{X}^k_{ni}(\hat{\boldsymbol{\beta}}^k)$, and $\hat{X}^k_{ni}(\hat{\boldsymbol{\beta}}^k)$ represents trade flows predicted by the structural gravity equation (6) given $\hat{\boldsymbol{\beta}}$. Because these conditions are analogous to aggregate versions of sample moment conditions in the class defined by (2)-(4), this index can be interpreted

[33] E.g., Anderson and Yotov (2010a,b), Caliendo and Parro (2015), and Levchenko and Zhang (2016).

analogously to the coefficients of an aggregate estimator. If $\boldsymbol{Z}$ were to contain country pair dummies, then $\hat{X}_{ni}(\bar{\boldsymbol{\beta}}) = \hat{X}_{ni}(\hat{\boldsymbol{\beta}})$ would hold for every country pair, and the coefficient index would be analogous to the Anderson and Neary (2003) index.[34] Otherwise, if $L < N^2$, then (12) imposes the weaker condition that deviations of $\hat{X}_{ni}(\bar{\boldsymbol{\beta}})$ from $\hat{X}_{ni}(\hat{\boldsymbol{\beta}})$ are uncorrelated with $\boldsymbol{z}_{ni}$. Note that because (6) and (7) are motivated independently from trade theory, this index generalizes to any application where one is interested in prediction conditional upon X_n^k and Y_i^k, rather than conditional upon the values of the fixed effects.

Pooled PPML is quite useful for calculating the coefficient index, as the following proposition shows.

Proposition 3. *The pooled product-level PPML estimator has the following properties:*

(i) *When applied to fitted values consistent with* (6) *and* (7)*, given parameter vector* $\hat{\boldsymbol{\beta}}$*, pooled PPML yields the coefficient index,* $\bar{\boldsymbol{\beta}}$*, defined by* (12)*.*

(ii) *When applied directly to product-level trade data, pooled PPML yields the coefficient index,* $\bar{\boldsymbol{\beta}}$*, defined by* (12)*, for the set of coefficients estimated by product-by-product PPML,* $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\text{PPML}}$*.*

Property (i) yields a straightforward method for calculating the coefficient index for *any* set of heterogeneous coefficients: simply perform pooled PPML using the fitted values based on those coefficients. Property (ii) states that, when applied to actual data, pooled PPML automatically estimates the coefficient index for product-by-product PPML. Property (ii) is very useful for inference because it implies that pooled PPML yields both an estimate of an ideal coefficient index and a valid estimate of the covariance matrix. Thus, any statistical tests that could be performed based on a sector-level estimation can be performed in the same way based on pooled PPML, and the results have a clear interpretation.[35]

4.3 Estimation with Missing Domestic Data

Data on production, expenditure, and domestic trade flows are typically not available at a level of disaggregation comparable to international trade data. This may appear to be a

[34] That is, $e^{\boldsymbol{z}'_{ni}\bar{\boldsymbol{\beta}}}$ is the uniform trade barrier that yields the same sector-level bilateral trade flows as the trade barriers implied by heterogeneous coefficients. This index differs in that it takes total product-level exports and imports as given, whereas the Anderson and Neary (2003) index requires specifying countries' expenditure functions over all products and assuming perfect competition. To borrow the language of Head and Mayer (2014), the coefficient index is a modular, rather than general equilibrium, index.

[35] Property (i) does not imply that PPML "estimation" using fitted values from a product-by-product estimation yields a valid estimate of the covariance matrix for the coefficient index, using standard techniques. A straightforward, though computationally intensive, method of obtaining this is the pairs cluster bootstrap discussed in detail by Cameron et al. (2011).

problem for structural estimators because (7) depends on the values of X_n and Y_i in the data. To see that this is not the case, note that $\boldsymbol{\beta}^k$ can always be specified to include country-specific border costs. These costs can be semi-parametrically estimated using the following specification:

Assumption 2′.

$$f(\boldsymbol{z}_{ni}, \boldsymbol{\beta}^k) = \tilde{f}(\tilde{\boldsymbol{z}}_{ni}, \boldsymbol{\gamma}^k) e^{\delta_n^k}$$

where $\tilde{\boldsymbol{z}}_{ni}$ *is a* $Q \times 1$ *vector of observable variables and* $\boldsymbol{\gamma}^k$ *is a* $Q \times 1$ *parameter vector.*[36]

In this specification, δ_n^k is an importer-specific border cost.[37] The structural estimators that I implement rely on the following result:

Proposition 4. *Given Assumption 2′, a structural gravity pseudo-maximum likelihood estimator for* $\boldsymbol{\gamma}^k$ *solves empirical moment conditions of the form*

$$\sum_{n=1}^{N} \sum_{i \neq n} r(X_{ni}^k, \tilde{X}_{ni}^k(\hat{\boldsymbol{\gamma}}^k)) \tilde{\boldsymbol{Z}}_{ni} = 0, \quad \forall k,$$

where

$$\tilde{X}_{ni}^k(\hat{\boldsymbol{\gamma}}^k) = \frac{M_n^k}{\tilde{\Phi}_n^k} \frac{E_i^k}{\tilde{\Psi}_i^k} \tilde{f}(\tilde{\boldsymbol{z}}_{ni}, \hat{\boldsymbol{\gamma}}^k), \tag{13}$$

for all $n \neq i$, *and where* $M_n^k = \sum_{i \neq n} X_{ni}^k$, $E_i^k = \sum_{n \neq i} X_{ni}^k$, *and* $\tilde{\Phi}_n^k$ *and* $\tilde{\Psi}_i^k$ *are given by*

$$\tilde{\Phi}_n^k = \sum_{i \neq n} \frac{E_i^k}{\tilde{\Psi}_i^k} \tilde{f}(\tilde{\boldsymbol{z}}_{ni}, \hat{\boldsymbol{\gamma}}^k) \quad \textit{and} \quad \tilde{\Psi}_i^k = \sum_{n \neq i} \frac{M_n^k}{\tilde{\Phi}_n^k} \tilde{f}(\tilde{\boldsymbol{z}}_{ni}, \hat{\boldsymbol{\gamma}}^k). \tag{14}$$

Equation (13) is a slight variation of (6) with two important distinctions. First, it is specified in terms of product-level exports and imports, rather than production and expenditure. Second, trade flows depend only on $\boldsymbol{\gamma}^k$, the vector of coefficients excluding border costs, $\boldsymbol{\delta}^k$. Thus, Proposition 4 implies structural gravity PML estimators can identify $\boldsymbol{\gamma}$ using highly disaggregated trade data without comparable domestic data and without estimating $\boldsymbol{\delta}$.

Proposition 4 also implies that it is possible to calculate the coefficient index for $\hat{\boldsymbol{\gamma}}$ without

[36] For notational completeness, under Assumption 2′, $\boldsymbol{\beta}^k = (\boldsymbol{\gamma}^{k\prime}, \boldsymbol{\delta}^{k\prime})'$, $\boldsymbol{z}_{ni} = (\boldsymbol{z}_{ni}', \mathbb{1}_{ni}')'$, $\boldsymbol{\delta}^k = (\delta_1^k, ..., \delta_n^k)'$, and $\mathbb{1}_{ni}$ is an $N \times 1$ vector with n^{th} element equal to 1 when $n \neq i$ and all other elements equal to zero.

[37] All the results that follow are isomorphic to the specification of border costs as importer-specific, exporter-specific, or some combination. While Waugh (2010) argues that exporter-specific border costs are more consistent with international data on prices of tradeable goods, Ramondo et al. (2016) find that domestic trade frictions can account for much of this phenomenon.

data on domestic trade flows or an estimate of $\boldsymbol{\delta}$. Using this result, (12) can be rewritten as

$$\sum_{n}\sum_{i\neq n}\left[\tilde{X}_{ni}(\bar{\boldsymbol{\gamma}})-\hat{X}_{ni}(\tilde{\boldsymbol{\gamma}})\right]\boldsymbol{z}_{ni}=0, \tag{15}$$

where $\tilde{X}_{ni}(\hat{\boldsymbol{\gamma}})=\sum_k \tilde{X}^k_{ni}(\hat{\boldsymbol{\gamma}}^k)$.

4.4 Identifying Border Costs

While identification of $\boldsymbol{\gamma}$ does not require identification of $\boldsymbol{\delta}$, border cost estimates are themselves of interest in many contexts. Though data on X^k_{nn} is typically not available, sector-level domestic trade flow data often is available. In such cases, it is possible to identify the sector-level coefficient index of $\hat{\boldsymbol{\delta}}$, even though the individual elements of $\hat{\boldsymbol{\delta}}$ cannot be identified. Given a consistent estimate of $\boldsymbol{\gamma}$, the predicted value of X^k_{nn}, based on (13), is

$$\hat{X}^k_{nn}=e^{-\hat{\delta}^k_n}\frac{M^k_n}{\tilde{\Phi}^k_n}\frac{E^k_i}{\tilde{\Psi}^k_i}\tilde{f}(\tilde{\boldsymbol{z}}_{nn},\hat{\boldsymbol{\gamma}}^k),$$

where $\tilde{\Phi}^k_n$ and $\tilde{\Psi}^k_i$ solve (14) given $\hat{\boldsymbol{\gamma}}^k$. If data on X^k_{nn} were available, then a valid method-of-moments estimator of δ^k_n would be

$$\hat{\delta}^k_n=\ln\left(\frac{M^k_n}{\hat{\tilde{\Phi}}^k_n}\frac{E^k_i}{\hat{\tilde{\Psi}}^k_i}\tilde{f}(\tilde{\boldsymbol{z}}_{nn},\hat{\boldsymbol{\gamma}}^k)\right)-\ln(X^k_{nn}).$$

Because the values of $\tilde{\Phi}^k_n$ and $\tilde{\Psi}^k_i$ do not depend on δ^k_n, the elements of $\bar{\boldsymbol{\delta}}$ can be easily calculated as

$$\bar{\delta}_n=\ln\left(\sum_k\frac{M^k_n}{\hat{\tilde{\Phi}}^k_n}\frac{E^k_i}{\hat{\tilde{\Psi}}^k_i}\tilde{f}(\tilde{\boldsymbol{z}}_{nn},\bar{\boldsymbol{\gamma}})\right)-\ln(X_{nn}), \tag{16}$$

which does not require knowledge of X^k_{nn}. Note that $\bar{\delta}_n$ is calculated based on the coefficient index $\bar{\boldsymbol{\gamma}}$, not on heterogeneous coefficient estimates.[38] This is because (12) defines the coefficient index for all coefficients simultaneously. While, it is straightforward to calculate an index of $\hat{\boldsymbol{\delta}}$ given a set of heterogeneous coefficients $\hat{\boldsymbol{\gamma}}$, it is not clear how one should interpret a uniform border cost index together with heterogeneous bilateral trade costs.

Together, Proposition 4 and equations (15) and (16) provide a method to estimate the coefficient index $\bar{\boldsymbol{\beta}}=(\bar{\boldsymbol{\gamma}}',\bar{\boldsymbol{\delta}}')'$ given product-level international trade flow data and sector-level domestic trade flow data. The procedure is as follows: (i) Estimate $\hat{\boldsymbol{\gamma}}$ based on (13), (ii) compute $\bar{\boldsymbol{\gamma}}$ using (15), and (iii) compute $\bar{\boldsymbol{\delta}}$ using (16).

[38] The fitted MR terms in (16) are also calculated given $\bar{\boldsymbol{\gamma}}$.

Table 1: Estimation Specifications

		Single-Sector			Multi-Sector		Product-
		Sector-Level	Pooled-IL	Pooled-PL	Sector-Level	Pooled-PL	by-Product
Log-LS	FE	X	X	X	X	X	X
	Struc.	X	X	X	X	X	X
GPML	FE	X					
	Struc.	X	X	X	X	X	X
PPML	FE	X	X	X	X	X	X
	Struc.	X	X	X	X	X	X
NLS	FE	X					
	Struc.	X	X	X	X	X	X

4.5 Estimation Specifications

In the empirical application, I perform sector- and product-level gravity estimations that vary along four dimensions: objective function, controls for fixed effects, pooling, and sector definition. Table 1 summarizes the specifications that I estimate. Based on the insights of my model and the known properties of the estimators discussed, pooled product-level PPML is my preferred estimator for applications where one is primarily concerned with the overall magnitude of trade barriers or the overall effect of one or more variables on trade flows. This is due to (a) the robustness of PPML estimates in the presence of heteroskedasticity and zeros and its lack of asymptotic bias, (b) the consistency between structural and FE PPML estimators, (c) the ease of interpretation and the comparability of pooled estimates with the aggregate estimates typical of the literature, and (d) the interpretability of pooled PPML as an ideal index of heterogeneous coefficients.

4.6 Data

I use data on bilateral product-level trade flows from the U.N. Comtrade database for 2003, classified at the 6-digit level of the Harmonized System (1996 revision). Data on bilateral relationships are taken from CEPII's Gravity dataset. For specifications that treat manufacturing as a single sector, total manufacturing output is taken from the OECD STAN database, where available, or the UNIDO INDSTAT database. Where not available from either source, it is imputed using manufacturing value added from the World Bank's WDI database. For multi-sector specifications, output by 2-digit ISIC (Rev. 3) industry is taken from the UNIDO INDSTAT database. The full sample consists of trade flows among 130 countries classified into 4,608 product categories. Table A1 lists the countries in the sample and the source of manufacturing output data for each. Industry-level output data is only available for a subset of these countries, so $\bar{\delta}_n^j$ is not identified for all countries in industry-by-industry estimations. Table A2 lists the set of industries and number of countries with output data for each industry. Further details are provided in Appendix C.

For the panel estimation in Section 6, trade data are from the NBER-UN dataset described by Feenstra et al. (2005). I use data for 1965-2000, in five-year increments, at the 4-digit SITC (Rev. 2) level, for a balanced panel of countries. This sample includes 87 countries, 8 time periods, and 641 manufacturing SITC codes. While not at the lowest level of aggregation available, this dataset features the largest number of countries and years for a consistent classification of product-level trade flows. Data on free trade agreements are from Baier and Bergstrand (2007).[39]

5 Empirical Results

In the baseline estimations, to keep the specification as parsimonious as possible and consistent with log-linear specifications in the literature, I specify $\tilde{f}(\tilde{\boldsymbol{z}}_{ni}, \boldsymbol{\gamma}^k) = e^{\tilde{\boldsymbol{z}}'_{ni}\boldsymbol{\gamma}^k}$, where $\tilde{\boldsymbol{z}}_{ni}$ includes (logged) bilateral distance and an indicator for whether countries shares a common border. This is comparable to many studies, such as Anderson and van Wincoop (2003) and Waugh (2010), that use distance and political borders as proxies for geographic and policy barriers to trade, respectively. In Section 6, I expand the set of covariates to include additional variables related to trade policy.

5.1 Baseline Estimates

Table 2 presents the estimates from the baseline single-sector specifications. Table 3 presents the coefficient indexes based on the multi-sector and product-by-product estimates. Full results for the multi-sector estimations are presented in Tables A3 and A4 in Appendix A.

The single-sector, sector-level estimates (Table 2a) are roughly in line with the literature. Bilateral trade is generally decreasing in distance and higher if countries share a border, and there is a great deal of variation across sector-level estimators. The distance elasticity varies by a factor of 2.5 from -0.94 to -2.41, and the effect of sharing a border varies from a statistically insignificant 0.20 to 0.96. The pooled industry-level (Table 2b) and industry-by-industry estimates (Table 3a) show a nearly identical pattern. Santos Silva and Tenreyro (2006) attribute such differences to bias due to heteroskedasticity and sample selection associated with Log-LS and to finite sample biases for the PML estimators.

[39]Updated data are available on Jeffrey Bergstrand's website (`https://www3.nd.edu/~jbergstr/`).

Table 2: Single-Sector Estimation Results

(a) Sector-Level Estimations

	Log LS		GPML		PPML		NLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean($\bar{\delta}_n$)	−0.83	−0.78	−0.18	−2.42	−3.70	−3.70	−3.79	−3.77
Distance	−1.95	−2.11	−2.41	−1.50	−0.96	−0.96	−0.94	−0.95
	(0.07)	(0.10)	(0.14)	(0.07)	(0.07)	(0.07)	(0.13)	(0.11)
Shared Border	0.96	0.54	0.66	0.20	0.54	0.54	0.49	0.50
	(0.19)	(0.35)	(0.29)	(0.24)	(0.12)	(0.12)	(0.18)	(0.21)
Exporter FEs	Y		Y		Y		Y	
Importer FEs	Y		Y		Y		Y	
Structural MR		Y		Y		Y		Y
Observations	11,193	11,193	16,770	16,770	16,770	16,770	16,770	16,770

(b) Pooled Industry-Level Estimations

	Log LS		GPML		PPML		NLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean($\hat{\delta}_n$)	−0.57	−0.81		−2.29	−3.33	−3.33		−3.30
Distance	−1.94	−1.95		−1.41	−1.01	−1.01		−1.04
	(0.06)	(0.09)		(0.09)	(0.07)	(0.07)		(0.13)
Shared Border	1.01	0.69		0.44	0.55	0.55		0.47
	(0.14)	(0.26)		(0.21)	(0.12)	(0.12)		(0.20)
Exp-Prod FEs	Y				Y			
Imp-Prod FEs	Y				Y			
Structural MR		Y		Y		Y		Y
Observations	11,193	11,193		16,770	16,770	16,770		16,770

(c) Pooled Product-Level Estimations

	Log LS		GPML		PPML		NLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean($\bar{\delta}_n$)	−2.74	−2.24		−2.96	−2.95	−2.95		−3.32
Distance	−1.15	−1.35		−1.20	−1.16	−1.16		−1.02
	(0.05)	(0.07)		(0.15)	(0.08)	(0.08)		(0.11)
Shared Border	0.79	0.75		0.37	0.51	0.51		0.55
	(0.10)	(0.11)		(0.32)	(0.12)	(0.12)		(0.21)
Exp-Prod FEs	Y				Y			
Imp-Prod FEs	Y				Y			
Structural MR		Y		Y		Y		Y
Observations	3,571,896	3,571,896		77,276,160	77,276,160	77,276,160		77,276,160

Notes: Standard errors (in parentheses) are robust to multi-way clustering by both importer and exporter.

Table 3: Multi-Sector Estimation Results

(a) Coefficient Index: Sector-Level Industy-by-Industry Estimations

	Log LS		GPML		PPML		NLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean($\bar{\bar{\delta}}_n$)	−0.75	−0.96		−2.18	−3.33	−3.33		−3.33
Distance	−1.88	−1.91		−1.49	−1.01	−1.01		−1.02
	[0.08]	[0.13]		[0.10]	[0.11]	[0.11]		[0.13]
Shared Border	0.99	0.58		0.31	0.55	0.55		0.52
	[0.19]	[0.44]		[0.27]	[0.16]	[0.16]		[0.22]
Exp-Ind FEs	Y				Y			
Imp-Ind FEs	Y				Y			
Structural MR		Y		Y		Y		Y
Observations	11,193	11,193		16,770	16,770	16,770		16,770

(b) Coefficient Index: Pooled Product-Level Industy-by-Industry Estimations

	Log LS		GPML		PPML		NLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean($\bar{\bar{\delta}}_n$)	−2.63	−2.14		−2.64	−2.95	−2.95		−3.05
Distance	−1.19	−1.38		−1.38	−1.16	−1.16		−1.13
	[0.06]	[0.08]		[0.13]	[0.10]	[0.10]		[0.19]
Shared Border	0.80	0.79		0.19	0.51	0.51		0.52
	[0.12]	[0.14]		[0.29]	[0.17]	[0.17]		[0.32]
Exp-Prod FEs	Y				Y			
Imp-Prod FEs	Y				Y			
Structural MR		Y		Y		Y		Y
Observations	3,571,896	3,571,896		77,276,160	77,276,160	77,276,160		77,276,160

(c) Coefficient Index: Product-by-Product Estimations

	Log LS		GPML		PPML		NLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean($\hat{\delta}_n$)	−2.28	−1.82		−2.23	−2.95	−2.95		−2.91
Distance	−1.30	−1.47		−1.48	−1.16	−1.16		−1.17
	[0.16]	[0.16]		[0.19]	[0.24]	[0.24]		[0.37]
Shared Border	0.86	0.88		0.36	0.51	0.51		0.56
	[0.32]	[0.32]		[0.41]	[0.40]	[0.40]		[0.59]
Exp-Prod FEs	Y				Y			
Imp-Prod FEs	Y				Y			
Structural MR		Y		Y		Y		Y
Observations	3,570,985	3,571,727		76,739,520	76,655,670	76,655,670		73,402,290

Notes: Standard errors are robust to multi-way clustering by both importer and exporter. Numbers in square brackets are median standard errors across industries for industry-by-industry estimations and across products for product-by-product estimations.

By contrast, the estimates are much more consistent across product-level estimators. To demonstrate the economic significance of these differences, Figure 1 plots the cumulative effect of distance on bilateral trade flows estimated by each specification.[40] It is clear that, over the observed range of distances, the sector-level estimates diverge markedly, while the product-level estimates are nearly indistinguishable. To place these numbers in perspective, given a trade cost elasticity equal to 6, single-sector, sector-level PPML implies a median bilateral trade cost equivalent to a 725% ad valorem tariff, versus 2,299% for FE Log-LS. The equivalent values for the respective product-level estimators are 858% and 818%. These results indicate that failing to control for PLCA significantly biases sector-level estimates, even at the 2-digit industry level. The similarity among product-level estimates suggests that most of the discrepancy among sector-level estimators results from misspecification of the conditional expectation of X_{ni}^{j} by omitting T_{ni}^{j}, rather than from heteroskedasticity, sample-selection, or finite sample biases.

Further, the results suggest that the choice between sector-level and product-level estimation is of greater consequence than the definition of a sector. Because pooled PPML automatically delivers the ideal coefficient index, all three product-level specifications estimate identical coefficients. Even for the other estimators, the differences across sector definitions tend to be smaller than the differences between sector-level and product-level estimates. Interestingly, the product-by-product estimates diverge somewhat more across estimators than the pooled estimates. This may indicate that the heteroskedasticity, sample-selection, and finite sample biases are more severe when allowing for more heterogeneity. The latter two seem especially plausible because the heterogeneous specifications necessarily have fewer observations to identify each parameter, and the heterogeneous coefficient estimates are considerably noisier than the pooled estimates. In addition, comparing the estimates in Tables 2b and 3a with those in Table 2a shows that industry-level versus fully aggregate estimation makes little difference.[41]

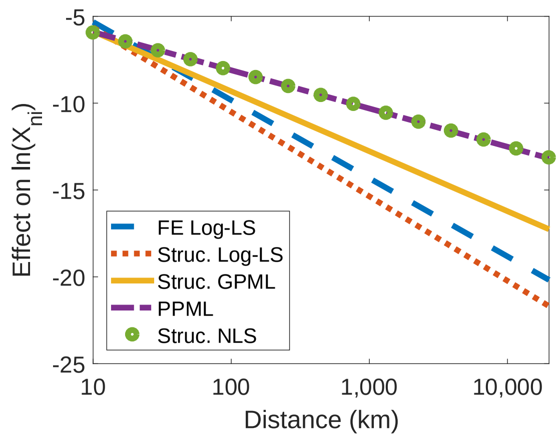
To gain insight into the underlying sources of bias in the sector-level estimates, I decompose the difference between sector-level and product-level estimates into components attributable to PLCA, parameter heterogeneity, and other forms of misspecification or finite sample biases that induce interactions between the estimation error and the components of T_{ni}. To do so, I repeat the sector-level estimations using aggregated trade flows predicted by (13) and (14), based on product-level estimates. The bias attributable to PLCA is the difference between sector-level estimates based on values predicted under the assumption that

[40] Figure 1 plots the value of mean$(\hat{\hat{\delta}}_n) + \hat{\hat{\gamma}}_{\text{Distance}} \ln(\text{Distance})$.

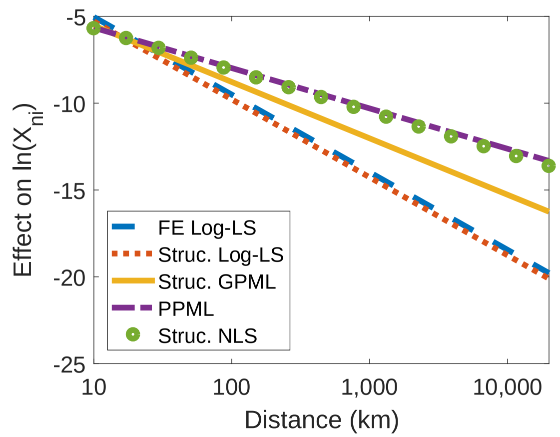
[41] Anderson and Yotov (2010a,b) find differences between aggregate and industry-by-industry estimates. However, they consider only a simple average, not an ideal index, of industry-level coefficients and do not test for statistical significance of the differences.

Figure 1: Estimated Distance-Related Trade Costs

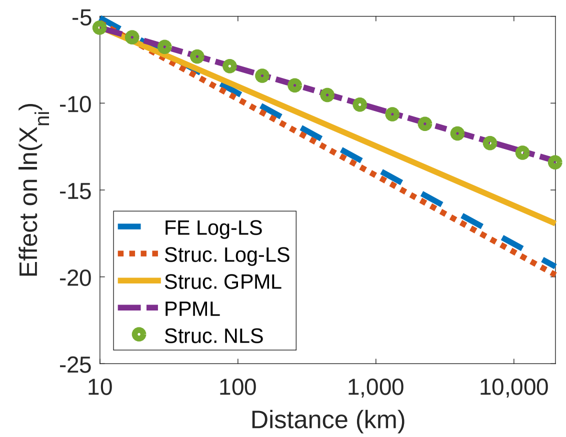
(a) Sector-Level Estimates



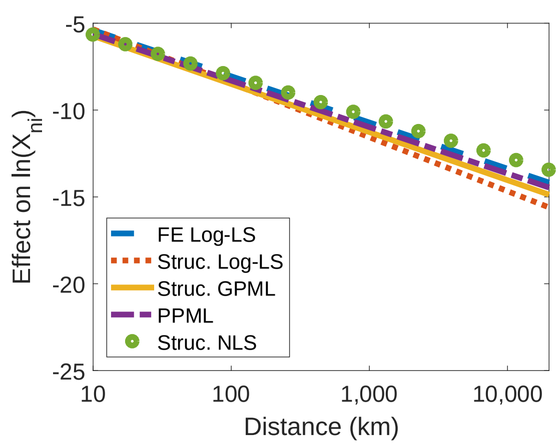
(b) Pooled Industry-Level Estimates



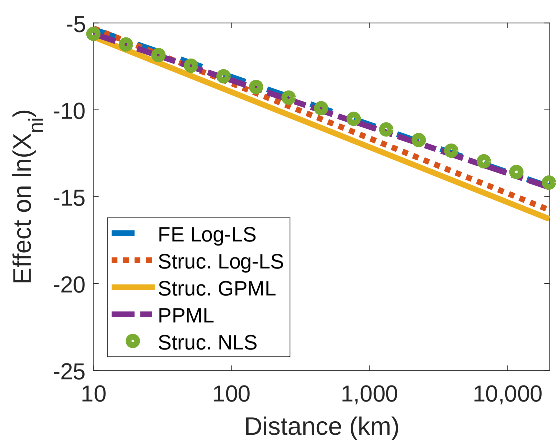
(c) Ind.-by-Ind. Estimates



(d) Pooled Product-Level Estimates



(e) Pooled PL Ind.-by-Ind. Estimates



(f) Product-by-Product Estimates

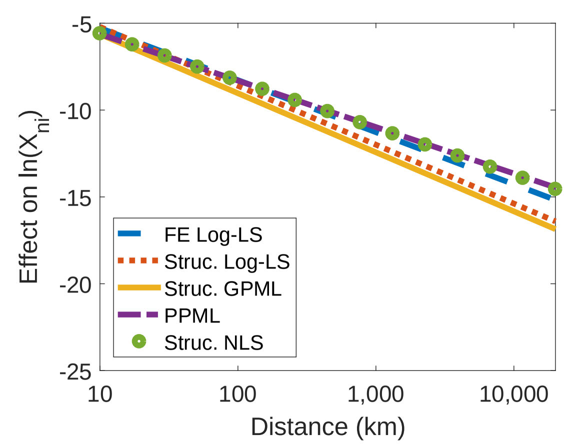


Table 4: Sources of Bias in Sector-Level Estimates

		Comparative Advantage	Heterogeneity	Finite Sample/ Misspecification	Total
FE Log-LS	Distance	0.13	0.04	-0.83	-0.65
	Border	0.01	0.18	-0.09	0.10
Struc. Log-LS	Distance	0.18	0.00	-0.82	-0.64
	Border	-0.11	0.07	-0.30	-0.34
Struc. GPML	Distance	0.33	0.02	-0.38	-0.02
	Border	0.25	0.07	-0.47	-0.16
PPML	Distance	0.21	0	0	0.21
	Border	0.02	0	0	0.03
Struc. NLS	Distance	0.21	-0.01	0.03	0.23
	Border	-0.02	-0.05	0.00	-0.06

each product-specific coefficient vector is equal to the coefficient index of product-by-product estimates, $\boldsymbol{\gamma}^k = \bar{\boldsymbol{\gamma}}$, and the coefficient index itself. The bias attributable to heterogeneity is the difference between sector-level estimates based on predicted values from the product-by-product estimates, and those under the assumption that $\boldsymbol{\gamma}^k = \bar{\boldsymbol{\gamma}}$. The bias attributable to the final source is the difference between the sector-level estimates from the actual data and those based on the product-by-product predicted values. By construction, the three bias measures sum to the difference between the sector-level estimates and the coefficient index of product-by-product estimates. To summarize, the first measure is the effect of aggregation on estimates given homogeneous parameters, the second is the additional effect of parameter heterogeneity, and the third is the effect of allowing T_{ni} to include the product-level estimation errors.

Table 4 reports the bias decomposition for each estimator. As expected, failing to control for PLCA biases the distance elasticity away from zero, and it does so by a similar magnitude for each estimator. The direction and magnitude of the bias in the shared border effect varies across estimators, which is not surprising given that shared border accounts for a relatively small share of overall trade costs. Ignoring heterogeneity generally results in relatively little bias once PLCA has been controlled for. PPML has exactly zero additional bias due to parameter heterogeneity because sector-level PPML estimates an ideal coefficient index under the assumption of no PLCA.[42] The final form of bias is very large for the log-LS and GPML

[42] Specifically, sector-level PPML imposes the condition that $\phi_n^k = \bar{\phi}_n \phi^k$ and $\psi_i^k = \bar{\psi}_i \psi^k$.

estimators. This phenomenon appears to be driven by zeros in the product-level data.[43] This is consistent with the findings of Santos Silva and Tenreyro (2006) of significant bias in log-LS and GPML estimates in the presence of zeros. These results indicate that this form of bias is small in estimates that control for PLCA, but the bias interacts with patterns of comparative advantage to produce large biases in sector-level estimates. This bias is also zero for PPML, which imposes that the product-level estimation errors, rather than nonlinear transformations of the errors, are orthogonal to the components of T_{ni}.[44]

5.2 Tests for Bias

I this section, I formally test for the bias in sector-level estimates indicated by the results summarized by Tables 2 and 3. I first consider tests based on the single-sector estimates in detail before summarizing the results of analogous tests based on the multi-sector estimates.

5.2.1 Single-Sector Tests

First, I perform an auxiliary estimation to test whether sector-level trade flows depend upon PLCA.[45] Pooled product-level estimation yields fitted values of T_{ni}, up to an importer- and exporter-specific scale factor. Using these fitted values, I perform a sector-level estimation that takes the following form:

$$E[X_{ni}] = \rho_n \upsilon_i e^{\tilde{\mathbf{z}}'_{ni}\hat{\boldsymbol{\gamma}}} \hat{T}_{ni}^{\alpha}$$

where ρ_n and υ_i are importer and exporter FEs. I do this for each of the sector-level estimators that have product-level analogues, using the values of $\hat{T}_{ni}$ calculated from the analogous product-level estimation. The first row of Table 5 presents the test statistics and p-values of the null hypothesis $\alpha = 0$, which is resoundingly rejected for each estimator. Thus, PLCA is a significant driver of sector-level trade flows.

The remaining rows of Table 5 present the results of three cluster-robust Hausman tests based on the comparison of estimates of $\boldsymbol{\gamma}$ across estimators. If the conditional expectation of X_{ni}^k is correctly specified and if sector-level trade flows are unaffected by PLCA, then all

[43] Equations (13) and (14) always predict positive (though potentially very small) trade flows as long as M_n^k and E_i^k are positive. Generating zeros by either imposing that $\hat{X}_{ni}^k = 0$ if $X_{ni}^k = 0$ or by rounding all values for which $\hat{X}_{ni}^k < \bar{x}$ to zero, where $\bar{x}$ is chosen to equate the number of zeros in the data and fitted values, yields estimates much closer to those obtained from the actual aggregated data.

[44] This is not to say that PPML cannot suffer from finite sample bias (see Head and Mayer (2014) for a counterexample), only that any such bias is identical in sector-level and product-level estimates.

[45] Directly testing for heterogeneity in the country-by-product fixed effects is not straightforward. The shear number of parameters makes it practically infeasible to compute the covariance matrix required to test such a hypothesis. In addition, as Cameron and Miller (2015) point out, if there is significant clustering in estimation errors, which appears to be the case with product-level trade data, the estimated variance matrix will be rank deficient, making such a test impossible even given sufficient computing power.

Table 5: Hausman Tests for Bias in Sector-Level Estimates

H_0	Test Statistic	FE Log LS	Structural Log LS	Structural GPML	Structural PPML	Structural NLS
$\alpha = 0$	$t(129)$	20.69	23.72	31.49	24.06	16.83
	$P > \lvert t \rvert$	0.000	0.000	0.000	0.000	0.000
$\gamma_{\text{Agg}} = \gamma_{\text{Pool}}$	$F(2, 129)$	57.84	73.45	2.78	21.42	3.07
	$P > F$	0.000	0.000	0.066	0.000	0.050
$\gamma_{\text{Agg}} = \gamma_{\text{Agg,PPML}}$	$F(2, 129)$	45.27	140.78	32.83		0.19
	$P > F$	0.000	0.000	0.000		0.828
$\gamma_{\text{Pool}} = \gamma_{\text{Pool,PPML}}$	$F(2, 129)$	1.05	9.88	0.10		1.86
	$P > F$	0.354	0.000	0.908		0.159

Notes: All tests are based on estimated variance matrices that are robust to multi-way clustering by both importer and exporter. In the calculation of p-values, the test statistic is treated as being distributed $t(N-1)$, for the first row, and $F(Q, N-1)$, for the remaining rows, where $N = 130$ is the number of clusters, and $Q = 2$ is the number of elements of $\boldsymbol{\gamma}$.

of the estimators presented in Table 2 will have the same probability limit.[46] Therefore, I test for bias in sector-level estimators by testing the equality of the sector- and product-level estimates. The second row of Table 5 presents the results of these tests. Equality is resoundingly rejected for both log-LS estimators and PPML. It is rejected at the 5% significance level for structural NLS and at the 10% level for structural GPML. This result permits the formal conclusion that sector-level gravity estimates are biased due to PLCA.

The final two rows of Table 5 test the hypotheses that the sector- and product-level estimates are equal to their PPML counterparts. Equality is resoundingly rejected for all sector-level estimators except NLS and cannot be rejected at any reasonable level of significance for the product-level estimators other than structural log-LS. Even in the latter case, the product-level test statistic is dramatically smaller. These tests formally confirm the patterns apparent in Table 2 and support the conclusion that differences in sector-level estimates can be attributed to failure to control for PLCA. In fact, it is not possible to reject the hypothesis that product-level FE log-LS is unbiased, despite the nearly forgone conclusion in the literature to the contrary.[47]

[46]In the case of log-LS estimators, this also requires that these estimators not be inconsistent due to heteroskedasticity and sample-selection bias.

[47]While the NLS estimates are surprisingly consistent across specifications, this is a notoriously unreliable estimator (see, e.g., Santos Silva and Tenreyro, 2006) that is very sensitive to outliers. This is consistent with the relatively large standard errors for pooled product-level NLS. Therefore, I caution against the adoption of sector-level NLS based on these results alone.

Table 6: Multi-Sector Hausman Tests for Bias: Share Rejected

H_0	FDR	FE Log LS	Structural Log LS	Structural GPML	Structural PPML	Structural NLS
$\alpha = 0$	0.10	1.00	1.00	1.00	1.00	1.00
	0.01	1.00	1.00	1.00	1.00	1.00
$\gamma^j_{\text{Agg}} = \gamma^j_{\text{Pool}}$	0.10	1.00	1.00	0.48	0.90	0.57
	0.01	1.00	1.00	0.29	0.81	0.52
$\gamma^j_{\text{Agg}} = \gamma^j_{\text{Agg,PPML}}$	0.10	1.00	1.00	0.86		0.05
	0.01	1.00	1.00	0.76		0.00
$\gamma^j_{\text{Pool}} = \gamma^j_{\text{Pool,PPML}}$	0.10	0.33	0.62	0.52		0.14
	0.01	0.14	0.57	0.24		0.10

Notes: False discovery rate (FDR), the expected share of rejected null hypotheses that are Type I errors, is controlled at the specified level using the method of Benjamini and Hochberg (1995). All tests are based on estimated variance matrices that are robust to multi-way clustering by both importer and exporter (except where indicated in Tables A3 and A4 in Appendix A). Test statistics are treated as being distributed $F(Q, N-1)$ where $N = 130$ is the number of clusters, and $Q = 2$ is the number of elements of $\boldsymbol{\gamma}$.

5.2.2 Multi-Sector Tests

Tables A3 and A4 in Appendix A present test statistics for the industry-by-industry estimates analogous to those in Table 5.[48] Table 6 summarizes these results, presenting the share of industries for which each hypothesis is rejected. To control for overrejection due to multiple inference, Table 5 reports the share of rejections after controlling the false discovery rate (FDR), or the expected share of rejected null hypotheses that are actually true, at the 10% and 1% levels, using the procedure of Benjamini and Hochberg (1995).

We can resoundingly reject the hypothesis that $\alpha = 0$ in the auxiliary estimations for every industry and estimator. Thus, PLCA is a significant driver of sector-level trade flows, even within 2-digit industries. As with the single-sector estimations, in most cases we can reject equality between sector-level and product-level estimates, indicating that significant bias remains at the 2-digit industry level. For all but NLS, we can reject equality between the PPML estimates and the others for almost all sector-level estimations, where the failure to reject for NLS is largely due to the inefficiency of this estimator. We cannot reject equality with PPML for the product-level estimators for a large share of sectors. Thus, the overall findings are consistent with the single-sector results and support the conclusion that sector-level estimates are significantly biased and that differences among sector-level estimators can be largely attributed to PLCA.

[48] For NLS, multi-way clustering by exporter and importer produced an estimated covariance matrix that was not positive semi-definite for a handful of industries. In these instances, which are indicated in Tables A3 and A4, the test statistic is robust to clustering by importer.

5.3 Tests for Pooling

The empirical results provide substantial evidence that sector-level gravity estimates are significantly biased due to PLCA, regardless of whether the estimation allows for heterogeneity across industries. I now turn to the question of whether heterogeneity in common parameters is a significant concern in its own right and thus whether the preferred alternative to sector-level estimation is pooled or product-by-product estimation. For brevity, I consider only PPML estimates because pooled PPML automatically delivers an ideal coefficient index.

There is substantial variation in the estimated heterogeneous coefficients. However, these coefficients are identified much less precisely than their homogeneous counterparts. I take two approaches to evaluate the significance of this heterogeneity. First, I test for equality between the heterogeneous and homogeneous coefficients. Second, I evaluate the out-of-sample predictive power of the homogeneous and heterogeneous estimators.

Standard Chow-type tests for poolability test whether every heterogeneous coefficient is equal to its homogeneous counterpart and thus rejects if *any* of the coefficients is significantly different.[49] Given the number of products in the dataset, it would be shocking if the null of homogeneity were not rejected by such a test. However, what is of practical concern in most applications is not whether there is any heterogeneity but *how much* heterogeneity is present in the parameter estimates and whether this heterogeneity significantly affects inference or predictions regarding the effects of the explanatory variables on overall trade flows. Therefore, I conduct a series of Hausman tests for equality of the heterogeneous coefficients with their homogeneous counterparts one-by-one for each industry and product.

Table 7 shows the share of industries and products for which we can reject equality with the homogeneous coefficients, controlling the FDR as in Table 6, for distance and shared border separately and joint tests for both coefficients. There is significant, though far from universal, heterogeneity across both industries and products within industries. Unlike the tests for bias, this conclusion is not qualitatively changed when we control for PLCA, though there is some disagreement between the sector- and product-level estimates regarding *which* industries and products are significantly heterogeneous. Interestingly, there is far more heterogeneity in the distance elasticity than the shared border effect, indicating that the heterogeneity lies primarily in the effect of distance on trade costs, not in the trade elasticity, because the latter would tend to shift coefficients proportionally for a given product.

Despite evidence of heterogeneity, the reporting, multiple inference, and overfitting con-

[49]See Baltagi (2005). These tests are typically only valid under the assumption of normally distributed errors. Further, if errors are clustered by country, which appears to be the case, such tests are not possible if the number or products exceeds the number of countries due to the rank deficiency of the covariance matrix for the heterogeneous coefficients (Cameron and Miller, 2015).

Table 7: Hausman Tests for Heterogeneity: Share Rejected

H_0	FDR	Distance	Shared Border	All Coeffs.
$\gamma^j_{\text{Agg}} = \gamma_{\text{Agg}}$	0.10	0.52	0.14	0.90
	0.01	0.33	0.05	0.43
$\gamma^{jk} = \gamma^j_{\text{Agg}}$	0.10	0.45	0.15	0.49
	0.01	0.19	0.03	0.26
$\gamma^j_{\text{Pool}} = \gamma_{\text{Pool}}$	0.10	0.48	0.10	0.71
	0.01	0.38	0.05	0.67
$\gamma^{jk} = \gamma^j_{\text{Pool}}$	0.10	0.40	0.14	0.45
	0.01	0.16	0.02	0.24

Notes: False discovery rate (FDR), the expected share of rejected null hypotheses that are Type I errors, is controlled at the specified level using the method of Benjamini and Hochberg (1995). All tests are based on estimated variance matrices that are robust to multi-way clustering by both importer and exporter. Test statistics are treated as being distributed $F(L, N-1)$ where $N = 130$ is the number of clusters, and L is the number of restrictions tested.

cerns mean there remains a pragmatic argument for using a pooled estimator when heterogeneity is not of primary interest. Therefore, I test the out-of-sample performance of each PPML estimator. Concerns with multiple inference and overfitting are not unique to gravity estimation. In a number of fields, split-sample techniques are often employed to address these issues.[50] I adapt these techniques to gravity estimation to develop a test for overfitting that is intuitive and straightforward to implement.

The basic strategy is to split the data into training and evaluation samples, use the former for estimation, and evaluate the performance of the estimated models in the latter. There are two non-trivial choices to make before applying this strategy to gravity estimation: the loss function and the evaluation sample. I use two loss functions: root mean square error (RMSE) and the (negative) Poisson log-likelihood. RMSE is the standard loss function in the time-series and panel forecasting literatures, while PPML maximizes the Poisson log-likelihood in the training samples.[51] In time-series and panel data models with a time dimension, it is natural to split the sample into earlier and later periods. With micro data, observations are split randomly into two samples that are identical in expectation. With bilateral trade data, neither practice is appropriate. There is no natural ordering of observations, and trade flows almost certainly violate the i.i.d. assumption implicit in random assignment. Instead, I split

[50] E.g., evaluation of time series models, based on Diebold and Mariano (1995), and panel estimators, as in Baltagi et al. (2000) and those surveyed by Baltagi (2008). Fafchamps and Labonne (2017) and Anderson and Magruder (2017) propose split-sample techniques to identify of causal effects in micro data.

[51] To a second-order approximation, the negative Poisson log-likelihood is equal to $\sum(X^k_{ni} - \hat{X}^k_{ni})^2/X^k_{ni}$ up to addition by a constant. Thus, it penalizes deviations for large observations less than RMSE.

Table 8: In- and Out-of-Sample Fit of PPML Estimators

	Criterion	Sector-Level Single-Sector	Sector-Level Multi-Sector	Pooled Product-Level Single-Sector	Pooled Product-Level Multi-Sector	Product-by-Product
In-Sample	Pois. LLF	1.055	1.071	1.102	1.102	1.104
	RMSE	4.602	3.352	1.608	1.551	1.445
Out-of-Sample	Pois. LLF	1.045	1.060	1.091	1.091	1.089
	RMSE	4.639	3.446	1.692	1.699	1.750
Best Predictor	Pois. LLF	0	0	0.400	0.595	0.005
	RMSE	0	0	0.444	0.311	0.245

Notes: RMSE $\times 10^6$. Poisson LLF $\times 10^6$. Poisson LLF $= \frac{1}{BN(N-1)K} \sum_b \sum_n \sum_{i \neq n} \sum_k X_{ni}^{k(-b)} \ln\left(\hat{X}_{ni}^{k(-b)}\right) - \hat{X}_{ni}^{k(-b)}$

the set of *countries* randomly, which preserves the network structure and error clustering present in the data. To ensure that the findings are representative of the overall sample, I repeat the analysis many times across random sub-samples.[52]

Specifically, I perform the following steps $b = 1, .., B$ times:

1. Randomly split the set of N countries into two sets of $N/2$ countries, $C^{(b)}$ and $C^{(-b)}$.
2. Construct the sets of trade flows, $X_{ni}^{jk(b)}$ and $X_{ni}^{jk(-b)}$, and independent variables, $\tilde{\boldsymbol{z}}_{ni}^{(b)}$ and $\tilde{\boldsymbol{z}}_{ni}^{(-b)}$, for the subsets $C^{(b)}$ and $C^{(-b)}$.
3. Estimate $\hat{\boldsymbol{\gamma}}^{(b)}$ based on the dataset $\{X_{ni}^{jk(b)}, \tilde{\boldsymbol{z}}_{ni}^{(b)}\}$.
4. Predict trade flows for $C^{(-b)}$ using (13) and (14) based on $\hat{\boldsymbol{\gamma}}^{(b)}$, given $\tilde{\boldsymbol{z}}_{ni}^{(-b)}$.

This procedure yields B sets of out-of-sample predicted trade flows. The average loss over the evaluation samples measures the out-of-sample performance of the estimator for a representative set of countries. Because the sampling procedure is consistent with the structure of trade flows in the model and errors that are clustered by country, the distribution of loss function values across bootstrap iterations can be used to test for the best out-of-sample predictor.[53]

Table 8 presents the results of 40,000 bootstrap iterations. The product-level estimators far outperform the sector-level estimators in all regards. As expected, allowing for more heterogeneity improves in-sample fit. However, the out-of-sample fit of single-sector pooled PPML dominates the others on average in RMSE. For Poisson loss, multi-sector pooled PPML slightly outperforms single-sector pooled PPML, with a difference in the sixth digit,

[52] This procedure is analogous to the bootstrap procedure of Snijders and Borgatti (1999).

[53] While it is well beyond the scope of this paper to derive the asymptotic properties of this evaluation measure, it can be seen as an application of the Reality Check Bootstrap of White (2003), which produces asymptotically valid p-values. It can also be seen as a bootstrapped form of cross-validation (Stone, 1978).

and the latter outperforms all others. The bottom two rows of Table 8 show the share of bootstrap iterations for which each estimator is the best out-of-sample predictor. For Poisson loss, product-by-product PPML dominates the pooled estimators less than 1% of time. For RMSE loss, none of the product-level estimators are clearly dominated, though pooled PPML is most often the best predictor. Because the loss values for single-sector and multi-sector pooled PPML are highly correlated, product-by-product PPML is the third best predictor in RMSE more than 70% of the time.

These results indicate that overfitting is a concern for the product-by-product estimator, consistent with the findings of Baltagi (2008) for a range of heterogeneous panel estimators. While the results do not clearly reject the product-by-product estimator, there is a strong argument to prefer the pooled estimators, unless heterogeneity is a primary concern, due to their relative parsimony, pragmatic usefulness, and other desirable properties. Therefore, I propose the following approach: Pool to the level of aggregation most relevant to the research question, while acknowledging the correct interpretation of the pooled coefficient estimates as indexes of coefficients that likely vary at the product level.

5.4 Heteroskedasticity

The baseline results suggest that bias due to heteroskedasticity is less of a concern than previously thought once we control for PLCA. However, it is still useful to evaluate the properties of the estimation errors to help in selecting among these estimators. I follow Manning and Mullahy (2001) in estimating the relationship between the squared residual and a power function of the model predicted values – referred to by Head and Mayer (2014) as a "MaMu" test – which is given by

$$(X_{ni}^k - \hat{X}_{ni}^k)^2 = \lambda_0(\hat{X}_{ni}^k)^{\lambda_1}.$$

Manning and Mullahy (2001) suggest estimating this relationship by OLS in its log-linear form, but Santos Silva and Tenreyro (2006) point out that this is only appropriate under the same conditions for which log-LS is consistent. Therefore, in each case, I estimate this relationship using the estimator that produced the residuals.[54]

The results of the MaMu tests are presented in Table 9. For each of the estimators, the assumed value of λ_1 can be rejected. The hypotheses that $\lambda_1 = 0$ and $\lambda_1 = 2$ can be easily rejected in all cases. The preponderance of the evidence suggests that λ_1 lies between 1 and

[54]Specifically, I use log-linear OLS with non-robust standard errors in the FE and Structural log-LS cases and the same PML estimator with multi-way cluster-robust standard errors for the PML cases. Because the log-LS estimators are valid when $\lambda_1 = 2$, the MaMu tests based on them are valid tests of this hypothesis. The PML-based MaMu tests provide asymptotically valid estimates of λ_1, given robust standard errors.

Table 9: MaMu Tests

	FE Log LS	Structural Log LS	Structural GPML	Structural PPML	Structural NLS
H_0	$\lambda_1 = 2$	$\lambda_1 = 2$	$\lambda_1 = 2$	$\lambda_1 = 1$	$\lambda_1 = 0$
Aggregate	1.78	1.81	1.66	1.42	1.27
	(0.009)	(0.009)	(0.044)	(0.050)	(0.108)
Product-Level	1.93	1.30	1.13	1.43	0.92
	(0.001)	(0.001)	(0.010)	(0.062)	(0.054)

Notes: Estimated values of $\hat{\lambda}_1$ and standard errors (in parentheses) are reported. Standard errors for log least squares specifications are non-robust. All other standard errors are robust to multi-way clustering by both importer and exporter.

2, and the MaMu tests based on product-level estimators suggest that λ_1 is likely closer to 1 than to 2, including the failure to reject $\lambda_1 = 1$ based on product-level structural NLS. Interestingly, however, product-level FE log-LS finds λ_1 very close to 2 in magnitude, which may explain why the heteroskedasticity bias appears much less significant in product-level data. Together with the other desirable properties of PPML, these results suggest that product-level PPML be the preferred gravity estimator, though product-level GPML and log-LS may be considered for robustness.

5.5 The Trade Impact of Changes in Border Costs

In addition to bias in parameter estimates, I also consider how ignoring PLCA biases the predicted effects of changes in trade costs. In the spirit of Anderson and van Wincoop (2003), Waugh (2010), and others, I consider a simple counterfactual experiment: the effect of eliminating border costs.

To this end, I use the Modular Trade Impact (MTI), so termed by Head and Mayer (2014) because it relies upon the modular nature of structural gravity models.[55] The MTI allows the MR terms to adjust to changes in trade costs but holds output and expenditure constant. It is not a full general equilibrium trade impact (GETI) measure because factor prices are held constant. The MTI is very useful for multi-sector and/or product-level gravity models because, as my model makes clear, it is not necessary to specify the form of demand across products, details of factor markets, or the sources of comparative advantage to estimate the parameters of a trade cost function.[56] Unlike the MTI, computing the GETI would involve specifying and parameterizing each of these as well as separately identifying the set of trade cost elasticities. Further, Head and Mayer (2014) and Anderson and van Wincoop (2003)

[55] See Anderson (2011) for a detailed analysis of this property.

[56] Behar and Nelson (2014) make the related point that, in their single-sector model with CES demand, the MTI does not require an estimate of the elasticity of substitution across products.

Table 10: Median MTI of Elimination of Border Costs

	Sector-Level		Pooled PL		Product-
	Single-Sector	Multi-Sector	Single-Sector	Multi-Sector	by-Product
FE Log-LS	62.8	82.4	55.1	59.1	44.0
Struc. Log-LS	59.1	48.3	42.0	40.0	29.3
Struc. GPML	119.1	89.3	58.3	46.4	37.4
PPML	130.8	158.1	59.0	68.5	67.2
Struc. NLS	130.3	144.6	68.6	69.8	59.2

Notes: Percentage changes in trade flows are reported – i.e., $100 \times (\text{MTI}_{ni} - 1)$.

report that the differences between an MTI and GETI tend to be relatively small in practice, especially for small changes in trade costs.

For the product-level gravity model, the MTI is defined as follows:

$$\text{MTI}_{ni} \equiv \frac{X_{ni}^1}{X_{ni}^0} = \frac{\sum_j \sum_k \frac{Y_i^{jk}}{\Psi_i^{jk1}} \frac{X_n^{jk}}{\Phi_n^{jk1}} e^{\boldsymbol{z}_{ni}^{1\prime} \hat{\boldsymbol{\beta}}^{jk}}}{\sum_j \sum_k \frac{Y_i^{jk}}{\Psi_i^{jk0}} \frac{X_n^{jk}}{\Phi_n^{jk0}} e^{\boldsymbol{z}_{ni}^{0\prime} \hat{\boldsymbol{\beta}}^{jk}}},$$

where superscript zeros denote baseline values and ones denote values after changes in trade costs. For the single-sector, sector-level specifications, this expression reduces to

$$\text{MTI}_{ni} \equiv \frac{X_{ni}^1}{X_{ni}^0} = e^{(\boldsymbol{z}_{ni}^1 - \boldsymbol{z}_{ni}^0)' \hat{\bar{\boldsymbol{\beta}}}} \frac{\Psi_i^0}{\Psi_i^1} \frac{\Phi_n^0}{\Phi_n^1}.$$

The values of Φ_n^{k1} and Ψ_i^{k1} are calculated according to (7). Because data on X_{nn}^k are unavailable, computations based on the product-level model use the predicted values of Y_i^k and X_n^k, calculated using (13) and (14) and the baseline estimates of $\boldsymbol{\gamma}$ and $\bar{\boldsymbol{\delta}}$.

Table 10 presents the median MTI of setting all border costs to zero, using estimates from the baseline specifications. As with the coefficient estimates, the MTIs based on sector-level specifications are much more heterogeneous than those based on product-level specifications. Also, the product-level MTIs tend to be much smaller. The former result is driven by the heterogeneity in border costs estimated by the sector-level specifications. The latter depends mostly on the effects of PLCA. Because the effects of trade costs are ameliorated by PLCA, the product-level models predict a smaller MTI for a given partial trade impact than the sector-level models. This effect can be offset if border effects are estimated to be much smaller by the sector-level specifications. This is true for the log-LS and GPML estimators, but the latter effect is never large enough to offset the former.

This simple exercise demonstrates both the ease and importance of accounting for PLCA in predicting the effects of changes in trade costs, even for changes that are uniform across products. However, it is not clear whether MTIs based on the heterogeneous or homogeneous specifications should be preferred. There is no clear pattern in the differences between the MTIs based on homogeneous and heterogeneous specifications, except that the dispersion across estimators is greater for the more heterogenous specifications. Because the more heterogenous estimates tend to be noisier and have worse out-of-sample predictive power, the same approach that I proposed for estimation – pool to the level of aggregation relevant to the research question – is reasonable for prediction, as well.

6 Panel Estimation

The baseline estimates demonstrate that failing to control for PLCA causes substantial bias in sector-level gravity estimates. To show this as clearly as possible, the baseline specification was kept simple. In this section, I include other common gravity variables that proxy for trade policy and cultural and political ties between countries, specifically whether country pairs share a common language, historical colonial ties, or a free trade agreement (FTA). These relationships are likely endogenous, as they are more likely to form between countries that trade intensively for other reasons. Controlling for PLCA, which is one important reason, will lessen but may not eliminate the endogeneity problem. To address this form of endogeneity, I follow the approach of Baier and Bergstrand (2007) and others in using panel data and including country-pair FEs to control for any time-invariant unobserved factors that determine bilateral trade flows.

As shown in Section 3, sector-level panel estimation is also biased when failing to control for PLCA. To evaluate the severity of the bias in practice, I perform sector-level and pooled product-level estimations. In the spirit of Baier and Bergstrand (2007), I use trade flows observed in five-year intervals and include FTA membership and its five-year lag to capture the effects of FTAs that "phase in" over time. Because geographic and historical relationship variables are not time-varying, they are absorbed by the county-pair FEs. Therefore, similar to Bergstrand et al. (2015), I interact these variables with a time trend to estimate changes in their effects over time.[57] For brevity, I focus on FE Log-LS and PPML estimators.

The estimates are shown in Table 11. As in the baseline estimates, we see substantial differences when controlling for PLCA. Product-level estimates of the FTA effect are smaller by one quarter and one half for PPML and FE-OLS when estimated in isolation. This phenomenon is much weaker when allowing for changes in the effects of other common

[57]Note that all specifications control for changing effects of national borders over time. Specifically, the $\text{INTER}_{ni,t}$ term estimated by Bergstrand et al. (2015), who assume changes in border costs over time are equal across countries, is subsumed in the country- and product-specific border costs, δ_n^k, which are allowed to vary over time in the panel estimation.

Table 11: Panel Estimation Results

(a) Sector-Level Estimations

	Log LS			PPML		
	(1)	(2)	(3)	(4)	(5)	(6)
FTA	0.171	0.142	0.129	0.205	0.244	0.236
	(0.128)	(0.126)	(0.124)	(0.063)	(0.053)	(0.051)
FTA_{t-5}	0.341	0.316	0.308	0.172	0.206	0.191
	(0.193)	(0.192)	(0.186)	(0.047)	(0.059)	(0.060)
Distance × t		−0.011	−0.010		−0.002	−0.002
		(0.015)	(0.014)		(0.010)	(0.009)
Sh. Border × t		0.000	0.008		−0.048	−0.050
		(0.032)	(0.033)		(0.018)	(0.021)
Language × t			0.008			0.005
			(0.023)			(0.019)
Colony × t			−0.121			−0.037
			(0.023)			(0.015)
Total FTA	0.512	0.458	0.437	0.377	0.450	0.427
	(0.157)	(0.161)	(0.158)	(0.079)	(0.070)	(0.070)
Exporter-Time FEs	Y	Y	Y	Y	Y	Y
Importer-Time FEs	Y	Y	Y	Y	Y	Y
Pair FEs	Y	Y	Y	Y	Y	Y
Observations	34,425	34,425	34,425	59,856	59,856	67,485

(b) Product-Level Estimations

	Log LS			PPML		
	(1)	(2)	(3)	(4)	(5)	(6)
FTA	0.112	0.163	0.151	0.163	0.228	0.221
	(0.039)	(0.041)	(0.036)	(0.056)	(0.039)	(0.037)
FTA_{t-5}	0.161	0.218	0.202	0.127	0.207	0.188
	(0.038)	(0.045)	(0.041)	(0.049)	(0.053)	(0.050)
Distance × t		0.031	0.030		0.019	0.019
		(0.012)	(0.012)		(0.009)	(0.008)
Sh. Border × t		−0.008	0.012		−0.054	−0.036
		(0.015)	(0.011)		(0.016)	(0.018)
Language × t			−0.040			−0.032
			(0.009)			(0.016)
Colony × t			−0.088			−0.030
			(0.014)			(0.014)
Total FTA	0.272	0.381	0.353	0.291	0.435	0.408
	(0.060)	(0.055)	(0.053)	(0.082)	(0.060)	(0.059)
Exporter-Time FEs	Y	Y	Y	Y	Y	Y
Importer-Time FEs	Y	Y	Y	Y	Y	Y
Pair FEs	Y	Y	Y	Y	Y	Y
Observations	2,188,041	2,188,041	2,188,041	38,367,696	38,367,696	38,367,696

Notes: Standard errors (in parentheses) are robust to multi-way clustering by importer, exporter, and year.

gravity variables. Instead, other differences emerge. In the product-level estimates, the distance elasticity falls in absolute value by an economically and statistically significant 2-3% per period, and the effect of a common language falls significantly, consistent with the "flattening world" hypothesis (Friedman, 2007). These results are consistent with a trend of strengthening bilateral comparative advantage between nearby countries, which tend to sign FTAs and are more likely to share a language, perhaps due to shifting investment from comparative disadvantage to comparative advantage products as trade barriers fall. These results clearly demonstrate that, even when it is possible to control for unobserved country-pair effects, it is still important to control for changes in PLCA over time.

7 Conclusion

This paper shows both theoretically and empirically that aggregate estimation of models with heterogeneous multiplicative two-way fixed effects leads to biased parameter estimates and misleading predictions. I propose a set of disaggregate PML estimators that are straightforward to implement. The ideal coefficient index and pooled estimators that I develop also obviate reporting, efficiency, and multiple inference concerns that may have dissuaded researchers from pursuing disaggregate estimation in the past. I apply these estimators to data on product-level bilateral trade flows, where the interacted effects embody product-level comparative advantage, and find that the bias in aggregate estimates is significant. Based on the theoretical and empirical results, I argue that estimation of models with unobserved multiplicative effects should always make use of the most disaggregated data available. In particular, the desirable properties of the pooled PPML estimator make it a strong candidate to be the workhorse estimator for such models.

The structural gravity PML estimators that I implement are efficient if the model errors obey the implied distributional assumptions. However, in general an optimal GMM estimator based on the structural gravity moment conditions may improve efficiency. Another potentially fruitful avenue for future research is to explore possible efficiency gains by explicitly modelling the heterogeneity of the fixed effects, for example by imposing a factor structure. In the context of trade flows, Lind and Ramondo (2018) and Hanson et al. (2015) make substantial progress in modelling countries' patterns of comparative advantage.

References

Anderson, James E. (2011), "The Gravity Model," *Annual Review of Economics*, 3 (1), 133–160.

Anderson, James E. and J. Peter Neary (2003), "The Mercantilist Index of Trade Policy," *International Economic Review*, 44 (2), 627–649.

Anderson, James E. and Eric van Wincoop (2003), "Gravity with Gravitas: A Solution to the Border Puzzle," *The American Economic Review*, 93 (1), 170–192.

——— (2004), "Trade Costs," *Journal of Economic Literature*, 42 (3), 691–751.

Anderson, James E. and Yoto V. Yotov (2010a), "Specialization: Pro- and Anti-Globalizing, 1990-2002," NBER Working Paper No. 16301.

——— (2010b), "The Changing Incidence of Geography," *American Economic Review*, 100 (5), 2157–86.

Anderson, Michael L. (2008), "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," 103 (484), 1481–1495.

Anderson, Michael and Jeremy Magruder (2017), "Split-Sample Strategies for Avoiding False Discoveries," NBER Working Paper No. 23544.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare (2015), "The Elusive Pro-Competitive Effects of Trade," NBER Working Paper No. 21370.

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare (2012), "New Trade Models, Same Old Gains?" *American Economic Review*, 102 (1), 94–130.

Arvis, Jean-François and Ben Shepherd (2013), "The Poisson Quasi-maximum Likelihood Estimator: A Solution to the 'adding Up' Problem in Gravity Models," *Applied Economics Letters*, 20 (6), 515–519.

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

Baltagi, Badi H. (2005), *Econometric Analysis of Panel Data*, Hoboken, Wiley, 3rd edition.

——— (2008), "Forecasting with Panel Data," *Journal of Forecasting*, 27 (2), 153–173.

Baltagi, Badi H., James M. Griffin, and Weiwen Xiong (2000), "To Pool or Not to Pool: Homogeneous Versus Heterogeneous Estimators Applied to Cigarette Demand," *Review of Economics and Statistics*, 82 (1), 117–126.

Behar, Alberto and Benjamin D. Nelson (2014), "Trade Flows, Multilateral Resistance, and Firm Heterogeneity," *The Review of Economics and Statistics*, 96 (3), 538–549.

Benjamini, Yoav and Yosef Hochberg (1995), "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society: Series B (Methodological)*, 57 (1), 289–300.

Bergstrand, Jeffrey H., Mario Larch, and Yoto V. Yotov (2015), "Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation," *European Economic Review*, 78, 307–327.

Caliendo, Lorenzo and Fernando Parro (2015), "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 82 (1), 1–44.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2011), "Robust Inference With Multiway Clustering," *Journal of Business & Economic Statistics*, 29 (2), 238–249.

Cameron, A. Colin and Douglas L. Miller (2015), "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 50 (2), 317–372.

Charbonneau, Karyne (2012), "Multiple Fixed Effects in Nonlinear Panel Data Models," Working Paper.

Costinot, Arnaud and Andrés Rodríguez-Clare (2014), "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in Elhanan Helpman Gita Gopinath and Kenneth Rogoff (Eds.), *Handbook of International Economics*, Vol. 4, Elsevier, Chap. 4, 197–261.

Diebold, Francis X. and Roberto S. Mariano (1995), "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13 (3), 253–263.

Disdier, Anne-Célia and Keith Head (2008), "The Puzzling Persistence of the Distance Effect on Bilateral Trade," *The Review of Economics and Statistics*, 90 (1), 37–48.

Eaton, Jonathan and Samuel Kortum (2002), "Technology, Geography, and Trade," *Econometrica*, 70 (5), 1741–1779.

Eaton, Jonathan and Akiko Tamura (1994), "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns," *Journal of the Japanese and International Economies*, 8 (4), 478 – 510.

Egger, Peter H. and Sergey Nigai (2015), "Structural Gravity with Dummies Only: Constrained ANOVA-Type Estimation of Gravity Models," *Journal of International Economics*, 97 (1), 86–99.

Fafchamps, Marcel and Julien Labonne (2017), "Using Split Samples to Improve Inference on Causal Effects," *Political Analysis*, 25 (4), 465–482.

Fally, Thibault (2015), "Structural Gravity and Fixed Effects," *Journal of International Economics*, 97 (1), 76–85.

Feenstra, Robert C., Robert E. Lipsey, Haiyan Deng, Alyson C. Ma, and Hengyong Mo (2005), "World Trade Flows: 1962-2000," NBER Working Paper No. 11040.

Fernández-Val, Iván and Martin Weidner (2016), "Individual and Time Effects in Nonlinear Panel Models with Large N, T," *Journal of Econometrics*, 192 (1), 291–312.

French, Scott (2015), "Welfare in the Eaton-Kortum Model of International Trade," Working Paper.

——— (2016), "The Composition of Trade Flows and the Aggregate Effects of Trade Barriers," *Journal of International Economics*, 98, 114 – 137.

——— (2017), "Revealed Comparative Advantage: What Is It Good For?" *Journal of International Economics*, 106, 83–103.

Friedman, Thomas L. (2007), *The World Is Flat 3.0: A Brief History of the Twenty-First Century*, Picador.

Gaure, Simen (2013), "OLS with Multiple High Dimensional Category Variables," *Computational Statistics & Data Analysis*, 66, 8–18.

Gourieroux, C., A. Monfort, and A. Trognon (1984), "Pseudo Maximum Likelihood Methods: Theory," *Econometrica*, 52 (3), pp. 681–700.

Guimarães, Paulo and Pedro Portugal (2010), "A Simple Geasible Procedure to Fit Models with High-Dimensional Fixed Effects," *Stata Journal*, 10 (4), 628–649.

Hallak, Juan Carlos (2006), "Product Quality and the Direction of Trade," *Journal of International Economics*, 68 (1), 238 – 265.

——— (2010), "A Product-Quality View of the Linder Hypothesis," *The Review of Economics and Statistics*, 92 (3), 453–466.

Hanson, Gordon H, Nelson Lind, and Marc-Andreas Muendler (2015), "The Dynamics of Comparative Advantage," NBER Working Paper No. 21753.

Head, Keith and Thierry Mayer (2014), "Gravity Equations: Workhorse, Toolkit, and Cookbook," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff (Eds.), *Handbook of International Economics*, Vol. 4, Elsevier, Chap. 3, 131–195.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (2008), "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123 (2), 441–487.

Hillberry, Russell H. (2002), "Aggregation Bias, Compositional Change, and the Border Effect," *The Canadian Journal of Economics*, 35 (3), 517–530.

Hillberry, Russell and David Hummels (2002), "Explaining Home Bias in Consumption: The Role of Intermediate Input Trade," NBER Working Paper No. 9020.

Jochmans, Koen (2017), "Two-Way Models for Gravity," *The Review of Economics and Statistics*, 99 (3), 478–485.

Kehoe, Timothy J., Pau S. Pujolàs, and Jack Rossbach (2017), "Quantitative Trade Models: Developments and Challenges," *Annual Review of Economics*, 9 (1), 295–325.

Levchenko, Andrei A. and Jing Zhang (2016), "The Evolution of Comparative Advantage: Measurement and Welfare Implications," *Journal of Monetary Economics*, 78, 96–111.

Lewbel, Arthur (1992), "Aggregation with Log-Linear Models," *The Review of Economic Studies*, 59 (3), 635–642.

Lind, Nelson and Natalia Ramondo (2018), "Trade with Correlation," NBER Working Paper No. 24380.

Linder, Staffan Burenstam (1961), *An Essay on Trade and Transformation*, Stockholm, Almqvist & Wiksell.

Manning, Willard G and John Mullahy (2001), "Estimating Log Models: To Transform or Not to Transform?" *Journal of Health Economics*, 20 (4), 461–494.

Mayer, Thierry and Soledad Zignago (2011), "Notes on CEPIIs Distances Measures (GeoDist)," CEPII Working Paper No. 2011-25.

Melitz, Marc J. and Gianmarco I. P. Ottaviano (2008), "Market Size, Trade, and Productivity," *The Review of Economic Studies*, 75 (1), 295–316.

Neuhaus, John M. and Nicholas P. Jewell (1993), "A Geometric Approach to Assess Bias Due to Omitted Covariates in Generalized Linear Models," *Biometrika*, 80 (4), 807–815.

Poissonnier, Aurélien (2019), "Iterative Solutions for Structural Gravity Models in Panels," *International Economics*, 157, 55–67.

Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez (2016), "Trade, Domestic Frictions, and Scale Effects," *American Economic Review*, 106 (10), 3159–84.

Santos Silva, J. M. C. and Silvana Tenreyro (2006), "The Log of Gravity," *The Review of Economics and Statistics*, 88 (4), pp. 641–658.

Snijders, Tom AB and Stephen P Borgatti (1999), "Non-Parametric Standard Errors and Tests for Network Statistics," *Connections*, 22 (2), 161–170.

Stoker, Thomas M. (1993), "Empirical Approaches to the Problem of Aggregation Over Individuals," *Journal of Economic Literature*, 31 (4), 1827–1874.

Stone, M. (1978), "Cross-Validation: A Review," *Statistics: A Journal of Theoretical and Applied Statistics*, 9 (1), 127–139.

van Garderen, Kees Jan, Kevin Lee, and M. Hashem Pesaran (2000), "Cross-Sectional Aggregation of Non-Linear Models," *Journal of Econometrics*, 95 (2), 285–331.

Waugh, Michael E. (2010), "International Trade and Income Differences," *American Economic Review*, 100 (5), 2093–2124.

White, Halbert (2003), "A Reality Check for Data Snooping," *Econometrica*, 68 (5), 1097–1126.

Yi, Kei-Mu (2003), "Can Vertical Specialization Explain the Growth of World Trade?" *Journal of Political Economy*, 111 (1), 52–102.

A Additional Tables

Table A1: Countries and Sources of Manufacturing Output Data

Country	Source	Country	Source	Country	Source
Albania	INDSTAT	Georgia	INDSTAT	Panama	INDSTAT(int.)
Argentina	WDI	Germany	STAN	Papua New Guinea	WDI
Australia	INDSTAT	Ghana	INDSTAT	Peru	INDSTAT
Austria	STAN	Greece	STAN	Philippines	INDSTAT
Azerbaijan	INDSTAT	Grenada	INDSTAT	Poland	STAN
Bahamas	WDI	Guatemala	WDI	Portugal	STAN
Bangladesh	WDI	Honduras	WDI	Qatar	INDSTAT
Barbados	WDI	Hungary	STAN	Rep. of Korea	STAN
Belarus	WDI	Iceland	STAN	Rep. of Moldova	INDSTAT
Belize	WDI	India	INDSTAT	Romania	INDSTAT
Benin	WDI	Indonesia	INDSTAT	Russian Federation	INDSTAT
Bolivia	WDI	Iran	INDSTAT	Rwanda	WDI
Botswana	INDSTAT	Ireland	STAN	Saint Kitts and Nevis	INDSTAT
Brazil	INDSTAT	Israel	STAN	St. Lucia	WDI
Brunei Darussalam	WDI	Italy	STAN	Samoa	WDI
Bulgaria	INDSTAT	Jamaica	WDI	Sao Tome and Princ.	WDI
Burkina Faso	WDI	Japan	STAN	Saudi Arabia	INDSTAT(int.)
Burundi	WDI	Jordan	INDSTAT	Senegal	WDI
Cambodia	WDI	Kazakhstan	INDSTAT	Slovakia	STAN
Cameroon	WDI	Kenya	INDSTAT	Slovenia	STAN
Canada	STAN	Kyrgyzstan	INDSTAT	South Africa	INDSTAT
Cape Verde	WDI	Latvia	INDSTAT	Spain	STAN
Central African Rep.	WDI	Lebanon	WDI	Sri Lanka	INDSTAT(int.)
Chile	INDSTAT	Lithuania	INDSTAT	Sudan	WDI
China	INDSTAT	Madagascar	INDSTAT	Swaziland	WDI
Colombia	INDSTAT	Malawi	WDI	Sweden	STAN
Costa Rica	WDI	Malaysia	INDSTAT	Switzerland	STAN
Cte d'Ivoire	WDI	Maldives	WDI	Syria	INDSTAT
Croatia	WDI	Malta	INDSTAT	TFYR of Macedonia	INDSTAT
Cuba	WDI	Mauritania	WDI	Thailand	INDSTAT(int.)
Cyprus	INDSTAT	Mauritius	INDSTAT	Togo	WDI
Czech Rep.	STAN	Mexico	STAN	Trinidad and Tobago	INDSTAT
Denmark	STAN	Morocco	INDSTAT	Tunisia	INDSTAT
Dominica	INDSTAT	Mozambique	WDI	Turkey	INDSTAT
Dominican Rep.	WDI	Namibia	WDI	USA	STAN
Ecuador	INDSTAT	Nepal	WDI	Uganda	WDI
Eritrea	INDSTAT	Netherlands	STAN	Ukraine	INDSTAT
Estonia	STAN	New Zealand	STAN	United Kingdom	STAN
Ethiopia	INDSTAT	Nicaragua	WDI	U. Rep. of Tanzania	INDSTAT
Fiji	INDSTAT	Niger	WDI	Uruguay	INDSTAT
Finland	STAN	Nigeria	INDSTAT	Venezuela	WDI
France	STAN	Norway	STAN	Viet Nam	INDSTAT
Gabon	WDI	Pakistan	INDSTAT(int.)	Zambia	WDI
Gambia	WDI				

Notes: INDSTAT(int.) indicates that output data were interpolated based on INDSTAT data for years before and after 2003.

Table A2: ISIC Rev. 3 Industries

ISIC	Industry Description	HS-6 Codes	Countries	Trade Share
15A	Food, beverages, and tobacco	428	76	6.6%
17	Textiles	541	63	3.3%
18	Wearing apparel; dressing and dyeing of fur	241	48	2.9%
19	Leather, leather products, and footwear	67	57	1.4%
20	Wood products, except furniture	69	75	1.3%
21	Paper and paper products	120	75	2.4%
22	Publishing, printing, reproduction of recorded media	36	78	0.9%
23	Coke, refined petroleum products, nuclear fuel	20	55	2.7%
24	Chemicals and chemical products	879	66	11.8%
25	Rubber and plastics products	121	76	3.0%
26	Non-metallic mineral products	170	79	1.5%
27	Basic metals	359	58	5.4%
28	Fabricated metal products, except mach. and equip.	221	74	2.7%
29	Other machinery and equipment	528	61	10.6%
30	Office, accounting and computing machinery	37	33	5.4%
31	Other electrical machinery and apparatus	134	62	4.7%
32	Radio, television, and communication equipment	101	48	8.5%
33	Medical, precision instruments, watches and clocks	212	53	3.9%
34	Motor vehicles, trailers and semi-trailers	54	56	13.6%
35	Other transport equipment	81	58	4.1%
36	Furniture, other manufacturing	189	64	3.2%

Notes: Column "Countries" lists number of sample countries with output data available for each ISIC industry.

Table A3: Multi-Sector Estimation Results: Sector-Level Industy-by-Industry Estimations

	Food/ Tobacco	Textiles	Apparel	Leather	Wood	Paper	Printing	Fuel	Chemical	Rubber/ Plastics	Minerals	Basic Metal	Fabr. Metal	Other Mach.	Comp. Mach.	Electrical	Comm. Equip.	Medical	Vehicles	Other Transp.	Furniture
Fixed Effects Log LS																					
mean($\hat{\delta}_n$)	−1.62	−0.08	−0.34	−0.26	−1.43	−0.24	−2.19	−0.96	−0.44	−0.48	−1.53	−0.01	−1.10	−0.34	−1.32	−0.46	−0.46	−1.10	−0.93	−1.43	−1.70
Distance	−1.73	−1.97	−2.05	−1.78	−2.00	−2.42	−2.11	−2.35	−2.06	−2.18	−2.10	−2.02	−2.12	−1.85	−1.87	−1.97	−1.70	−1.48	−1.85	−1.39	−1.88
	(0.07)	(0.09)	(0.10)	(0.11)	(0.08)	(0.09)	(0.09)	(0.14)	(0.08)	(0.08)	(0.07)	(0.11)	(0.07)	(0.07)	(0.09)	(0.08)	(0.09)	(0.07)	(0.08)	(0.07)	(0.08)
Shared Border	1.32	0.97	0.82	1.32	0.81	0.83	1.07	1.32	0.87	1.08	1.43	0.80	1.11	0.92	0.55	1.00	0.85	0.95	1.06	1.14	0.96
	(0.19)	(0.19)	(0.21)	(0.21)	(0.21)	(0.22)	(0.18)	(0.24)	(0.18)	(0.18)	(0.20)	(0.18)	(0.17)	(0.18)	(0.19)	(0.22)	(0.19)	(0.18)	(0.20)	(0.19)	(0.16)
$\mathrm{H}_0 : \gamma^j_{\mathrm{Agg}} = \gamma^j_{\mathrm{Agg,PPML}}$	53.36	38.59	19.63	36.62	26.22	85.39	48.74	27.29	91.19	55.28	67.65	31.18	57.61	82.60	52.90	46.57	53.80	58.03	23.29	62.41	50.73
Structural Log LS																					
mean($\hat{\delta}_n$)	−1.59	0.03	−0.43	−0.32	−1.74	−0.71	−2.10	−0.76	−0.95	−0.78	−1.66	−0.11	−1.34	−0.62	−2.16	−0.58	−0.53	−1.44	−0.97	−1.06	−2.60
Distance	−1.83	−2.19	−1.95	−1.82	−2.10	−2.25	−2.19	−2.55	−1.94	−2.11	−2.18	−1.84	−2.21	−1.81	−1.69	−2.08	−1.90	−1.62	−1.89	−1.58	−1.93
	(0.12)	(0.22)	(0.21)	(0.18)	(0.13)	(0.14)	(0.17)	(0.14)	(0.11)	(0.11)	(0.10)	(0.13)	(0.11)	(0.10)	(0.12)	(0.13)	(0.13)	(0.12)	(0.14)	(0.20)	(0.12)
Shared Border	1.03	0.43	1.00	1.08	0.04	0.49	0.97	1.12	0.35	0.75	0.88	1.20	0.39	0.54	−0.15	0.46	0.16	0.12	0.85	1.14	−0.39
	(0.38)	(0.69)	(0.67)	(0.53)	(0.51)	(0.37)	(0.53)	(0.32)	(0.36)	(0.29)	(0.29)	(0.32)	(0.33)	(0.35)	(0.48)	(0.46)	(0.55)	(0.59)	(0.40)	(0.66)	(0.44)
$\mathrm{H}_0 : \gamma^j_{\mathrm{Agg}} = \gamma^j_{\mathrm{Agg,PPML}}$	94.74	97.43	32.42	43.64	39.71	114.75	99.12	85.77	133.98	88.53	184.19	82.53	143.97	163.13	19.66	115.92	71.67	104.16	42.06	106.36	51.82
Structural GPML																					
mean($\hat{\delta}_n$)	−2.63	−1.17	−1.96	−1.79	−2.80	−1.88	−3.47	−2.58	−1.89	−1.84	−2.49	−1.64	−2.17	−1.42	−2.45	−1.61	−1.45	−1.90	−2.13	−2.64	−3.57
Distance	−1.57	−1.66	−1.61	−1.51	−1.75	−1.65	−1.74	−1.85	−1.52	−1.63	−2.14	−1.68	−1.88	−1.46	−1.44	−1.60	−1.37	−1.30	−1.33	−0.66	−1.50
	(0.09)	(0.08)	(0.10)	(0.17)	(0.10)	(0.11)	(0.08)	(0.11)	(0.08)	(0.17)	(0.09)	(0.10)	(0.10)	(0.13)	(0.09)	(0.08)	(0.11)	(0.10)	(0.10)	(0.18)	(0.07)
Shared Border	0.28	0.16	−0.03	0.10	−0.36	0.56	0.26	−0.08	0.26	0.54	−0.30	−0.51	0.12	0.38	0.10	0.39	0.37	0.36	0.77	1.74	−0.38
	(0.29)	(0.31)	(0.33)	(0.29)	(0.28)	(0.23)	(0.20)	(0.43)	(0.21)	(0.36)	(0.24)	(0.37)	(0.26)	(0.31)	(0.25)	(0.20)	(0.20)	(0.27)	(0.22)	(0.44)	(0.19)
$\mathrm{H}_0 : \gamma^j_{\mathrm{Agg}} = \gamma^j_{\mathrm{Agg,PPML}}$	22.56	11.53	4.08	5.80	14.20	12.92	29.71	3.06	33.14	9.24	106.17	21.03	51.11	38.41	13.27	24.09	27.88	35.68	2.29	33.18	38.62
PPML																					
mean($\hat{\delta}_n$)	−2.63	−1.17	−1.96	−1.79	−2.80	−1.88	−3.47	−2.58	−1.89	−1.84	−2.49	−1.64	−2.17	−1.42	−2.45	−1.61	−1.45	−1.90	−2.13	−2.64	−3.57
Distance	−1.10	−1.27	−1.27	−0.93	−1.26	−1.35	−1.14	−1.59	−1.03	−1.27	−1.23	−1.11	−1.20	−0.88	−1.01	−1.11	−0.81	−0.73	−1.12	−0.33	−0.89
	(0.10)	(0.11)	(0.19)	(0.21)	(0.12)	(0.09)	(0.13)	(0.14)	(0.09)	(0.10)	(0.08)	(0.09)	(0.11)	(0.07)	(0.14)	(0.09)	(0.09)	(0.07)	(0.17)	(0.15)	(0.12)
Shared Border	0.64	0.30	0.43	0.28	0.80	0.46	0.70	0.66	0.28	0.61	0.66	0.68	0.65	0.56	0.03	0.46	0.44	0.24	0.66	1.10	0.86
	(0.18)	(0.15)	(0.19)	(0.21)	(0.17)	(0.16)	(0.23)	(0.26)	(0.11)	(0.15)	(0.12)	(0.14)	(0.14)	(0.12)	(0.18)	(0.16)	(0.21)	(0.13)	(0.23)	(0.27)	(0.15)
Structural NLS																					
mean($\hat{\delta}_n$)	−3.67	−1.89	−3.00	−1.38	−3.50	−2.61	−4.75	−3.22	−3.23	−2.31	−3.90	−2.06	−3.14	−2.65	−3.39	−2.49	−2.72	−3.42	−2.30	−4.24	−3.56
Distance	−1.12	−1.31	−0.96	−1.82	−0.88	−1.45	−1.14	−1.12	−1.01	−1.51	−1.20	−1.06	−1.34	−0.82	−1.15	−1.18	−0.71	−0.78	−1.35	−0.13	−1.08
	(0.11)	(0.10)	(0.02)	(0.35)	(0.24)	(0.10)	(0.17)	(0.17)	(0.07)	(0.26)	(0.11)	(0.08)	(0.25)	(0.07)	(0.29)	(0.19)	(0.13)	(0.12)	(0.12)	(0.21)	(0.27)
Shared Border	0.36	0.24	0.73	−0.34	1.43	0.21	0.53	1.12	0.12	0.29	0.57	0.71	0.46	0.65	−0.26	0.56	0.75	0.04	0.50	1.54	0.80
	(0.23)	(0.14)	(0.26)	(0.08)	(0.41)	(0.19)	(0.25)	(0.51)	(0.08)	(0.24)	(0.16)	(0.19)	(0.22)	(0.18)	(0.27)	(0.33)	(0.37)	(0.21)	(0.22)	(0.41)	(0.11)
$\mathrm{H}_0 : \gamma^j_{\mathrm{Agg}} = \gamma^j_{\mathrm{Agg,PPML}}$	1.96	$0.40^{\dagger}$	2.83	$5.33^{\dagger}$	12.29	24.08	90.94	20.16	6.88	$2.64^{\dagger}$	2.88	0.43	$0.55^{\dagger}$	1.45	11.19	97.70	$1.75^{\dagger}$	3.76	$3.52^{\dagger}$	2.91	0.81

Notes: Standard errors (in parentheses) and test statistics are robust to multi-way clustering by both importer and exporter, except where indicated.
† Test statistic is robust to clustering by importer due to non-negative semi-definite estimated multi-way cluster robust covariance matrix.

Table A4: Multi-Sector Estimation Results: Pooled Product-Level Industy-by-Industry Estimations

	Food/ Tobacco	Textiles	Apparel	Leather	Wood	Paper	Printing	Fuel	Chemical	Rubber/ Plastics	Minerals	Basic Metal	Fabr. Metal	Other Mach.	Comp. Mach.	Electrical	Comm. Equip.	Medical	Vehicles	Other Transp.	Furniture
Fixed Effects Log LS																					
mean($\hat{\delta}_n$)	−3.31	−2.19	−2.54	−2.19	−3.47	−1.81	−3.45	−2.48	−2.34	−1.87	−3.83	−2.09	−2.96	−1.69	−2.07	−1.70	−1.39	−1.97	−1.76	−2.56	−2.98
Distance	−1.14	−1.05	−1.14	−1.12	−1.12	−1.50	−1.50	−1.65	−1.11	−1.44	−1.09	−1.11	−1.22	−1.09	−1.40	−1.29	−1.19	−0.95	−1.34	−0.81	−1.09
	(0.06)	(0.07)	(0.11)	(0.08)	(0.08)	(0.07)	(0.07)	(0.08)	(0.05)	(0.06)	(0.05)	(0.07)	(0.06)	(0.05)	(0.07)	(0.06)	(0.06)	(0.04)	(0.07)	(0.05)	(0.06)
Shared Border	1.06	0.62	0.60	0.84	0.99	0.94	0.97	0.94	0.72	0.99	1.04	0.81	0.93	0.64	0.47	0.82	0.69	0.74	0.85	0.83	0.87
	(0.12)	(0.13)	(0.16)	(0.14)	(0.14)	(0.11)	(0.13)	(0.13)	(0.09)	(0.12)	(0.11)	(0.10)	(0.13)	(0.10)	(0.13)	(0.14)	(0.12)	(0.11)	(0.12)	(0.08)	(0.14)
$H_0: \gamma^j_{Agg} = \gamma^j_{Pool}$	63.24	152.48	135.86	110.72	158.18	135.56	100.28	63.57	182.43	153.63	210.93	82.69	190.71	175.41	64.41	94.86	51.33	76.14	71.75	95.73	180.17
$H_0: \gamma^j_{Pool} = \gamma^j_{Pool,PPML}$	10.00	8.60	1.64	3.16	21.83	5.51	8.58	1.06	4.53	4.45	3.81	14.50	1.58	1.64	16.20	2.93	6.65	9.67	3.67	4.22	0.44
Structural Log LS																					
mean($\hat{\delta}_n$)	−2.82	−1.59	−2.29	−2.02	−2.87	−1.31	−3.07	−2.16	−1.81	−1.56	−3.26	−1.41	−2.47	−1.18	−2.21	−1.28	−1.16	−1.77	−1.33	−1.93	−2.64
Distance	−1.39	−1.34	−1.28	−1.27	−1.39	−1.74	−1.59	−1.80	−1.33	−1.55	−1.36	−1.40	−1.42	−1.33	−1.39	−1.45	−1.29	−1.08	−1.52	−1.08	−1.27
	(0.08)	(0.09)	(0.13)	(0.10)	(0.10)	(0.07)	(0.08)	(0.12)	(0.08)	(0.07)	(0.07)	(0.08)	(0.07)	(0.06)	(0.07)	(0.07)	(0.08)	(0.07)	(0.08)	(0.08)	(0.09)
Shared Border	0.91	0.52	0.43	0.58	0.90	0.90	1.22	1.02	0.74	1.09	0.97	0.81	0.91	0.64	0.32	0.91	0.71	0.62	0.93	0.80	0.76
	(0.13)	(0.13)	(0.19)	(0.16)	(0.16)	(0.13)	(0.19)	(0.17)	(0.13)	(0.13)	(0.11)	(0.12)	(0.14)	(0.10)	(0.19)	(0.14)	(0.16)	(0.14)	(0.16)	(0.13)	(0.15)
$H_0: \gamma^j_{Agg} = \gamma^j_{Pool}$	50.62	137.11	130.88	82.73	59.18	35.48	52.62	91.68	63.82	82.57	224.99	55.12	166.95	102.41	16.09	87.83	51.84	47.69	32.59	36.44	56.63
$H_0: \gamma^j_{Pool} = \gamma^j_{Pool,PPML}$	1.49	5.18	1.45	3.41	8.85	11.65	38.18	4.93	23.32	29.24	2.75	4.64	13.38	35.03	15.75	35.24	20.50	30.45	27.38	36.07	4.91
Structural GPML																					
mean($\hat{\delta}_n$)	−3.66	−1.32	−2.81	−2.05	−2.96	−2.46	−3.69	−3.71	−2.85	−2.14	−2.65	−2.91	−2.27	−1.75	−2.44	−1.58	−1.66	−0.54	−1.75	−1.17	−2.14
Distance	−1.04	−1.67	−0.99	−1.50	−1.53	−1.57	−1.57	−1.23	−1.08	−1.43	−1.82	−0.82	−1.82	−2.14	−1.58	−1.33	−1.12	−1.37	−1.65	−1.50	−1.25
	(0.18)	(0.06)	(0.13)	(0.17)	(0.12)	(0.16)	(0.13)	(0.26)	(0.16)	(0.10)	(0.21)	(0.17)	(0.11)	(0.21)	(0.12)	(0.09)	(0.10)	(0.19)	(0.12)	(0.11)	(0.13)
Shared Border	0.90	−0.12	0.74	−0.17	0.37	−0.10	0.45	0.43	0.20	0.67	0.41	0.68	−0.02	−2.02	−0.43	0.86	0.54	1.37	0.03	0.56	1.45
	(0.23)	(0.11)	(0.26)	(0.42)	(0.19)	(0.24)	(0.27)	(0.39)	(0.38)	(0.32)	(0.48)	(0.53)	(0.29)	(0.65)	(0.50)	(0.26)	(0.20)	(0.49)	(0.26)	(0.24)	(0.48)
$H_0: \gamma^j_{Agg} = \gamma^j_{Pool}$	6.89	0.92	20.49	1.05	5.09	11.62	1.73	4.94	13.12	1.57	2.31	21.22	1.33	13.22	1.26	5.38	4.37	7.24	7.46	25.11	12.32
$H_0: \gamma^j_{Pool} = \gamma^j_{Pool,PPML}$	6.19	8.18	4.85	2.41	12.00	11.42	5.47	5.13	0.57	0.28	5.34	17.87	13.84	32.33	9.30	5.14	3.51	29.61	10.21	40.27	5.48
PPML																					
mean($\hat{\delta}_n$)	−2.73	−1.35	−1.91	−2.50	−2.28	−1.65	−4.40	−2.43	−2.60	−2.32	−3.48	−1.31	−3.03	−2.01	−3.14	−2.26	−2.25	−2.80	−2.43	−3.00	−3.18
Distance	−1.50	−1.50	−1.45	−1.14	−1.69	−1.72	−1.19	−1.78	−1.13	−1.37	−1.32	−1.51	−1.28	−0.99	−1.08	−1.19	−0.96	−0.73	−1.14	−0.71	−1.06
	(0.09)	(0.10)	(0.17)	(0.18)	(0.09)	(0.09)	(0.12)	(0.14)	(0.10)	(0.10)	(0.09)	(0.08)	(0.10)	(0.08)	(0.14)	(0.09)	(0.08)	(0.06)	(0.16)	(0.12)	(0.13)
Shared Border	0.69	0.36	0.39	0.42	0.77	0.47	0.73	0.62	0.35	0.60	0.80	0.61	0.65	0.55	−0.02	0.45	0.31	0.42	0.59	0.63	0.75
	(0.18)	(0.15)	(0.20)	(0.19)	(0.14)	(0.15)	(0.21)	(0.26)	(0.13)	(0.16)	(0.15)	(0.10)	(0.13)	(0.14)	(0.19)	(0.17)	(0.19)	(0.15)	(0.22)	(0.19)	(0.17)
$H_0: \gamma^j_{Agg} = \gamma^j_{Pool}$	71.41	54.82	18.81	18.83	68.52	106.46	4.21	22.33	25.47	59.08	32.90	25.85	27.44	20.04	27.93	28.39	5.61	13.61	3.00	7.54	24.88
Structural NLS																					
mean($\hat{\delta}_n$)	−3.61	−1.01	−1.73	−0.18	−2.59	−1.80	−4.59	−3.15	−3.38	−2.05	−3.46	−0.93	−2.57	−2.06	−3.29	−1.70	−2.26	−2.49	−2.71	−1.81	−3.24
Distance	−1.28	−1.63	−1.41	−2.34	−1.59	−1.72	−1.25	−1.26	−0.86	−1.62	−1.34	−1.73	−1.51	−0.90	−1.08	−1.44	−0.89	−0.89	−0.98	−1.06	−0.87
	(0.31)	(0.11)	(0.08)	(0.25)	(0.16)	(0.19)	(0.19)	(0.14)	(0.15)	(0.24)	(0.06)	(0.20)	(0.24)	(0.10)	(0.23)	(0.31)	(0.12)	(0.17)	(0.20)	(0.53)	(0.48)
Shared Border	0.23	0.42	0.75	−0.22	0.65	0.27	0.33	1.11	0.23	0.22	0.77	0.47	0.53	0.76	−0.19	0.39	0.48	0.33	0.69	1.01	1.25
	(0.37)	(0.29)	(0.39)	(0.14)	(0.63)	(0.26)	(0.32)	(0.55)	(0.18)	(0.22)	(0.23)	(0.36)	(0.19)	(0.27)	(0.23)	(0.46)	(0.32)	(0.39)	(0.43)	(0.46)	(0.39)
$H_0: \gamma^j_{Agg} = \gamma^j_{Pool}$	0.79†	15.99	106.64	42.70	35.56	30.64	2.78	2.90	1.26	5.59	17.36	13.62	44.10	9.34	0.57	11.81	3.20	21.60	25.33	3.59	1.81
$H_0: \gamma^j_{Pool} = \gamma^j_{Pool,PPML}$	47.52	6.38	3.40†	88.96	11.95	6.19†	4.79	39.66	17.87	20.99	0.04	1.17	2.38	3.80	6.96†	12.14	0.68	4.19	5.04†	4.38	9.40

Notes: Standard errors (in parentheses) and test statistics are robust to multi-way clustering by both importer and exporter, except where indicated.
† Test statistic is robust to clustering by importer due to non-negative semi-definite estimated multi-way cluster robust covariance matrix.

B Proofs

Proof of Proposition 1. Given Assumption 1, summing (1) over k and taking expectations with respect to the data yields

$$\mathrm{E}[X_{ni}|\boldsymbol{Z},\boldsymbol{\Gamma}] = \sum_k \phi_n^k \psi_i^k f(\boldsymbol{z}_{ni};\boldsymbol{\beta}^k),$$

where $\boldsymbol{\Gamma} = \left(\boldsymbol{\Gamma}^{1\prime},...,\boldsymbol{\Gamma}^{K\prime}\right)'$. Multiplying and dividing by $f(\boldsymbol{z}_{ni};\bar{\boldsymbol{\beta}})$, $\bar{\phi}_n$, and $\bar{\psi}_i$ yields

$$\begin{aligned} \mathrm{E}[X_{ni}|\boldsymbol{Z},\boldsymbol{\Gamma}] &= \bar{\phi}_n\bar{\psi}_i f(\boldsymbol{z}_{ni};\bar{\boldsymbol{\beta}}) \sum_k \frac{\phi_n^k}{\bar{\phi}_n}\frac{\psi_i^k}{\bar{\psi}_i}\frac{f(\boldsymbol{z}_{ni};\boldsymbol{\beta}^k)}{f(\boldsymbol{z}_{ni};\bar{\boldsymbol{\beta}})} \\ &\equiv \bar{\omega}_{ni} T_{ni} \\ &= \mathrm{E}[X_{ni}|\boldsymbol{Z},\bar{\boldsymbol{\Gamma}},T_{ni}]. \end{aligned} \tag{17}$$

Note that $r(\cdot)$ and $\tilde{r}(\cdot)$ are defined such that $\mathrm{E}[r(x,\mathrm{E}[x|\boldsymbol{\Omega}])|\boldsymbol{\Omega}] = 0$ and $\mathrm{E}[\tilde{r}(x,\mathrm{E}[x|\boldsymbol{\Omega}])|\boldsymbol{\Omega}] = 0$ for $x \in \mathbb{R}^+$, given information set $\boldsymbol{\Omega}$. Therefore, (17) implies that

$$\mathrm{E}[r(X_{ni},\bar{\omega}_{ni}T_{ni})|\boldsymbol{Z},\bar{\boldsymbol{\Gamma}},T_{ni}] = \mathrm{E}\left[r(X_{ni},\mathrm{E}[X_{ni}|\boldsymbol{Z},\bar{\boldsymbol{\Gamma}},T_{ni}])\middle|\boldsymbol{Z},\bar{\boldsymbol{\Gamma}},T_{ni}\right] = 0$$

$\square$

Proof of Proposition 2. Assumptions 1′ and 2 imply that

$$\hat{X}_{ni} = \bar{\phi}_n\bar{\psi}_i e^{\boldsymbol{z}'_{ni}\bar{\boldsymbol{\beta}}} T_{ni}, \tag{18}$$

where

$$T_{ni} = \sum_k \frac{T_i^k c_i^k}{\bar{\phi}_i}\frac{X_n^k}{\Phi_n^k\bar{\phi}_n} e^{\boldsymbol{z}'_{ni}(\boldsymbol{\beta}^k-\bar{\boldsymbol{\beta}})}. \tag{19}$$

Totally differentiating (19), holding constant all values of T_i^k, c_i^k, and X_n, yields

$$d\ln(T_{ni}) = \sum_k \frac{\hat{X}_{ni}^k}{\hat{X}_{ni}}\left[d\boldsymbol{z}'_{ni}(\boldsymbol{\beta}^k-\bar{\boldsymbol{\beta}}) - d\ln(\Phi_n^k) + d\ln(X_n^k) - d\ln(\bar{\phi}_n)\right] \tag{20}$$

Totally differentiating Φ_n^k, given by Assumption 1′, yields

$$d\ln(\Phi_n^k) = \sum_i \frac{\hat{X}_{ni}^k}{X_n^k} d\boldsymbol{z}'_{ni}\boldsymbol{\beta}^k,$$

Totally differentiating (18) yields

$$d \ln \hat{X}_{ni} = d\boldsymbol{z}'_{ni}\bar{\boldsymbol{\beta}} + d \ln(\bar{\phi}_n) + d \ln(T_{ni}),$$

using the fact that ψ_i is an index of $T_i^k c_i^k$, which are held fixed, and thus $d\psi_i = 0$. Multiplying $d \ln \hat{X}_{ni}$ by $\hat{X}_{ni}/X_n$ and summing over i yields

$$d \ln(\bar{\phi}_n) = -\sum_i \frac{\hat{X}_{ni}}{X_n} d\boldsymbol{z}'_{ni}\bar{\boldsymbol{\beta}},$$

which imposes the normalization $\sum_i \frac{\hat{X}_{ni}}{X_n} dT_{ni} = 0$ and uses the fact that $\sum_i d\hat{X}_{ni} = 0$, because X_n is held constant. Substituting these result into (20) yields (10).

To verify that the normalization imposed on T_{ni} is valid, multiply both sides of (20) by $\hat{X}_{ni}/X_n$ and sum over i. This implies that

$$\begin{aligned}
\sum_i \frac{\hat{X}_{ni}}{X_n} d \ln(\tilde{T}_{ni}) &= \sum_i \frac{\hat{X}_{ni}}{X_n} \sum_k \frac{\hat{X}^k_{ni}}{\hat{X}_{ni}} \left[d\boldsymbol{z}'_{ni}(\boldsymbol{\beta}^k - \bar{\boldsymbol{\beta}}) - \sum_m d\boldsymbol{z}'_{nm} \left(\frac{\hat{X}^k_{nm}}{X^k_n}\boldsymbol{\beta}^k - \frac{\hat{X}_{nm}}{X_n}\bar{\boldsymbol{\beta}} \right) + d \ln(X^k_n) \right] \\
&= \sum_k \sum_i \frac{\hat{X}^k_{ni}}{X_n} d\boldsymbol{z}'_{ni}\boldsymbol{\beta}^k - \sum_i \frac{\hat{X}_{ni}}{X_n} d\boldsymbol{z}'_{ni}\bar{\boldsymbol{\beta}} \\
&- \sum_k \sum_m \frac{\hat{X}^k_{nm}}{X_n} d\boldsymbol{z}'_{nm}\boldsymbol{\beta}^k + \sum_m \frac{\hat{X}_{nm}}{X_n} d\boldsymbol{z}'_{nm}\bar{\boldsymbol{\beta}} + \sum_k \frac{dX^k_n}{X_n} \\
&= 0,
\end{aligned}$$

where the last equality uses the fact that $\sum_k dX^k_n = 0$, because X_n is held constant.

□

Proof of Proposition 3.

Part (i). Given any set of heterogeneous coefficients, $\hat{\boldsymbol{\beta}}$, the coefficient index solves the following set of equations

$$\sum_n \sum_i \sum_k \left[\hat{X}^k_{ni}(\hat{\boldsymbol{\beta}}^k) - \hat{X}^k_{ni}(\bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}})) \right] \boldsymbol{z}_{ni} = 0.$$

The structural (and FE) pooled product-level PML estimator estimates a parameter vector $\hat{\bar{\boldsymbol{\beta}}}$ that solves first-order conditions of the form

$$\sum_n \sum_i \sum_k \left[X^k_{ni} - \hat{X}^k_{ni}(\hat{\bar{\boldsymbol{\beta}}}_{\text{PPML}}) \right] \boldsymbol{z}_{ni} = 0. \tag{21}$$

When the pooled PPML estimator is applied to fitted values $\hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}^k)$, these two conditions are identical, which implies that $\hat{\bar{\boldsymbol{\beta}}}_{\text{PPML}} = \bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}})$.

Part (ii). The structural (and FE) product-by-product Poisson PML estimator estimates a parameter vector $\hat{\boldsymbol{\beta}}^k$ that solves first-order conditions of the form

$$\sum_n \sum_i \left[X_{ni}^k - \hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}_{\text{PPML}}^k) \right] \boldsymbol{z}_{ni} = 0, \tag{22}$$

for all k. The coefficient index based on product-by-product PPML estimates solves

$$\sum_n \sum_i \sum_k \left[\hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}_{\text{PPML}}^k) - \hat{X}_{ni}^k(\bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}}_{\text{PPML}})) \right] \boldsymbol{z}_{ni} = 0.$$

This condition implies that

$$\begin{aligned}
\sum_n \sum_i \sum_k \hat{X}_{ni}^k(\bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}}_{\text{PPML}}))\boldsymbol{z}_{ni} &= \sum_n \sum_i \sum_k \hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}_{\text{PPML}}^k)\boldsymbol{z}_{ni} \\
&= \sum_k \left[\sum_n \sum_i \hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}_{\text{PPML}}^k)\boldsymbol{z}_{ni} \right] \\
&= \sum_n \sum_i \sum_k X_{ni}^k \boldsymbol{z}_{ni} \\
&= \sum_n \sum_i \sum_k \hat{X}_{ni}^k(\hat{\bar{\boldsymbol{\beta}}}_{\text{PPML}})\boldsymbol{z}_{ni},
\end{aligned}$$

where the third equality is a result of (22), and the last is a result of (21). Thus, it must be the case that $\bar{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}}_{\text{PPML}}) = \hat{\bar{\boldsymbol{\beta}}}_{\text{PPML}}$.

□

Proof of Proposition 4. Given Assumption 2′, (6) becomes

$$\hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}^k) = \frac{X_n^k}{\Phi_n^k} \frac{Y_i^k}{\Psi_i^k} \tilde{f}(\tilde{\boldsymbol{z}}_{ni}' \hat{\boldsymbol{\gamma}}^k) e^{\hat{\delta}_n^k}, \tag{23}$$

for all $n \neq i$. Summing over all $n \neq i$ implies

$$\hat{E}_i^k = \frac{Y_i^k}{\Psi_i^k} \sum_{n \neq i} \frac{X_n^k}{\Phi_n^k} \tilde{f}(\tilde{\boldsymbol{z}}_{ni}' \hat{\boldsymbol{\gamma}}^k) e^{\hat{\delta}_n^k} \equiv \frac{Y_i^k}{\Psi_i^k} \tilde{\Psi}_i^k.$$

Substituting back into (23) yields

$$\hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}^k) = \frac{X_n^k}{\Phi_n^k} \frac{\hat{E}_i^k}{\tilde{\Psi}_i^k} \tilde{f}(\tilde{\boldsymbol{z}}_{ni}' \hat{\boldsymbol{\gamma}}^k) e^{\hat{\delta}_n^k},$$

for all $n \neq i$. Summing over all $i \neq n$ implies

$$\hat{M}_n^k = \frac{X_n^k}{\Phi_n^k} e^{\hat{\delta}_n^k} \sum_{n \neq i} \frac{\hat{E}_i^k}{\tilde{\Psi}_i^k} \tilde{f}(\tilde{\boldsymbol{z}}'_{ni} \hat{\boldsymbol{\gamma}}^k) \equiv \frac{X_n^k}{\Phi_n^k} e^{\hat{\delta}_n^k} \tilde{\Phi}_n^k.$$

Substituting this expression yields

$$\hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}^k) = \frac{\hat{M}_n^k}{\tilde{\Phi}_n^k} \frac{\hat{E}_i^k}{\tilde{\Psi}_i^k} \tilde{f}(\tilde{\boldsymbol{z}}'_{ni} \hat{\boldsymbol{\gamma}}^k),$$

for all $n \neq i$. Summing again over all $n \neq i$ implies that

$$\tilde{\Psi}_i^k = \sum_{n \neq i} \frac{\hat{M}_n^k}{\tilde{\Phi}_n^k} \tilde{f}(\tilde{\boldsymbol{z}}'_{ni} \hat{\boldsymbol{\gamma}}^k)$$

It remains to be shown that $\hat{E}_i^k = E_i^k$ and $\hat{M}_n^k = M_n^k$. Note that

$$\hat{X}_{nn}^k(\hat{\boldsymbol{\beta}}^k) = \frac{\hat{M}_n^k}{\tilde{\Phi}_n^k} \frac{\hat{E}_n^k}{\tilde{\Psi}_n^k} \tilde{f}(\tilde{\boldsymbol{z}}'_{nn} \hat{\boldsymbol{\gamma}}^k) e^{-\hat{\delta}_n^k}.$$

Structural gravity PML estimators maximize an objective function of the form

$$\mathcal{L} = \sum_n \sum_i X_{ni}^k c(\hat{X}_{ni}^k) - a(\hat{X}_{ni}^k) + b(X_{ni}^k),$$

where $a(\cdot)$ and $c(\cdot)$ are defined such that $[c'(\mu)]^{-1} a'(\mu) = \mu$, for a given scalar μ. Thus, $\hat{\boldsymbol{\beta}}^k$ satisfies first-order conditions of the form

$$\sum_n \sum_i \left(X_{ni}^k - \hat{X}_{ni}^k \right) c'(\hat{X}_{ni}^k) \frac{\partial \hat{X}_{ni}^k}{\partial \hat{\boldsymbol{\beta}}^k} = 0. \tag{24}$$

Recall that $\boldsymbol{\beta}^k$ is of dimension $L = Q + N$, with the last N elements containing the border cost coefficients δ_n^k. The first-order condition with respect to δ_n^k simplifies to

$$\left(X_{nn}^k - \hat{X}_{nn}^k \right) c'(\hat{X}_{ni}^k) = \sum_n \sum_i \left(X_{ni}^k - \hat{X}_{ni}^k \right) c'(\hat{X}_{ni}^k) \left[\frac{\partial \hat{M}_n^k}{\partial \hat{\delta}_n^k} + \frac{\partial \hat{E}_i^k}{\partial \hat{\delta}_n^k} - \frac{\partial \tilde{\Phi}_n^k}{\partial \hat{\delta}_n^k} - \frac{\partial \tilde{\Psi}_i^k}{\partial \hat{\delta}_n^k} \right]. \tag{25}$$

Let us conjecture that the values of $\hat{\delta}_n^k$ are such that $X_{nn}^k = \hat{X}_{nn}^k$. This implies that $\hat{E}_i^k = E_i^k$

and $\hat{M}_n^k = M_n^k$ and that

$$\hat{X}_{ni}^k(\hat{\boldsymbol{\beta}}^k) = \frac{M_n^k}{\tilde{\Phi}_n^k}\frac{E_i^k}{\tilde{\Psi}_i^k}\tilde{f}(\tilde{\boldsymbol{z}}_{ni}'\hat{\boldsymbol{\gamma}}^k) \equiv \tilde{X}_{ni}^k(\hat{\boldsymbol{\gamma}}^k),$$

where

$$\tilde{\Phi}_n^k = \sum_{i\neq n}\frac{E_i^k}{\tilde{\Psi}_i^k}\tilde{f}(\tilde{\boldsymbol{z}}_{ni}'\hat{\boldsymbol{\gamma}}^k) \quad \text{and} \quad \tilde{\Psi}_i^k = \sum_{n\neq i}\frac{M_n^k}{\tilde{\Phi}_n^k}\tilde{f}(\tilde{\boldsymbol{z}}_{ni}'\hat{\boldsymbol{\gamma}}^k).$$

Thus,

$$\frac{\partial\hat{M}_n^k}{\partial\hat{\delta}_n^k} = \frac{\partial\hat{E}_i^k}{\partial\hat{\delta}_n^k} = \frac{\partial\tilde{\Phi}_n^k}{\partial\hat{\delta}_n^k} = \frac{\partial\tilde{\Psi}_i^k}{\partial\hat{\delta}_n^k} = 0,$$

(25) is satisfied, and (24) becomes

$$\sum_n\sum_{i\neq n}\left(X_{ni}^k - \tilde{X}_{ni}^k\right)c'(\tilde{X}_{ni}^k)\frac{\partial\tilde{X}_{ni}^k}{\partial\hat{\boldsymbol{\gamma}}^k} = 0. \tag{26}$$

Thus, the conjecture is verified. Further, Gourieroux et al. (1984) show that the solution to (24) is unique, which implies that the value of $\hat{\boldsymbol{\gamma}}^k$ that solves (26) also constitutes the first Q elements of the unique value of $\hat{\boldsymbol{\beta}}^k$ that satisfies (24).

□

C Data

C.1 Trade Data

Bilateral, product-level trade data are from the U.N. Comtrade database. The data are classified into six-digit Harmonized System (HS), 1996 revision, product codes. The sample consists of trade flows for the year 2003, which was chosen to maximize the number of countries for which both gross output data from INDSTAT and trade data from Comtrade were available. The sample consists of trade flows reported by exporters because these values are more likely to be consistent with the gross output data, which is reported by the producing country, and because exports are typically reported “free on board”, as opposed to “cost, insurance, and freight”, and the former is consistent with the measure of trade flows in the model.

The trade flow data were combined with manufacturing gross output data from several sources. The manufacturing output data are classified according Revision 3 of the International Standard Industrial Classification (ISIC). To match the trade and output data, the HS1996 codes were mapped to ISIC (Revision 3) codes using the concordance available from the U.N. Statistics Division.[58] All HS codes not mapped to manufacturing ISIC codes (2-digit industries 15-37) were dropped. This reduced the number of HS codes in the sample to 4,608.

C.2 Gravity Variables

The bilateral relationship variables used to estimate trade costs are from the Gravity dataset available from CEPII (see Mayer and Zignago, 2011). The estimations use the following variables: population-weighted distance (*distw*), whether countries share a common border (*contig*), whether they have a common official language (*comlang_off*), whether they have ever had a colonial link (*colony*), whether they are currently parties to a regional trade agreement (*rta*), and whether the share a common currency (*comcur*).

C.3 Manufacturing Output

Gross manufacturing output data come from three sources. First, the data are taken from the OECD STAN database, where available. If countries are not included in this database, data are from the Industrial Statistics Database (INDSTAT4), 2011 Edition, CD-ROM published by UNIDO. Where data are available for years before and after, but not including, 2003, log output is linearly interpolated based on the closest values before and after 2003. Where data are not available from either sources, output is imputed from total manufacturing value added obtained from the World Development Indicators database of the World Bank. Gross output is obtained by scaling value added by a factor of 3.04.[59]

Gross output data at the 2-digit ISIC (Revision 3) level were obtained from the INDSTAT2, 2014 Edition, CD-ROM published by UNIDO. Where countries reported data in combined or aggregated categories ISIC categories, these observations were excluded. Table A2 lists the ISIC categories, their descriptions, the number of 6-digit HS codes matched to each ISIC industry, the number of countries that reported output data in each industry, and the industry's share in total world manufacturing Trade.

C.4 Constructing the Sample

To be included in the sample, data must be available for a country from the Comtrade database and at least one of the STAN, INDSTAT, or WDI databases. To avoid problems

[58]This is available for free download from the following url:
http://unstats.un.org/unsd/cr/registry/regdntransfer.asp?f=183.

[59]This value is obtained from a cross-sectional regression of gross output on value added, omitting the constant term. The regression R^2 was equal to 0.99.

related to entrepot trade, China, Hong Kong, and Macao are merged into a single country. There were also several other cases in which there were apparent problems of entrepot trade – i.e. reported exports exceeded reported gross output – which resulted in 8 countries being dropped from the sample.[60] Once the trade and manufacturing data were merged, domestic absorbtion of domestic manufacturing output, X_{ii}, was then calculated as total manufacturing output minus total manufacturing exports to all countries (including non-reporters), and total manufacturing absorbtion, X_i, was calculated as X_{ii} plus total imports from countries in the sample, yielding an internally consistent bilateral trade flow matrix. For the industry-level sample, values of X_{ii} that were computed to be less than zero were excluded. The final sample consists of total gross manufacturing output and bilateral trade flows for 130 countries and 4,608 6-digit manufacturing HS products.[61]

[60]The excluded countries are Armenia, Belgium, the Federated States of Micronesia, Guyana, Luxembourg, Mali, Mongolia, and Singapore.

[61]Note that the industry-level estimations are based on the full sample of 130 countries. The lack of industry-level output only reduces the number of border cost parameters ($\bar{\delta}_n^j$) that can be identified.