DISAGGREGATED GRAVITY: BENCHMARK ESTIMATES AND STYLIZED FACTS FROM A NEW DATABASE

Ingo Borchert, Mario Larch,
Serge Shikher, and Yoto Yotov

ECONOMICS WORKING PAPER SERIES
Working Paper 2020–06–A

U.S. INTERNATIONAL TRADE COMMISSION
500 E Street SW
Washington, DC 20436

June 2020

Disaggregated Gravity: Benchmark Estimates and Stylized Facts from a New Database
Ingo Borchert, Mario Larch, Serge Shikher, and Yoto Yotov
Office of Economics Working Paper 2020–06–A
June 2020

Abstract

The objective of this paper is threefold. First, we test and validate the new International Trade and Production Database for Estimation (ITPD-E) for disaggregated gravity estimations. Second, we capitalize on the rich industry dimension of the ITPD-E to obtain benchmark gravity estimates for a wide range (170) of industries. We document differences and similarities of the impact of the standard gravity variables across the broad sectors of Agriculture, Mining and Energy, Manufacturing, and Services. Third, we use the large number of disaggregated gravity estimates to evaluate the stylized facts and best practice recommendations for gravity estimations. We compare the results obtained using Poisson Pseudo Maximum Likelihood (PPML) estimation with the results from several alternative specifications, including OLS, PPML without the zero trade flows, PPML with interval data, PPML without domestic trade, and PPML without the proper set of fixed effects to control for the multilateral resistances in gravity regressions that pool across industries. The findings from these experiments confirm and reinforce some results from the literature while challenging others.

Ingo Borchert, University of Sussex Business School

Mario Larch, University of Bayreuth, CEPII, CESifo, GEP, ifo

Serge Shikher, United States International Trade Commission

Yoto Yotov, Drexel University, ifo, CESifo

1 Introduction

Due to its solid theoretical foundations, intuitive appeal, and remarkable empirical success, the gravity equation has established itself as the empirical workhorse model of international trade.[1] Along with the latest developments in the theoretical gravity literature, the trade profession has made significant progress in establishing a series of best practices for estimating gravity equations, c.f., Head and Mayer (2014) and Yotov et al. (2016). We have also witnessed improvements in the quality and availability of trade data, both aggregate and disaggregated. Despite the wider availability of more reliable disaggregated trade data, to our best knowledge, there are no studies that obtain comprehensive disaggregated gravity estimates and offer analysis of the heterogeneous impact of trade costs across different industries.[2] While the separability of the trade gravity equation means that the estimations can be performed industry-by-industry, it is not clear whether the estimation practices and recommendations that perform well for aggregate gravity estimations also lead to plausible estimates with highly disaggregated data. Furthermore, the vast majority of the gravity estimates from the literature are obtained from data on international trade flows only, thereby ignoring the theory-motivated use of domestic trade, mainly due to the lack of necessary data.

To fill this gap, we use the newly available International Trade and Production Database for Estimation (ITPD-E) that includes international and domestic trade for 170 industries, 243 countries, and 17 years, 2000-2016 (Borchert et al., 2020). We take advantage of the rich cross-industry dimension of the ITPD-E to obtain benchmark gravity estimates for the whole range of 170 ITPD-E industries, and we document wide but intuitive differences between the estimates of a series of standard gravity covariates across the broad ITPD-E sectors of Agriculture, Mining and Energy, Manufacturing, and Services. We confirm some stylized facts and best estimation practices from the related literature, while we challenge others. These results demonstrate the usefulness of the ITPD-E for structural gravity analysis and open many opportunities for future research.

We start by using the latest gravity methodology to obtain estimates for the impact of the standard gravity covariates (e.g., distance, contiguity, language, etc.) for each of the 170 industries in the ITPD-E. As expected, we find that distance is a very significant impediment to disaggregated trade, while sharing a common border, speaking the same official language, sharing colonial ties, and being members of the same free trade agreement promote international trade. We also estimate significant border barriers, especially in Services and Agriculture. On average, our industry-level ITPD-E gravity estimates are readily comparable to existing estimates from the literature and, therefore, they offer encouraging support for using the ITPD-E for disaggregated gravity estimations.

[1] We refer the reader to Anderson (2011), Arkolakis, Costinot and Rodríguez-Clare (2012), Costinot and Rodríguez-Clare (2014), and Yotov et al. (2016) for excellent surveys of the evolution of the theoretical gravity literature and the alternative microfoundations of structural gravity. Baldwin and Taglioni (2006), Head and Mayer (2014), and Yotov et al. (2016) offer informative discussions of the data and econometric challenges with gravity estimations.

[2] Traditionally, the gravity model has been estimated mostly with aggregate data and there is enough robust evidence for the impact of a series of traditional determinants of trade flows such as bilateral distance, contiguity, sharing an official language, colonial ties, etc. Head and Mayer (2014) offer benchmark gravity estimates and discussion in the form of an excellent mata analysis.

We document wide but intuitive variation in gravity coefficient estimates across ITPD-E industries. In order to identify systematic patterns, we focus on the differences in the gravity estimates across the four broad sectors of Agriculture, Mining and Energy, Manufacturing, and Services. Our estimates reveal that geography (proxied by distance and contiguity) has a stronger impact on trade in Mining and Energy and Agriculture, and a smaller impact in Manufacturing and, especially, Services. We also find that language is a stronger determinant of trade in Manufacturing and, even more so, Services, while it does not play such an important role in Agriculture and Mining and Energy. The impact of colonial relationships is similar across the four broad sectors. Interestingly, our estimates suggest that FTAs play a more significant role in promoting trade in Agriculture and Mining and Energy and less so in Manufacturing and in Services. Finally, our border estimates are largest in Services, followed by Agriculture, Mining and Energy, and then Manufacturing.

We capitalize on the large number of industries, inclusion of domestic trade, and significant fraction of zeros in the ITPD-E to shed light on the issues and estimation considerations that often arise in the empirical gravity literature. Specifically, we compare our main/benchmark Poisson Pseudo Maximum Likelihood (PPML) estimates with the results from the following alternative specifications: (i) OLS estimates that are obtained with the same sample; (ii) PPML estimates without taking into account the zero trade flows in ITPD-E; (iii) PPML estimates with interval data; (iv) PPML estimates without domestic trade; and (v) PPML estimates that do not include the proper set of fixed effects to control for the multilateral resistances in gravity regressions that pool across industries. The findings from these experiments confirm and reinforce some stylized facts from the literature while challenging others.

Comparing the industry-level gravity estimates that are obtained with PPML vs. OLS we draw three conclusions. First, consistent with the results from Santos Silva and Tenreyro (2006, 2011), we confirm that OLS and PPML deliver substantially different gravity estimates also at the disaggregated level. Second, we observe that, as compared to the standard OLS estimator, PPML tends to give larger estimates for large effects and smaller estimates for small effects of the determinants of bilateral trade flows. Correspondingly, the variation among the PPML estimates across industries is significantly larger as compared to the variation in the corresponding OLS estimates. Third, consistent with the findings of Hinz et al. (2020), our border gravity estimates reveal that the PPML and OLS estimates converge when the effects of the determinants of trade are allowed to vary by country. The intuition for this result is that PPML weights observations by country-size.

In our second experiment we investigate the impact of including the zero trade flows in gravity regressions by comparing the industry-level estimates that are obtained with and without taking into account the zero trade flows. Our main finding is that the results with and without taking into account the zero trade flows are virtually identical. Based on this, we draw three conclusions. First, consistent with the key argument of Santos Silva and Tenreyro (2006, 2011), the main value in using PPML is to account for heteroskedasticity and not to take into account the information contained in zero trade flows. Second, consistent with the analysis of Hinz et al. (2020), a possible explanation for the finding that the zeros do not matter is that zero trade flows are more likely to arise from small countries, which are discounted in the PPML first order conditions. Third, while ITPD-E is relatively disaggregated as compared to other gravity datasets that include domestic trade,we do recognize the possibility that zeros may play a much more significant role with more disaggregated data and, of course, the significant presence of zeros and their proper modeling and estimation are very important for analysis of the extensive margin of trade.

The goal of our third experiment is to compare the industry-level ITPD-E gravity estimates that are obtained with the full sample with the corresponding gravity estimates that are obtained with interval data. We use two alternative interval samples, one with 2-year intervals and one with 4-year intervals. Regardless of the interval employed, our results reveal that the industry-level estimates that we obtain with consecutive-year data vs. interval data are almost identical. Thus, consistent with recent findings of Egger, Larch and Yotov (2020), our conclusion is that gravity estimations can/should be performed without throwing away data. This will not only improve estimation efficiency but also allow for more proper quantification of the impact of trade policies whose effects may have been biased due to the 'arbitrary' dropping of observations for estimations with interval data.

Our fourth experiment compares estimates of the standard gravity variables that are obtained with and without the domestic trade data. Estimating gravity with international trade data only is still a common practice in trade literature, mainly due to lack of reliable domestic trade data. The key finding from this analysis is that the industry-level estimates of standard gravity variables are almost identical whether we use the domestic trade data or not. Our explanation for this result is that the time-invariant country specific dummy variables that we employ for domestic trade do a good job overall in capturing the impact of various domestic trade costs. Of course, by construction, the country-specific domestic trade dummies absorb the impact of all time-invariant domestic trade costs, e.g., geography. However, our results imply that during the period covered by ITPD-E there has not been much time variation in domestic (relative to international) trade costs. A possible explanation for this result is that ITPD-E covers a period of two recessions and slow recovery after the great recession of 2008, where trade costs fluctuated a lot but remained unchanged on average.

We conclude the analysis with an investigation of the importance of properly controlling for the multilateral resistance terms of Anderson and van Wincoop (2003). To this end we compare two sets of gravity estimates with data pooled across ITPD-E industries within each of the four broad ITPD-E sectors, i.e. Agriculture, Mining and Energy, Manufacturing, and Services, respectively. The first set of results are obtained with theory-motivated fixed effects (i.e., exporter-industry-time and importer-industry-time fixed effects) whereas the second set of estimates are obtained with exporter-time, importer-time, and industry fixed effects instead. Comparison between the two sets of estimates reveals that while some of them are similar many of them are significantly different, especially in the case of our distance and border estimates. Therefore, and consistent with Anderson and van Wincoop (2003) and Baldwin and Taglioni (2006), we conclude that proper control for the multilateral resistances is indeed important in the case of disaggregated gravity estimations.

The rest of the paper is organized as follows. Section 2 offers a brief review of the industry-level gravity model and of best practices in estimating structural gravity. Section 3 showcases the main features of the International Trade and Production Database for Estimation (ITPD-E). Section 4 presents benchmark disaggregated gravity estimates and documents some intuitive systematic differences in gravity estimates across the four broad sectors in the ITPD-E. Section 5 uses the large number of industry-level gravity estimates to offer support for some stylized facts from the related literature and to challenge others. Section 6 concludes with a summary of our findings and with directions for future work.

2 Disaggregated Gravity: Theory and Estimation

Many trade policies (e.g. tariffs but also non-tariff measures) are designed and implemented at a disaggregated level (e.g. at the industry or at the product level). Therefore, it is important and desirable to be able to evaluate the effects of such policies at a level of aggregation as close as possible to the one at which they are applied; in terms of industry disaggregation, ITPD-E reaches down further than any other publicly available data source. Furthermore, as noted by Yotov et al. (2016), even for policies that are applied at the aggregate level, e.g. some regional trade trade agreements, it may be desirable to also obtain industry-specific effects because the effects of these non-discriminatory policies may actually be quite heterogeneous across industries. These examples point to the need for proper disaggregated gravity theory and estimations. Accordingly, the objective of this section is to review the industry-specific version of the structural gravity model and the associated best estimation practices and recommendations.

The good news on that front is that one of the most attractive properties of the workhorse model of empirical trade—the structural gravity model—is that it is separable at the industry level, i.e. it can be derived and estimated following theory both at the aggregate level and at any level of disaggregation for which data are available. Following Anderson and van Wincoop (2003), Anderson and van Wincoop (2004) derive a disaggregated gravity model on the demand side. Similarly, following Eaton and Kortum (2002), Shikher (2011) and Costinot, Donaldson and Komunjer (2012) derive a supply-side disaggregated gravity model.[3] Motivated by Arkolakis, Costinot and Rodríguez-Clare (2012), who demonstrate that a very large class of theoretical trade models converge to the gravity equation, Yotov et al. (2016) employ the notation of Anderson and van Wincoop (2003) to demonstrate the equivalence between the industry-level structural gravity systems on the demand side and on the supply side by deriving the following gravity equation:

$$X_{ij,t}^k = \frac{Y_{i,t}^k E_{j,t}^k}{Y_t^k} \left(\frac{t_{ij,t}^k}{\Pi_{i,t}^k P_{j,t}^k} \right)^{1-\sigma^k}, \tag{1}$$

where: i denotes the exporting/source country, j denotes the importing/destination country, k denotes the industry,[4] and t is a time subscript; $X_{ij,t}^k$ is the nominal trade flow (in levels) from i to j in industry k at time t; $Y_{i,t}^k$ is the value of output (at delivered prices) at time t and industry k in origin i; $E_{j,t}^k$ is the expenditure at time t and industry k in destination j; $t_{ij,t}^k$ denotes the bilateral trade frictions between i and j in industry k at time t; and,

[3]For more recent derivations of disaggregated demand-side gravity models, we refer the reader to Larch and Wanner (2017), and for more recent derivations of disaggregated gravity models on the supply side see Caliendo and Parro (2015) and Donaldson (2016).

[4]While k may denote any classification, e.g. product, sector, or industry, for convenience and simplicity we will refer to class k as industry in the rest of the analysis as corresponding to the most disaggregated level in ITPD-E.

finally $\Pi_{i,t}^k$ and $P_{j,t}^k$ are the outward and inward multilateral resistances of Anderson and van Wincoop (2003). σ^k denotes the elasticity of substitution between the different varieties in industry k.

Equation (1) has three implications for gravity estimations. First, according to (1), if the estimating gravity sample pools across different industries, then the bilateral trade costs $t_{ij,t}^k$, including the effects of any trade policy or time-invariant trade costs, should be industry-specific. Second, in addition, the fixed effects that are now routinely used in gravity estimations to account for the multilateral resistances should be of dimensions exporter-industry-time and importer-industry-time, respectively. Of course, the gravity model can be estimated separately for each industry, in which case the fixed effects that account for multilateral resistances should be of dimension exporter-time and importer-time. Third, consistent with theory and as recommended by Yotov et al. (2016), gravity estimations should be performed with international trade flows and domestic trade. As emphasized below, this is an important adjustment that is often overlooked in the related literature. A key feature of the new ITPD-E is its inclusion of consistently constructed domestic trade, which will be employed throughout and thus harnessed for the analyses in this paper.

In addition to estimating disaggregated gravity with (i) industry-specific bilateral trade costs, (ii) exporter-industry-time and importer-industry-time fixed effects, and (iii) domestic trade, Yotov et al. (2016) recommend that gravity estimations are performed: (iv) with pair fixed effects (to comprehensively account for any time-invariant trade costs and to mitigate endogeneity concerns with respect to bilateral trade policies), (v) with the PPML estimator (to account for heteroskedasticity and to take into account the information contained in the zero trade flows); and (vi) with interval data instead of consecutive years (in order to allow for adjustments of trade flows in response to trade policy changes.) Given that one of the main objectives of this paper is to validate the new ITPD-E for gravity estimations and that the vast majority of gravity estimates in the literature are obtained with standard gravity variables, we will estimate our specifications without pair fixed effects. However, we will implement and discuss the implications of all other best practice recommendations.

Taking all of the above considerations into account, we specify the following industry-level estimable gravity equation:

$$\begin{aligned} X_{ij,t}^k &= \exp[\gamma_1^k DIST_{ij} + \gamma_2^k CNTG_{ij} + \gamma_3^k LANG_{ij} + \gamma_4^k CLNY_{ij}] \times \\ &\quad \exp[\gamma_5^k FTA_{ij,t} + \sum_i \gamma_{6,i}^k SMCTRY_{ij} + \psi_{i,t}^k + \phi_{j,t}^k] \times \varepsilon_{ij,t}^k, \end{aligned} \tag{2}$$

where, in order to demonstrate the representativeness of the ITPD-E, we employ as covariates the set of standard and most widely used gravity variables from the literature including the logarithm of bilateral distance between trading partners ($DIST_{ij}$), as well as indicator variables for contiguity ($CNTG_{ij}$), common official language ($LANG_{ij}$), colonial relationships ($CLNY_{ij}$), and the presence of a free trade agreement ($FTA_{ij,t}$), respectively. All standard gravity variables employed in this paper come from the recently constructed Dynamic Gravity Dataset (DGD) of the U.S. International Trade Commission (Gurevich and

Herman, 2018).[5] Finally, each specification includes a set of country-specific dummy variables ($\sum_i SMCTRY_{ij}$) that take a value of one for domestic trade and are set to zero for international trade.

3 The International Trade and Production Database

This section summarizes the key features of the International Trade and Production Database for Estimation (ITPD-E) developed by Borchert et al. (2020). ITPD-E is constructed from four main different sources for Agriculture, Mining and Energy, Manufacturing, and Services. For Agriculture the trade and production data come from the Food and Agriculture Organization of the United Nations Statistics Division (FAOSTAT). Manufacturing and Mining and Energy trade data are obtained from the UN Commodity Trade Statistics Database (COMTRADE), while the production data are from the UNIDO United Nations Industrial Statistics (INDSTAT) Database. For services trade, ITPD-E uses information from the WTO-UNCTAD-ITC Annual Trade in Services Database and the UN Trade in Services Database (UN TSD). Services gross output data are from the UN System of National Accounts (UN SNA) Database.

ITDP-E is based on reported import flows for Agriculture, Mining and Energy, Manufacturing, and on export flows for Services trade. Mirror exports reported by partner countries (mirror imports for services) are used to fill missing import (export) values. Domestic trade is calculated as the difference between the values of total (gross value) production and total exports. ITPD-E is balanced across the exporter, importer, industry and time dimension by filling all remaining missing observations with zeros. In order to drop irrelevant zeros, the final dataset keeps only observations that are retained when estimating a gravity model using the PPML estimator with exporter-time, importer-time, and directional bilateral fixed effects. ITPD-E drops countries not included in the comprehensive geopolitical Dynamic Gravity Database of the US ITC.

In terms of time coverage, ITPD-E covers the 17-year period between 2000 and 2016. ITPD-E covers 170 industries, of which 26 are in Agriculture, 7 are in Mining and Energy, 120 are in Manufacturing, and 17 are in Services. The number of countries in ITPD-E is 243. Hence, ITPD-E provides a comprehensive dataset for many industries and many years that includes international as well as domestic trade. All these features, which are unique in their combination, render this new dataset particularly suited for detailed industry-level gravity estimation and analyses.

4 Disaggregated Gravity Estimates

This section presents gravity estimates based on Equation (2) for each of the 170 industries in the ITPD-E. Our goal is twofold. First, we investigate the suitability of the newly

[5] Three features of the ITC Dynamic Gravity Dataset (DGD) make it appealing for our purposes. First, the DGD includes a time dimension which accounts for the separation or combination of countries and territories. Second, the country coverage between ITPD-E and DGD is very close. Third, in addition to the standard gravity covariates employed here, DGD includes a series of additional bilateral variables as well as many country-specific variables which can be used in more comprehensive gravity specifications.

constructed ITPD-E for gravity estimation by comparing gravity coefficients on a series of standard variables obtained with ITPD-E data to estimates from existing studies, e.g. to the benchmark meta analysis gravity estimates of Head and Mayer (2014). To the best of our knowledge, there are no corresponding theory-consistent gravity estimates at a similar level of disaggregation as the ones presented in this paper. Therefore, we hope that the current estimates can be used as benchmark results for disaggregated gravity. Second, ITPD-E offers a unique opportunity to obtain, compare, and analyze industry-level gravity estimates across a wide range of industries within each the four broad sectors (Agriculture, Mining and Energy, Manufacturing, and Services) in ITPD-E. Thus, our second objective is to provide a set of consistent and comparable industry-level gravity estimates and to discuss the main differences across the main industries in ITPD-E.

4.1 Disaggregated Gravity Estimates: Benchmark Results

Our main results appear in Table 1.[6] The table reports OLS gravity estimates (in Panel A) and PPML gravity estimates (in Panel B) that are obtained with the full ITPD-E data. Whereas, as discussed earlier, our preferred estimator is PPML, we also report and discuss OLS estimates for comparison purposes; both because OLS has been the standard estimator for a long time, and in order to compare the PPML vs. the OLS results. Thus, the dependent variable in Panel A of Table 1 is the logarithm of nominal bilateral trade flows, while the dependent variable in Panel B is nominal trade in levels. All estimates in Table 1 are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. We also omit for brevity the standard errors and t-statistics of the estimates. Instead, we just report significance levels as follows: * $p < 0.10$, ** $p < .05$, *** $p < .01$. Finally, since we obtain a very large number of country-specific $SMCTRY$ estimates (potentially 243 estimates for each industry), we aggregate them to a single index per industry by constructing averages across the country-specific $SMCTRY$ estimates. The full set of $SMCTRY$ estimates with their standard standard errors are available upon request.

Several findings stand out from the estimates in Table 1. First, we are able to obtain OLS and PPML estimates of Equation (2) for each industry in ITPD-E.[7] Second, we confirm many of the standard and well-accepted results from the empirical gravity literature. Specifically, we find that distance ($DIST$) is a significant impediment to trade. All of the OLS estimates on $DIST$ are negative and highly statistically significant and all but five of the PPML estimates on $DIST$ are negative and statistically significant (with four of the insignificant estimates being negative as well). We also confirm the standard finding in the gravity literature that sharing a common border ($CNTG$) promotes trade. All OLS estimate of the impact of $CNTG$ are positive and only 3 of them are not statistically significant, while 15 of the PPML estimates on $CNTG$ are negative but only one of them is statistically significant.

As indicated by our estimates, sharing a common official language also promotes trade; only 4 of the OLS estimates on $LANG$ are negative and none of them is statistically significant. With 24 negative values (three of them statistically significant), the PPML estimates

[6] Table 1 is ordered by ITPD-E industry IDs shown in the first column. Please see the Appendix for full industry descriptions.

[7] We also confirm that we can obtain convergence when we replace the standard time-invariant gravity covariates with a rich set of directional pair fixed effects. These estimates are available by request.

on language are a bit weaker but still offer strong evidence that sharing a common official language promotes bilateral trade. We also confirm the strong link between past colonial relationships and bilateral trade. Only 8 of the OLS estimates on $CLNY$ are negative and only one of them is statistically significant. In the case of PPML, we obtain 17 negative estimates on $CLNY$ but only 4 of them are statistically significant.

Turning to the policy variable in our gravity specification, we see that FTAs promote trade significantly.[8] Only 4 of the OLS estimates on FTA are negative and none of them are statistically significant. The PPML estimates are similar, only 7 of them are negative, and only one of them is statistically significant. Finally, we obtain large, positive, and statistically significant estimates on $SMCTRY$, which suggest that bilateral trade flows are subject to large international border effects. Without any exception, both the OLS and the PPML estimates on $SMCTRY$ (for which we are able to obtain estimates[9]) are negative, large, and statistically significant at any conventional level. Home bias is a natural explanations for the large border estimates that we obtain, which may be due to demand-side factors such as differences in preferences or other aspects of tradability such as value-to-weight ratio or unobserved policy frictions. We confirm this intuition in the next section, in which we demonstrate that the border estimates vary intuitively across the four main sectors in ITPD-E.

4.2 On the Heterogeneity of Trade Costs Across Main Sectors

On average, the results from Table 1 are exactly as expected and readily comparable with corresponding aggregate estimates from the related literature, e.g. with the meta-analysis gravity estimates of Head and Mayer (2014). Thus, our results offer good support for the representativeness of ITPD-E for industry-level gravity estimations.

At the same time, the large number of 170 sets of disaggregated gravity estimates also exhibit wide variation in the estimates for each gravity covariate, and these differences are systematically linked to the four broad sectors in the ITPD-E (Agriculture, Mining and Energy, Manufacturing, and Services). For example, it seems that distance is a stronger impediment to trade in Mining and Energy and Agriculture than in Manufacturing and Services. This makes intuitive sense. The objective of this section is to gauge the significance of the variation in our estimates across the four main industry categories in the ITPD-E and to identify and document systematic differences in the impact of the standard gravity determinants of trade costs across Agriculture, Mining and Energy, Manufacturing, and Services.

Our key findings are presented in Table 2. Panel A of Table 2 reports simple averages and standard deviations (in parentheses), which are obtained from the industry-level estimates in Table 1. Comparison of the estimates across the four main ITPD-E sectors reveals several intuitive patterns. First, we find that the impact of distance is the strongest in Mining and Energy, followed by Agriculture. Weight to value transportation costs are the natural

[8] We are aware that the FTA coefficient estimates may be subject to endogeneity concerns. However, our purpose here is to establish the representativeness of ITPD-E by comparing the FTA estimates with the vast majority of existing results that do not take FTA endogeneity into account.

[9] There are 11 ITPD-E industries for which domestic trade cannot be constructed; all these industries are either part of Agriculture or Services.

explanation for this result. We also see that the effect of distance on both manufacturing and services trade is significantly lower. The latter can be explained with the intangible nature of services trade. The variation in the negative impact of distance is mirrored by the positive effects of contiguity, which are the strongest for Mining and Energy and Agriculture and weakest for Manufacturing and Services. In sum, our estimates reveal the extent to which geography plays an appreciably stronger role for Mining and Energy and Agriculture and is less important for trade in Manufacturing and especially for Services.

We also document some interesting cross-sectoral patterns for the impact of language. Specifically, our estimates reveal that common language is a stronger determinant of trade in Manufacturing and particularly so in Services trade, in which the pronounced language effect is driven specifically by Travel, Financial services, and Education, respectively (see industries 157, 160 and 166 in Table 1). Communication and therefore language is important for the exchange of services, especially in the aforementioned industries. By contrast, language does not play a strong role for trade in agricultural and mining products, respectively. We find this variation intuitive, against the backdrop that trade in standardized commodities, which is less dependent on language, is common in Agriculture, Mining and Energy, whereas trade in differentiated goods is more prevalent in Manufacturing and Services. The estimates in Panel A of Table 2 also suggest that past colonial relationships have relatively strong and homogeneous impact across Agriculture, Manufacturing and Services. However, the estimate for Mining and Energy is very small and in fact negative, which is driven by the large and negative estimate on "Gas production and distribution" at the industry-level (see Table 1).

The large positive average estimates of the impact of FTAs reveal that regional trade agreements are an important determinant of international trade for each of the four main sectors in ITPD-E. Interestingly, our estimates suggest that FTAs have played a more significant role in promoting trade in Agriculture and Mining and Energy and somewhat less so in Manufacturing and in Services. The FTA effect for services trade, approximately on par with the one for manufacturing, points to linkages between goods and services trade as the FTA variable effectively captures merchandise goods trade provisions and, to the extent that these are accompanied by services chapters, the provisions therein often only bind existing levels of market access.

Finally, we turn to the average estimates on $SMCTRY$, which denote the extent to which international borders reduce international trade. By some margin, the border effect is the largest in Services, followed by Mining and Energy and Agriculture, whereas Manufacturing trade exhibits by far the lowest estimate. Highly localized and regulated consumption or services, home bias in Agriculture and Mining and Energy, and pronounced specialization and global value chains in Manufacturing are natural explanations for these results.

Panel B of Table 2 reports average gravity estimates for the same four broad sectors; yet this time the estimates are obtained with pooled industry-level data and from panel PPML gravity specifications that correspond to equation (2), and where, in accordance with theory, the set of fixed effects are exporter-industry-time and importer-industry-time. We first note that all estimates in Panel B have signs as expected. In addition, all but one of the estimates (the exception is the estimate on language for Mining and Energy) are statistically significant. Finally, while the magnitude of the estimates varies across covariates and industries, they are all economically sizable. The estimates in Panel B of Table 2 also confirm most of the qualitative patterns that we identified and discussed based on the results in Panel A. Specifically, we see that (i) the negative impact of distance is stronger for Mining and Energy and Agriculture than in Manufacturing and Services, while (ii) the positive impact of contiguity is stronger in Mining and Energy and Agriculture and weaker in Manufacturing and Services. We also confirm that (iii) language is more important for trade in Services compared to Manufacturing or Mining and Energy. The estimates of the impact of colonial ties between Panels A and B are also comparable, the main difference being that in Panel A we also find that colonial relationships have a significant impact on trade in Mining and Energy. The pattern in the effects of FTAs is also consistent between Panels A and B. We do note, however, that, echoing the FTA literature's struggle to find significant impact of FTAs on services trade, our estimate for Services is the smallest. Yet this result is plausible given that the trade agreement variable employed in these estimates only captures free trade agreements in the sense of GATT Article XXIV and, therefore, does not have a direct bearing on services trade.

In sum, the analysis in this section demonstrates that the industry-level gravity estimates from Tables 1 and 2 are readily comparable to existing estimates from the literature. For example, our estimates are very close to the meta-analysis results from Head and Mayer (2014). In addition, comparison across the four broad sectors in the ITPD-E reveals significant and intuitive variation patterns across the impact of standard gravity covariates on trade in Agriculture vs. Mining and Energy, Manufacturing and Services, respectively. This further validates the use of the ITPD-E for gravity estimations.

5 Gravity Estimations and Results: Some Stylized Facts

In this section we capitalize on the rich dimensionality of ITPD-E (e.g., wide industry coverage, inclusion of domestic trade, and significant fraction of zeros) to shed light on some issues and estimation considerations that often arise in the empirical gravity literature. Specifically, we compare our main/benchmark PPML estimates from Table 1 with the results from the following alternative specifications: (i) OLS estimates that are obtained with the same sample, in Section 5.1; (ii) PPML estimates without taking into account the zero trade flows in ITPD-E, in Section 5.2; (iii) PPML estimates with interval data, in Section 5.3; (iv) PPML estimates without domestic trade, in Section 5.4; and (v) PPML estimates that do not include the proper set of fixed effects to control for the multilateral resistances in gravity regressions that pool across industries, in Section 5.5.

5.1 Industry-level Gravity with OLS vs. PPML

While the PPML estimator has established itself as the leading gravity estimator (cf. Santos Silva and Tenreyro, 2011, 2006), many researchers still rely on the OLS estimator. In addition, while we do not question the deficiency of OLS in terms of handling heteroskedasticity, we also agree with arguments put forward by Head and Mayer (2014), Egger and Staub (2014), and Hinz et al. (2020) that there is value in obtaining gravity results with both estimators; for instance, in order to detect areas for improvement in the specification of the bilateral trade cost vector, c.f. Hinz et al. (2020). Therefore, in the first experiment, we compare industry-level ITPD-E gravity estimates that are obtained with PPML vs. OLS.

Our findings are presented in Figure 1. Each panel of the figure compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. Both sets of estimates are based on the positive trade data of ITPD-E, and each of them is obtained with the proper set of exporter-time and importer-time fixed effects, according to specification (2). The only difference between the two sets of estimates in Figure 1 is that one of them is obtained with the PPML estimator (excluding zero trade flows which cannot be handled by OLS) while the other one is obtained with the OLS estimator. Each dot in each panel represents an estimate for a particular industry and the estimates are ordered from the smallest to the largest PPML estimate with all data.

The results in Figure 1 reveal several patterns. First, consistent with Santos Silva and Tenreyro (2006, 2011), we confirm that OLS and PPML deliver industry-level gravity estimates that can be quite different from each other. By extension, OLS and PPML also deliver different industry rankings according to the estimated coefficients. Second, we note that the shapes of the distribution of the PPML and the corresponding OLS estimates for each of the gravity variables in our specifications are similar, i.e., they are positively correlated. However, we do observe some notable patterns and differences.

Panel A of Figure 1 shows that the industry-level PPML estimates of distance elasticities are usually smaller (in absolute value) as compared to the corresponding OLS estimates, especially so for the industries with small (in absolute value) distance estimates. The converse, however, is the case for the industries with the largest—i.e. most negative—distance elasticities, where OLS estimates are appreciably smaller than PPML estimates. In other words, PPML delivers larger estimates (in absolute value) for large distances elasticities and smaller ones for small values, compared to OLS coefficients.

Panels B, C, and D of Figure 1 reveal that PPML usually produces lower estimates of the impact of contiguity ($CNTG$), common official language ($LANG$), and colonial relationships ($CLNY$). Yet the relationship is reversed for industries with largest estimates of these effects. Panel E of Figure 1 reveals significant overlap between the PPML and the OLS estimates of the effects of FTAs for most industries. Once again, though, we observe differences in the tails of the distribution. Specifically, the PPML estimates are larger than the corresponding OLS estimates for industries with very large estimates and they are smaller for industries with very small estimates. Finally, Panel F of Figure 1 reveals that the $SMCTRY$ estimates obtained with PPML and OLS are quite similar.

Based on these observations and on the corresponding estimates in Figure 1, we draw four conclusions about the relationship between the performance of PPML and OLS for structural gravity estimations at the industry level. First, consistent with Santos Silva and Tenreyro (2006, 2011), we find that OLS and PPML deliver substantially different disaggregated gravity estimates. Second, as compared to the standard OLS estimator, PPML tends to lead to large estimates for large effects and smaller estimates for small effects of the determinants of bilateral trade flows. Third, on a related note, the variation among the PPML estimates across industries is significantly larger as compared to the corresponding variation in OLS estimates.[10] Fourth, consistent with the findings of Hinz et al. (2020), our estimates on

[10]Without any exception, the standard deviations of the PPML estimates across the ITPD-E industries are always larger as compared to their OLS counterparts, i.e., 0.437 vs. 0.304 for distance (DIST), 0.396 vs. 0.338 for contiguity (CNTG), 0.591 vs. 0.511 for colony (CLNY), 0.275 vs. 0.260 for language (LANG), and 0.372 vs. 0.286 for FTAs.

SMCTRY from Panel F of Figure 1, reveal that the PPML and OLS estimates converge when the effects of the determinants of trade are allowed to vary by country.

5.2 Industry-level Gravity With and Without Zeros

Zero trade flows have been the focus of interest in a number of influential trade studies. Helpman, Melitz and Rubinstein (2008) offer an informative graph that shows that about 50% of the possible bilateral links for aggregate trade in the world are zeros and that this statistic has been stable over time. The problem with zero trade flows becomes more pronounced the more disaggregated the trade data are, and it is especially severe for services trade (due to the highly localized consumption and highly specialized production). As documented earlier, even when only the relevant zeros are kept for gravity estimations with ITPD-E, they still comprise a very significant fraction of the observations that are used for our estimations within each industry.

Until recently, most gravity studies just ignored the information contained in the zero trade flows by employing the OLS estimator.[11] However, in addition to addressing the issue of heteroskedasticity and due to its multiplicative form, the PPML estimator is an easy and convenient solution to take into account the information contained in the zero bilateral trade flows. In our next experiment we investigate the importance of including zero trade flows in gravity estimations with the ITPD-E by comparing, for each standard gravity variable, the industry-level estimates obtained with and without taking into account zero trade flows.

Our findings are reported in Figure 2. Each panel of this figure compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. The first set of estimates is obtained with the PPML estimator and all data from the ITPD-E whereas the second set of results is obtained also with the PPML estimator but data on positive trade flows only. Both sets of estimates are obtained with the proper set of exporter-time and importer-time fixed effects, according to specification (2). Each dot in each panel represents an estimate for a particular industry and the estimates are ordered from the smallest to the largest PPML estimate with all data.

The main message from Figure 2 is that the results with and without taking into account the zeros are virtually identical. Based on this, we draw three conclusions. First, consistent with the main argument made by Santos Silva and Tenreyro (2006, 2011), the principal value of using PPML is to account for heteroskedasticity and not to take into account the information contained in zero trade flows. That said, there is a discernible difference for large values of the distance elasticity (Panel A) and border effects (Panel F), respectively. Without zeros, both distance elasticities and border barriers are underestimated, which seems plausible. For instance, zero trade flows are more likely to arise over long distances and therefore, if these observations were included in the sample, distance will be found to exert a stronger effect in industries in which distance matters significantly. Second, consistent with the analysis of Hinz et al. (2020), a possible explanation for the finding that the zeros do not matter is that

[11] Over the years various studies have tried to account for the presence of zeros. For example, Eaton and Tamura (1995) and Martin and Pham (2008) propose the use of Tobit estimators as an econometric solution to the presence of zeros. More recently, Helpman, Melitz and Rubinstein (2008) offer a theoretically-founded two-step selection process, where exporters must absorb some fixed costs to enter a market. See Yotov et al. (2016) for further discussion.

the zero trade flows are usually associated with small countries, which are discounted in the PPML first order conditions. Third, while ITPD-E is relatively disaggregated as compared to other gravity datasets that include domestic trade, we do recognize the possibility that zeros may play a much more significant role with more disaggregated data and, of course, their presence and modeling is very important on the extensive margin of trade.

5.3 Industry-level Gravity Estimations with Interval Data

Adjustment of trade flows in response to trade policy changes will not be instantaneous. Cheng and Wall (2005) argue that the challenge of adjustment is even more pronounced in econometric specifications with fixed effects such as gravity estimations: "[f]ixed-effects estimation is sometimes criticized when applied to data pooled over consecutive years on the grounds that dependent and independent variables cannot fully adjust in a single year's time." (Footnote 8, p. 52, Cheng and Wall, 2005). Trefler (2004) also criticizes trade estimations pooled over consecutive years. In order to avoid this critique, researchers have used panel data with intervals instead of data pooled over consecutive years. For example, Trefler (2004) uses 3-year intervals, Anderson and Yotov (2011) use 4-year intervals, Baier and Bergstrand (2007) use 5-year intervals, and Olivero and Yotov (2012) experiment with 3-year and 5-year interval trade data. More recently, Egger, Larch and Yotov (2020) challenge the use of data with intervals for estimating the impact of trade policy in favor of gravity estimations that use all data, employ pair fixed effects, and allow for phasing in trade policy estimates.

The goal of our next experiment is to compare industry-level ITPD-E gravity estimates that are obtained with the full sample vs. corresponding gravity estimates that are obtained with interval data. The results appear in Figures 3 and 4. Each panel of Figure 3 compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. The first set of estimates is obtained with the PPML estimator and with all data from the ITPD-E, while the second set of results is obtained also with the PPML estimator but with 2-year interval data. Both sets of estimates are obtained with the proper set of exporter-time and importer-time fixed effects, according to specification (2). Each dot in each panel represents an estimate for a particular industry and the estimates are ordered from the smallest to the largest PPML estimate with all data. Figure 4 reproduces the results from Figure 3 but with data over 4-year intervals.

The main message from Figures 3 and 4 is that the industry-level estimates that we obtain with consecutive-year data vs. 2-year interval data vs. 4-year interval data are almost identical. Thus, consistent with recent findings of Egger, Larch and Yotov (2020), our conclusion is that gravity estimations can/should be performed without throwing away data. This will not only improve estimation efficiency but also allow for more proper quantification of the impact of trade policies whose effects may have been biased due to the 'arbitrary' dropping of observations for estimations with interval data.

5.4 Industry-level Gravity With and Without Domestic Trade

Simulated general equilibrium trade analysis has always *required* the use and modeling of domestic trade costs. More recently, the importance of domestic trade costs has been recognized and emphasized for gravity estimations too.[12], as well as non-discriminatory effects of trade policies (see Heid, Larch and Yotov, 2020). One of the main advantages of the ITPD-E is that it includes consistently constructed domestic trade for a large number of industries across all broad sectors of the economy. This feature of ITPD-E offers the unique opportunity of studying in detail the variation and determinants of industry-level domestic trade; we pursue this exciting agenda separately in future work. Instead, in our next experiment we compare the gravity estimates of the standard gravity variables that are obtained with the full ITPD-E (where we model domestic trade costs with time-invariant country-specific fixed effects) vs. estimates obtained by using only the international trade flow observations in the ITPD-E, which—mainly for want of reliable data—is still the standard practice in the trade literature.

Our findings are reported in Figure 5. Each panel of this figure compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. The first set of estimates is obtained with the PPML estimator and with all data from the ITPD-E whereas the second set of results is obtained also with the PPML estimator but only using the ITPD-E data on international trade flows, i.e. without domestic trade. Both sets of estimates are obtained with proper sets of exporter-time and importer-time fixed effects, according to specification (2). Each dot in each panel represents an estimate for a particular industry and estimates are ordered from smallest to largest PPML estimate with all data.

The main message from Figure 5 is that the industry-level estimates of the impact of the standard gravity covariates that we obtain with data that include domestic trade vs. data that only include international observations are almost identical. The explanation for this, and our main conclusion based on these results, is that the time-invariant country-specific $SMCTRY$ variables that we employ do a good job overall in capturing the impact of various domestic trade costs. Of course, by construction, the country-specific $SMCTRY$ dummies absorb the impact of all time-invariant domestic trade costs including, for instance, geography. However, our results imply that during the period covered by ITPD-E there has not been much time variation in domestic (relative to international) trade costs. A possible explanation for this result is that ITPD-E covers a period of two recessions and slow recovery after the great recession of 2008, where trade costs fluctuated a lot but remained unchanged on average.

[12]The importance of properly accounting for domestic trade costs is demonstrated in a series of papers including Anderson and Yotov (2010) who study the impact of Canada's Agreement on Internal Trade (AIT); Yotov (2012) who argues that the use of domestic trade ensures proper measurement of the evolving impact of distance and, thus, resolves the 'distance puzzle' in international trade; Dai, Yotov and Zylkin (2014) who use domestic trade to capture trade-diversion effects of regional trade agreements; Bergstrand, Larch and Yotov (2015) who rely on domestic trade to resolve the 'missing globalization puzzle' and to improve on the estimation of trade agreement effects; Ramondo, Rodríguez-Clare and Saborío-Rodríguez (2016) who demonstrate that the introduction of domestic trade frictions removes the counterfactual prediction that larger countries should be much richer than smaller ones; and Agnosteva, Anderson and Yotov (2019) who demonstrate that domestic trade costs are quite heterogeneous, even among Canada's provinces. Finally, the inclusion of domestic trade allows for identification of the effects of country-specific determinants of trade flows (see Beverelli et al., 2018) and country-specific determinants of border barriers (see Anderson et al., 2018)

5.5 Industry-level Multilateral Resistances

In their seminal paper Anderson and van Wincoop (2003) introduce the multilateral trade resistance and resolve the famous border puzzle for Canada's trade, c.f., McCallum (1995). More recently, Felbermayr and Yotov (2019) demonstrate that proper control for the multilateral resistances (MRs) also resolves the mystery of the excess trade balances, c.f., Davis and Weinstein (2002). As discussed earlier, and as captured by equation (2), the proper (theory-motivated) treatment of the MRs in disaggregated gravity regressions is with exporter-industry-time and importer-industry-time fixed effects in panel gravity regressions. However, for various reasons, both academic and policy researchers have used alternative sets of fixed effects to control for the MRs. For example, in regressions that are pooled across products/industries/sectors, we sometimes see exporter-time, importer-time and product/industry/sector fixed effects, respectively, instead of the theory motivated exporter-product/industry/sector-time and importer-product/industry/sector-time fixed effects. The objective of our last experiment is to see whether the treatment of MRs has an impact on ITPD-E's gravity estimates.

To this end we compare two sets of gravity estimates, which are presented in the lower part of Table 2. The estimates in Panel B of this table are obtained with the theory-motivated fixed effects (e.g., exporter-sector-time and importer-sector-time fixed effects) whereas the results in Panel C are obtained with exporter-time, importer-time and sector fixed effects instead. Comparison of estimates across Panels B and C reveals that some of the gravity estimates are not statistically different from each other, e.g. the estimates on $CLNY$, $LANG$, and FTA for Agriculture. However, we also observe some large and statistically significant differences. For example, the estimate of the effect of distance for Agriculture is more that 30 percent smaller when the MRs are not controlled for properly. The downward bias in estimated distance elasticities in Panel C is present for all broad sectors. We also observe a downward bias in the estimates on $SMCTRY$ across all broad sectors. Once again, the bias is most pronounced for Agriculture, followed by Manufacturing. Finally, we see biases in the estimates of the other gravity covariates, e.g. the impact of FTAs on trade in Mining and Energy. Based on these results, and consistent with Anderson and van Wincoop (2003) and Baldwin and Taglioni (2006), we conclude that proper control for the multilateral resistances is indeed important in the case of disaggregated gravity estimations.

6 Concluding Remarks

This paper validates the use of the newly-created ITPD-E from Borchert et al. (2020) for disaggregated gravity estimations by demonstrating that industry-level gravity ITPD-E results are comparable to existing benchmark gravity estimates obtained with aggregate data. A byproduct of our analysis is that we are the first to offer theory-consistent gravity estimates across a range of 170 industries that nearly completely describe an economy, and which are based upon a very large number of countries from across all world regions and per-capita income groups.

We also document wide but intuitive variation in the ITPD-E gravity estimates across Agriculture, Mining and Energy, Manufacturing, and Services, respectively. Specifically, our estimates reveal that geography (proxied by distance and contiguity) plays a stronger role for trade in Mining and Energy and Agriculture, while language is a stronger determinant of trade in Manufacturing and, even more so, in Services trade. We confirm that colonial relationships and FTAs promote trade. Interestingly, our estimates suggest that FTAs have played a more significant role in promoting trade in Agriculture and Mining and Energy, and less so in Manufacturing and in Services. Finally, our border estimates are the largest in Services, followed by Agriculture and Mining and Energy.

We employ the large number of industry-level gravity estimates to check the validity of some stylized facts from the related literature. Our results confirm that PPML and OLS deliver significantly different estimates. In addition, we find that PPML tends to lead to larger estimates for large effects and smaller estimates for small effects of the determinants of bilateral trade flows. Thus, the variation among the PPML estimates across industries is significantly larger. We also find that the presence of zeros, domestic trade, or using time intervals do not play significant roles for the estimates of standard gravity variables. Finally, we offer further evidence for the importance of properly controlling for multilateral resistances.

This paper presents some of the first results obtained using the new ITPD-E. We see many opportunities for further use of these data. For example, the ITPD-E offers the opportunity of quantifying the impact of various trade agreements, e.g. WTO, FTAs, etc., and to offer insights into their potentially heterogeneous impact across a wide range of industries or broad sectors. Since the ITPD-E includes domestic trade, which is one of its defining features, it could be used to study the determinants of domestic trade costs. We also believe that it would be beneficial if the latest econometric techniques and insights for gravity estimations are used to construct a fully balanced database corresponding to ITPD-E, which can be employed for general equilibrium simulation analysis. We plan to build such a database, which would constitute the second member within the family of ITPD databases. We may call this version of the database the International Trade and Production Database for Simulation (ITPD-S).

References

Agnosteva, Delina E., James E. Anderson, and Yoto V. Yotov. 2019. “Intra-national Trade Costs: Assaying Regional Frictions.” *European Economic Review*, 112(C): 32–50.

Anderson, James E. 2011. “The Gravity Model.” *Annual Review of Economics*, 3: 133–160.

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

Anderson, James E., and Eric van Wincoop. 2004. “Trade Costs.” *Journal of Economic Literature*, 42(3): 691–751.

Anderson, James E., and Yoto V. Yotov. 2010. “The Changing Incidence of Geography.” *American Economic Review*, 100(5): 2157–2186.

Anderson, James E., and Yoto V. Yotov. 2011. “Terms of Trade and Global Efficiency Effects of Free Trade Agreements, 1990-2002.” *NBER Working Paper No. 17003.*

Anderson, James E., Ingo Borchert, Aaditya Mattoo, and Yoto V. Yotov. 2018. “Dark Costs, Missing Data: Shedding Some Light on Services Trade.” *European Economic Review*, 105: 193–214.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. “New Trade Models, Same Old Gains?” *American Economic Review*, 102(1): 94–130.

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.

Baldwin, Richard E., and Daria Taglioni. 2006. “Gravity for Dummies and Dummies for Gravity Equations.” *NBER Working Paper No. 12516.*

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.

Beverelli, Cosimo, Alexander Keck, Mario Larch, and Yoto V. Yotov. 2018. “Institutional Quality, Trade and Development: A Quantitative Analysis.” *CESifo Working Paper No. 6920.*

Borchert, Ingo, Mario Larch, Serge Shikher, and Yoto V. Yotov. 2020. “The International Trade and Production Database.” *International Economics,* under review.

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

Cheng, I-Hui, and Howard J. Wall. 2005. “Controlling for Heterogeneity in Gravity Models of Trade and Integration.” *Federal Reserve Bank of St. Louis Review*, 87(1): 49–63.

Costinot, Arnaud, and Andrés Rodríguez-Clare. 2014. "Trade Theory with Numbers: Quantifying the Consequences of Globalization." Chapter 4 in the Handbook of International Economics Vol. 4, eds. Gita Gopinath, Elhanan Helpman, and Kenneth S. Rogoff, Elsevier Ltd., Oxford.

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

Dai, Mian, Yoto V. Yotov, and Thomas Zylkin. 2014. "On the Trade-diversion Effects of Free Trade Agreements." *Economics Letters*, 122(2): 321–325.

Davis, Donald R., and David E. Weinstein. 2002. "The Mystery of the Excess Trade (Balances)." *American Economic Review*, 92(2): 170–174.

Donaldson, Dave. 2016. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review,* forthcoming, available for download at http://economics.mit.edu/faculty/ddonald/publications.

Eaton, Jonathan, and Akiko Tamura. 1995. "Bilateralism and Regionalism in Japanese and U.S. Trade and Direct Foreign Investment Patterns." *NBER Working Paper No. 4758.*

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

Egger, Peter H., and Kevin E. Staub. 2014. "GLM Estimation of Trade Gravity Models with Fixed Effects." unpublished manuscript.

Egger, Peter H., Mario Larch, and Yoto V. Yotov. 2020. "Structural Gravity with Interval Data." unpublished manuscript.

Felbermayr, Gabriel, and Yoto Yotov. 2019. "From Theory to Policy with Gravitas: A Solution to the Mystery of the Excess Trade Balances." LeBow College of Business, Drexel University School of Economics Working Paper Series 2019-5.

Gurevich, Tamara, and Peter Herman. 2018. "The Dynamic Gravity Dataset: 1948-2016." *USITC Working Paper 2018-02-A*.

Head, Keith, and Thierry Mayer. 2014. "Gravity Equations: Workhorse, Toolkit, and Cookbook." Chapter 3 in the Handbook of International Economics Vol. 4, eds. Gita Gopinath, Elhanan Helpman, and Kenneth S. Rogoff, Elsevier Ltd., Oxford.

Heid, Benedikt, Mario Larch, and Yoto V. Yotov. 2020. "Estimating the Effects of Non-discriminatory Trade Policies within Structural Gravity Models." *Canadian Journal of Economics,* forthcoming.

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

Hinz, Julian, Mario Larch, Joschka Wanner, and Yoto V. Yotov. 2020. "PPML vs. OLS: Implications for Modeling Trade Costs in Structural Gravity Estimations." unpublished manuscript.

Larch, Mario, and Joschka Wanner. 2017. "Carbon Tariffs: An Analysis of the Trade, Welfare, and Emission Effects." *Journal of International Economics*, 109: 195–213.

Martin, William J., and Cong S. Pham. 2008. "Estimating the Gravity Equation when Zero Trade Flows are Frequent." *Working Paper No. 2008-03. Deakin University, School of Accounting, Economics and Finance.*

McCallum, John. 1995. "National Borders Matter." *American Economic Review*, 85(3): 615–623.

Olivero, María Pía, and Yoto V. Yotov. 2012. "Dynamic Gravity: Endogenous Country Size and Asset Accumulation." *Canadian Journal of Economics*, 45(1): 64–92.

Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez. 2016. "Trade, Domestic Frictions, and Scale Effects." *American Economic Review*, 106(10): 3159–3184.

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

Santos Silva, João M.C., and Silvana Tenreyro. 2011. "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator." *Economics Letters*, 112(2): 220–222.

Shikher, Serge. 2011. "Capital, Technology, and Specialization in the Neoclassical Model." *Journal of International Economics*, 83(2): 229–242.

Trefler, Daniel. 2004. "The Long and Short of the Canada-U.S. Free Trade Agreement." *American Economic Review*, 94(4): 870–895.

Yotov, Yoto V. 2012. "A Simple Solution to the Distance Puzzle in International Trade." *Economics Letters*, 117(3): 794–798.

Yotov, Yoto V., Roberta Piermartini, José-Antonio Monteiro, and Mario Larch. 2016. *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model.* Geneva, Switzerland, available for download at http://vi.unctad.org/tpa/index.html:United Nations and World Trade Organization.

Table 1: Industry-Level Gravity Estimates, OLS & PPML, 2000-2016

	Panel A. OLS Estimates						Panel B. PPML Estimates					
ID	DIST	CNTG	LANG	CLNY	FTA	SMCTRY	DIST	CNTG	LANG	CLNY	FTA	SMCTRY
1	-1.07***	1.1***	.44***	0.0300	.42***	5.820	-1.63***	.52**	0.180	1.14***	.37**	4.730
2	-.96***	1.3***	0.220	-0.220	.4**	9.960	-1.43***	1**	0.360	-1.46*	1.31**	10.47
3	-1.25***	.88***	0.110	-0.250	.37***	5.570	-2.25***	0.0400	-0.0300	-1.42**	.58**	4.200
4	-1.29***	.8***	0.0800	-0.320	0.110	5.490	-2.05***	-0.0300	.3**	-0.370	0.110	5.190
5	-.94***	1.15***	0.150	-.64*	.27*	.	-1.45***	.45*	.63**	-1.330	.88**	.
6	-.83***	1.19***	.35**	0.530	-0.160	5.610	-1.5***	0.570	-0.130	.77***	.75**	6.730
7	-.96***	.8***	.32***	0.120	.34***	6.050	-.97***	.57**	-0.100	0.0800	.39**	6.390
8	-1.14***	.81***	.29**	.51**	-0.140	.	-1.23***	.85***	0.0700	0.0500	0.140	.
9	-.9**	1.16**	-0.490	1.320	-0.580	7.830	-1.81**	2.92***	-0.710	1.87**	-0.510	9.430
10	-.49***	.79***	.55***	0.270	.26**	6.880	-.61**	0.0200	.65**	0.440	.7**	8.820
11	-.9***	.59***	.33***	.48**	.33***	5.570	-.91***	0.200	.56**	0.440	.45**	6.410
12	-1.21***	.92***	.73***	.96***	.79***	6.340	-1.06***	.54***	0.210	.86***	1.03***	6.310
13	-1.38***	.88***	.6***	.79***	.86***	6.610	-1.38***	.46***	.45***	1.61***	1.38***	6.310
14	-.88***	.57***	.45***	.63***	.43***	.	-1.09***	-0.170	0	.86***	.85***	.
15	-.86***	0.340	0.340	0.0700	0.0700	.	-2.65***	.68*	-1.23***	0.470	1.23**	.
16	-1.04***	.37**	.46***	0.170	.28***	4.290	-.89***	0.200	.46**	-0.650	0.0900	5.170
17	-1.22***	.9***	0.110	0.0200	.53**	.	-1.91***	1.59***	0.210	0.240	1.03***	.
18	-1.41***	1.36***	.58**	0.570	.84**	.	-2.38***	1.33***	.78**	0.520	2.37***	.
19	-1.17***	1.18***	0.0600	0.0800	.21*	5.820	-1.63***	.89***	0.250	0.280	.43**	6.260
20	-1.05***	1.24***	.37***	.57***	.25***	3.050	-1.27***	.93***	.7***	.92**	-.32*	2.320
21	-.47***	.56**	.29*	0.310	.56***	5.130	-.6**	.96**	-0.0200	-0.190	.6**	7.480
22	-.78***	.83***	.5***	.78***	.24***	4.950	-.81***	.65**	.52***	0.310	0.0100	6.170
23	-.5***	.64***	.26***	0.210	.24**	5.390	-1.09***	.61**	.48***	0.0800	0.180	6.060
24	-.39***	.27**	0.130	.86***	0.100	4.580	-.53***	0.0600	-0.0500	.81**	.41**	6.160
25	-.79***	.82***	.41***	.92***	.38***	5.030	-.77***	0.160	.42**	.89***	.27**	5.570
26	-1.06***	.75***	.43***	.77***	.35***	1.660	-.82***	0.0600	.25**	.42*	.55***	1.530
27	-1.02***	1.27***	-0.190	.75**	0.140	6.180	-2.01***	0.300	-0.0200	-0.250	0.120	3.660
28	-.79***	2.13***	0.0500	0.220	.55*	9.400	-2.14***	1.51**	-0.410	-0.300	1.3**	8.080
29	-1.53***	.5*	0.150	-0.540	.37**	4.970	-1.52***	0.350	0.180	0.300	.49**	3.350
30	-.73***	1.08**	0.260	.82*	0.420	6.870	-1.9***	.83**	.31**	.53**	0.340	4.070
31	-1.4***	1.04***	.56***	1***	.59***	5.290	-.71***	.37*	-0.200	1.84***	.5***	5.750
32	-1.97***	1.55***	0.370	-2.41***	.61**	6.080	-1.97***	1.11***	.73**	-2.33*	.48**	5.400
33	-.64***	0.0500	-0.0400	0.520	0.0800	13.53	-0.310	1.54***	-.77**	0.140	1.36**	15.04
34	-1.18***	1.05***	.63***	.69***	.75***	6.190	-1.03***	.55***	.33**	0.460	.81***	5.480
35	-1.03***	.96***	.64***	.88***	.46***	4.530	-.69***	.63***	0.190	.8***	.16*	4.160
36	-1.3***	.92***	.85***	1.29***	.64***	4.670	-.96***	.37***	.24**	.96***	.48***	4.080
37	-1.39***	1.31***	.6***	1.43***	.75***	4.710	-.83***	.58***	.44***	.58*	.41***	4.230
38	-1.38***	1.37***	.57***	1.01***	1.07***	6.250	-1.05***	.66***	.61***	0.270	.76***	5.480
39	-1.43***	1.26***	.7***	1.39***	.68***	6.280	-1.5***	.6***	.31**	.53**	.54***	4.450
40	-1.39***	1***	.55***	.49**	.65***	4.950	-.98***	.62***	0.0400	0.0800	.88***	4.380
41	-1.44***	.85***	.43***	.56**	.4***	5.430	-1.11***	.75***	.25**	0.120	.26**	5.350
42	-1.76***	1.02***	.98***	1.34***	.81***	5.770	-1.4***	.62***	.55***	1***	.88***	5.290
43	-1.23***	1.17***	.45***	1.45***	.47***	7.140	-1.19***	.88***	0.0300	1.05**	.79***	4.840
44	-1.5***	.74***	.9***	1.3***	1***	4.380	-.96***	.63***	.5***	.78***	.85***	4.450
45	-1.4***	1.27***	.88***	1.26***	.53***	5.720	-1.09***	1.1***	.48***	0.250	.52***	5.290
46	-1.52***	.8***	.84***	1.34***	.69***	4.450	-1.05***	.4***	.51***	.61***	.59***	4.130
47	-1.19***	.93***	.64***	.79***	.65***	6.120	-.86***	0.190	.35**	.48**	0.140	5.260
48	-1.02***	1.17***	.79***	1.23***	.65***	7.200	-.69***	0.250	.39**	1.57***	0.0800	7.150
49	-1.29***	1.47***	.84***	1.33***	.48***	7.680	-1.12***	.95***	.51**	1.06***	0.110	6.620
50	-1.38***	1.57***	.8***	1.25***	.72***	7.010	-1.25***	1.17***	.45**	.69**	.64***	5.930
51	-.94***	.49***	.34***	.47**	.74***	6.980	-1.13***	-0.300	.44**	1.12***	1.09***	5.100
52	-1.43***	.88***	.53***	.65***	.74***	3.970	-1.06***	.19**	.12*	.75***	.43***	3.260
53	-1.42***	1.26***	.69***	1.02***	.63***	3.920	-1.2***	.43***	.41***	.5**	.31**	3.400
54	-1.2***	.75***	.53***	.68***	.58***	4.300	-1.05***	0.140	0.150	0.300	.72***	3.610
55	-1.19***	.96***	.45***	.91***	.4***	4.310	-1.09***	.33**	0.0700	.75**	.39**	3.310
56	-1.53***	.68***	.75***	.7***	.64***	2.980	-1.1***	.21**	.17**	.49**	.32***	2.800
57	-1.45***	1.01***	.53***	.96***	.41***	4.150	-1.01***	0.180	.47***	0.450	.38***	4.120
58	-1.43***	.98***	.62***	1.19***	.5***	4.430	-.9***	.38**	.43***	.65***	.34***	4.620
59	-.76***	.72***	.28***	0.180	0.0600	5.490	-.58***	0.260	0.200	.69**	.29*	5.840
60	-1.14***	.74***	.44***	.45**	.32***	4.120	-.98***	0.150	-0.0200	0.450	0.0600	3.520
61	-1.28***	.82***	.57***	.8***	.3***	4.190	-0.140	.71***	-0.0800	.78***	0	6.410
62	-1.35***	1.39***	.55***	1.06***	.41***	4.670	-.91***	.41***	-0.120	.86***	.31**	4.470
63	-1.32***	.97***	.4***	1.03***	.25***	4.960	-1.37***	.86***	-0.0200	0.320	0.0300	3.870
64	-1.51***	.71***	.7***	.82***	.39***	3.650	-1.14***	.6***	.2*	.58**	.77***	3.490
65	-1.24***	.84***	.62***	1.32***	.37***	5.670	-.91***	.88***	.48***	0.530	.89***	6.160
66	-1.59***	1.12***	.47***	.61***	.35***	4.820	-1.62***	.64**	0.110	0.350	-0.320	3.230
67	-1.33***	.92***	.63***	1.02***	.28***	4.530	-.99***	.56***	0.130	.65***	.38***	4.670
68	-1.61***	.72***	.7***	.69***	.79***	3.630	-.98***	.38***	.15**	.37**	.66***	3.680
69	-2.01***	1.26***	.94***	1.23***	.51***	4.400	-1.59***	.63***	.42***	.85***	.53***	3.660
70	-1.82***	.99***	1.01***	1.05***	.7***	3.860	-1.38***	.52***	.38***	.61***	.56***	3.220
71	-1.45***	.77***	1.17***	1.76***	.56***	5.760	-.68***	.84***	.91***	1.03***	.3**	6.240
72	-1.72***	1.33***	1.44***	1.67***	.37***	7.720	-1.39***	.96***	1.12***	1.38***	.39**	7.250
73	-1.35***	.72***	.8***	1.16***	.64***	3.620	-.73***	.28**	0.0900	.64***	.87***	5.040
74	-1.54***	.56***	.86***	1.46***	.33***	4.330	-.94***	0.270	.32**	1.04***	0.120	4.110
75	-1.54***	1***	1.01***	1.27***	.46***	7.200	-1.23***	.83***	.24**	1.04***	.24*	6.170
76	-1.1***	.68***	.52***	.53***	0.0700	6.750	-.85***	.62***	.47***	.83**	0.210	6.810
77	-.91***	.42**	.35**	.62**	-0.0200	8.050	-1.18***	.44**	0.0400	.7**	-0.110	6.860
78	-1***	.9***	0.190	0.0300	0.180	7.600	-1.3***	.45**	0.190	0.0800	.46**	6.710
79	-1.99***	.99***	.55***	0.200	.39***	4.680	-1.28***	.36**	.39***	-0.200	0.130	3.820
80	-.92***	.68***	.3**	-0.150	-0.0800	5.720	-0.280	.73**	.47**	0.210	0.130	6.910
81	-1.39***	.9***	.64***	.53***	.56***	3.220	-.92***	0.0700	0.110	.49*	0.0700	3.190
82	-1.39***	1.16***	.25***	.79***	.49***	4.290	-.98***	.62***	.26**	0.0300	.46***	3.690
83	-1.61***	.53***	.62***	.32**	.6***	2.750	-.98***	.31***	0.110	0.130	.7***	3.060
84	-1.33***	.57***	.36***	.47**	.68***	3.600	-.96***	.34***	0.0400	0.0700	.72***	3.540
85	-1.65***	1.03***	.86***	.94***	.63***	4.320	-1.23***	.39***	.39***	.76***	.5***	3.880
86	-1.37***	.41***	.92***	.75***	.8***	3.930	-.47***	0.0100	0.110	.55***	0.180	5.130

Continued on next page

87	-1.73***	1***	.83***	.86***	.8***	3.720	-1.15***	.38***	.29***	.39**	.24**	3.320
88	-1.52***	.98***	.74***	1***	.52***	3.560	-.83***	.12*	.16**	.53***	.32***	3.870
89	-1.21***	.56***	.33***	0.200	.58***	3.420	-.79***	.24**	.18**	0.120	.42***	3
90	-1.14***	.77***	.7***	.3**	.49***	4.220	-.94***	.39***	.14**	0.130	.81***	3.700
91	-1.3***	.99***	.65***	.84***	.47***	3.970	-.79***	.25**	.15**	.74***	.62***	4.040
92	-1.77***	1.24***	.85***	1.04***	.75***	3.710	-1.14***	.39***	0.120	1.04***	.38***	3.620
93	-1.62***	1.14***	.77***	.88***	.64***	4.470	-1.08***	.46***	0	1***	.22**	4.180
94	-1.28***	.9***	.77***	1.11***	.42***	4.250	-.7***	.68***	0.120	1.12***	.36**	4.450
95	-1.32***	.85***	.5***	.57***	.21**	4.230	-.9***	.49***	.17*	.51*	.4***	4.270
96	-1.32***	1.29***	.84***	.98***	.49***	5.750	-1.06***	.94***	0.240	0.420	.58***	4.850
97	-1.71***	2.14***	.58***	.41**	.33***	6.580	-1.85***	.84***	.54***	0.150	.34**	4.400
98	-1.62***	1.44***	.87***	1.03***	.46***	6.530	-1.55***	.91***	.56***	1.29***	.68***	5.710
99	-1.11***	1.21***	.67***	.99***	.17**	5.810	-1.12***	.7***	0.0200	.7**	-.3*	4.570
100	-1.45***	.99***	.8***	.81***	.44***	4.050	-.9***	.38***	.36***	.6***	.58***	4.450
101	-1.54***	1.15***	.69***	.85***	.78***	3.950	-1.09***	.48***	.13**	.34**	.63***	3.280
102	-1.49***	.41***	.93***	.77***	.73***	3.270	-.86***	0.0800	.49***	.57**	.47***	2.970
103	-1.43***	.79***	.68***	1.21***	.47***	5	-.8***	.39***	.22**	.94***	.28***	5.020
104	-1.58***	.74***	.7***	1.29***	.58***	4.750	-1.19***	.48***	.49***	1.06***	.64***	4.480
105	-1.36***	.74***	.61***	.68***	.45***	4.940	-1.12***	.51***	.45***	.8***	.37***	4.240
106	-1.09***	.65***	.38***	1.08***	.32***	5.210	-.52***	.91***	.22*	.72***	.61***	5.620
107	-1.25***	.88***	.64***	1.1***	.48***	4.370	-.66***	.41***	0.120	.71**	.37***	4.580
108	-1.64***	1.17***	.85***	1.12***	.56***	3.400	-.94***	.47***	.21**	.71**	.35***	3.740
109	-.85***	.67***	.34***	.65***	.17**	5.320	-.52***	.35**	0.130	.87***	0.0900	5.220
110	-1.3***	.76***	.63***	.85***	.52***	2.780	-.6***	.38***	.19**	.68***	.47***	3.400
111	-1.16***	.7***	.67***	.8***	.43***	3.230	-.64***	.19**	.15**	.56***	.49***	3.250
112	-1.17***	.63***	.55***	.95***	.27***	3.630	-.7***	.43***	0.100	1.05***	.44***	3.990
113	-1.38***	.62***	.63***	.85***	.46***	3.320	-.81***	.31***	.16**	.75***	.51***	3.630
114	-1.37***	.75***	.74***	.9***	.53***	3.080	-.59***	.5***	0.120	.72***	.47***	3.860
115	-1.34***	.71***	.65***	.81***	.51***	3.950	-.84***	.35***	0.110	.61***	.6***	3.970
116	-1.3***	.7***	.65***	1.01***	.44***	2.840	-.57***	.27**	0.0500	.5**	.32***	3.690
117	-.93***	.69***	.24***	0.190	.29***	4.340	-.63***	.35**	.28**	.43**	.4***	4.520
118	-1.32***	.62***	.57***	.82***	.45***	2.870	-.78***	.33***	.14*	.5***	.35***	2.820
119	-1.24***	.5***	.59***	1.02***	.37***	3.680	-.85***	.3**	.28***	.73***	.17**	3.370
120	-1.14***	.66***	.5***	.91***	.34***	2.560	-.58***	.27**	0.160	.44**	0.150	2.790
121	-.74***	.59***	.25***	1.05***	-0.0400	5.760	-.24**	.35**	.34**	1.3**	.62***	6.440
122	-1.24***	.64***	.64***	.83***	.42***	3.240	-.64***	0.130	.18**	.53**	.2**	3.610
123	-1.48***	.78***	.85***	.67***	.66***	4.090	-1.08***	.25**	0.160	.64**	.52***	3.380
124	-1.24***	.54***	.5***	1.1***	.35***	4.690	-.78***	-.21**	0.0400	0.110	.27***	4.870
125	-1.28***	1***	.71***	1.01***	.45***	4.380	-.77***	.22**	.29***	.83***	.3***	4.390
126	-1.33***	.76***	.8***	1.3***	.5***	4.050	-.89***	0.0700	.28***	1.05***	.28***	3.860
127	-1.51***	1.03***	.96***	1.28***	.8***	4.430	-1.13***	.35***	.52***	1.11***	.53***	3.420
128	-1.31***	.93***	.71***	.76***	.68***	4.410	-1.03***	.18*	.16*	.52***	.3***	3.590
129	-1.35***	.93***	.68***	1.1***	.48***	4.050	-.72***	.43***	0.160	1.22***	.39***	4.360
130	-1.26***	.8***	.63***	.93***	.55***	4.460	-.76***	.21**	.2**	.6**	.54***	4.610
131	-1.14***	.9***	.58***	.88***	.35***	4.890	-.58***	-0.0400	.24**	.56***	0.0600	5.550
132	-1.13***	.45***	.51***	1.05***	.28***	4.270	-.74***	-0.0800	0.140	.47**	.25**	3.850
133	-1.26***	.68***	.55***	.85***	.48***	4.650	-.89***	0.0600	0.0600	0.200	.34***	4.200
134	-1.01***	.75***	.5***	1.15***	.44***	3.750	-.47***	-0.0100	.22**	.55***	.14*	4.120
135	-1.13***	.7***	.57***	1.06***	.39***	3.380	-.44***	.28**	.2**	.46***	.35***	4.080
136	-1.12***	.82***	.4***	.94***	.32***	4.120	-.79***	0.310	0.240	0.500	-0.110	4.280
137	-1.12***	.85***	.52***	.95***	.22***	4.620	-.27***	.39*	0	0.470	-0.0400	5.560
138	-1.12***	.66***	.61***	.55***	.74***	5.120	-.7***	.56***	-0.0600	0.250	1***	4.460
139	-1.44***	.54***	.6***	.95***	.48***	4.420	-1.49***	-.33*	0.240	1.66***	.89***	3.320
140	-1.26***	.88***	.71***	.76***	.7***	4.620	-.87***	.49***	-0.0100	.5**	.94***	4.100
141	-.52***	.79***	.2**	0.130	.4***	6.790	-.39***	.41**	0.120	0.240	.27**	5.260
142	-.66***	.83***	.22***	0.230	.34***	4.770	0.0800	1.2***	0.0400	-0.390	1.17***	6.920
143	-1.13***	.96***	.76***	.7***	.48***	5.440	-.73***	.54***	.65***	1.47***	1.22***	5.150
144	-.82***	.47***	.3***	.55***	.2***	4.890	-0.110	.97***	-0.0800	.76***	.26**	4.770
145	-1.06***	.91***	.42***	.58***	.58***	6.140	-.55***	.48***	0	0.200	.65***	5.650
146	-1.05***	1.18***	.36***	.82***	.39***	4.590	-.88***	.77***	-0.250	-0.230	.34**	3.940
147	-1.08***	.73***	.56***	1.02***	.22***	4.840	-1.01***	.49***	.3**	.84**	.26**	4.630
148	-1.47***	.93***	.78***	1.05***	.61***	4.550	-1.05***	.49***	0.150	.91***	.49***	4.600
149	-1.01***	.77***	.72***	.81***	.34***	6.210	-.43***	0.110	.28*	-0.240	.44**	7.010
150	-.96***	.71***	.36***	.92***	0.0700	3.710	-.28***	.56***	-0.0500	0.0300	.38***	4.960
151	-1.17***	.92***	.43***	1.28***	.09*	3.050	-.65***	.55***	0.110	0.230	.24*	3.530
152	-1.25***	.88***	.55***	1.12***	.45***	3.530	-1.16***	-0.0600	-0.140	.71**	0.0300	2.660
153	-1.44***	1.1***	.71***	1.12***	.49***	3.720	-.93***	.33***	0.0300	.77***	0.0200	3.710
154	-1.41***	.39**	-0.100	1.65***	0.0900	.	-1.67***	-0.180	-0.280	2.09***	0.210	.
155	-.91***	.29**	.27**	0.120	.35**	.	-.59***	-0.0100	.62***	-.61**	.34**	.
156	-1.21***	.41**	.54***	1.11***	.13*	4.110	-.74***	.17*	.24**	.44**	0.0200	5.180
157	-1.02***	.88***	.61***	1.72***	.41***	3.930	-1.01***	.58***	.46***	1.02***	0.120	4.070
158	-1.2***	.57***	0.0600	0.0900	0.110	6.430	-.8***	.26*	.29*	-0.0100	.84***	8.400
159	-.99***	.5***	.53***	.5**	.24**	6.960	-.67***	0.160	.47**	-0.180	0.320	8.390
160	-1.2***	0.200	.55***	.61**	0.160	5.570	-.71***	-0.170	.4**	0.280	0.210	7.130
161	-.89***	.23*	.36***	.68**	0.0600	.	-.47***	-0.170	0.160	-0.430	0.140	.
162	-1.13***	.32**	.38***	1.22***	.2**	5.530	-.78***	-0.0500	0.0800	.98***	.64***	6.680
163	-1.22***	0.170	.38***	1.26***	.23***	4.610	-.61***	0.160	0.100	.5**	.4***	6.630
164	-.96***	.36**	.35**	-0.390	2.45**	9.670	-1.07***	1.04***	.43*	2.75**	1.28***	10.42
165	-1.16***	.4**	.45***	1.36***	.27**	8.230	-1.22***	.61**	.58***	.96**	.5**	8.950
166	-.96***	0.150	.53***	.76***	.38***	7.540	-1.25***	0.0900	.39**	0.300	0.360	7.050
167	-.57***	-0.0700	.41***	.79***	.24**	.	-.44***	0.0400	-0.130	.93***	.41**	.
168	-1.1***	.36**	.33***	1.45***	.37***	.	-.7***	.22**	.21**	.81***	.53***	.
169	-1.17***	.32**	.24***	0.290	-0.150	6.970	-.84***	0.220	-0.0900	0.0600	.45**	8.310
170	-.29**	.31*	.31**	-0.350	.94**	9.730	-.36*	.46*	.44**	0.420	1.6***	11.50

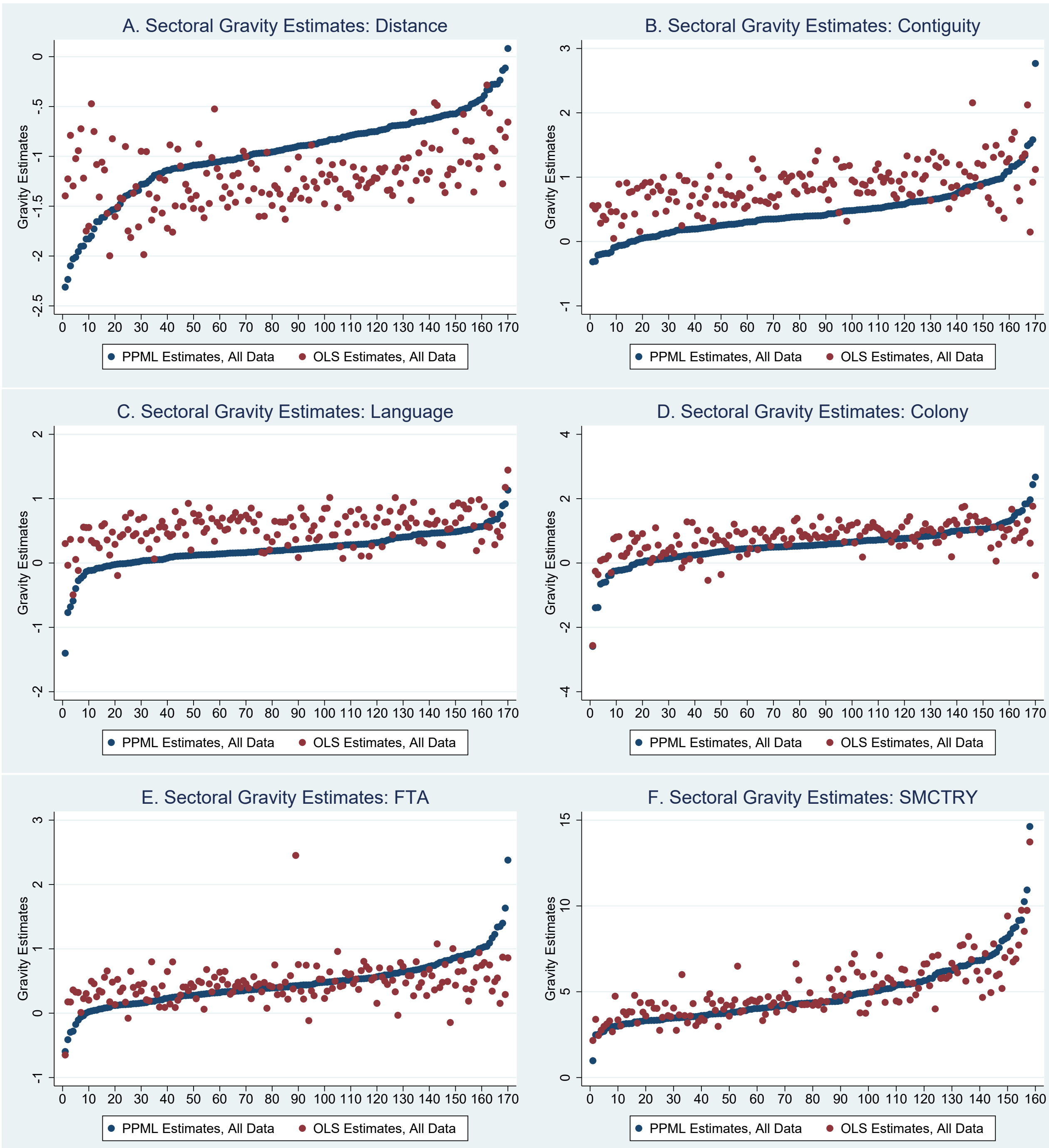
Notes: This table reports OLS gravity estimates (in Panel A) and PPML gravity estimates (in Panel B) that are obtained with the full ITPD-E. The dependent variable in Panel A is the logarithm of nominal bilateral trade flows, while the dependent variable in panel B is nominal trade in level. All estimates are obtained with exporter-time and importer-time fixed effects, whose estimates are omitted for brevity. We also omit for brevity the standard errors and t-statistics of the estimates. Instead, we just report significance levels as follows ** $p < 0.10$, *** $p < .05$, **** $p < .01$. The SMCTRY indexes in each panel are obtained as averages across the country-specific SMCTRY estimates that are obtained from each specification. The full set of estimates and standard errors are available by request from the authors.

Table 2: Broad Sector Gravity Estimates

	(1) Agriculture	(2) Mining&Energy	(3) Manufacturing	(4) Services
A. Industry-level PPML Averages				
DIST	-1.335	-1.509	-0.898	-0.819
	(.574)	(.718)	(.326)	(.335)
CNTG	0.618	0.859	0.433	0.202
	(.64)	(.542)	(.287)	(.322)
LANG	0.200	-0.0260	0.229	0.257
	(.438)	(.493)	(.205)	(.256)
CLNY	0.294	-0.0100	0.611	0.606
	(.832)	(1.249)	(.372)	(.848)
FTA	0.588	0.656	0.415	0.492
	(.593)	(.48)	(.284)	(.413)
SMCTRY	6.086	6.479	4.476	7.726
	(2.088)	(4.105)	(1.098)	(2.071)
B. Pooled Industry-level PPML Estimates, All Data				
DIST	-1.078	-1.383	-0.837	-0.717
	(0.028)**	(0.101)**	(0.012)**	(0.025)**
CNTG	0.548	0.429	0.328	0.187
	(0.051)**	(0.187)*	(0.020)**	(0.044)**
LANG	0.231	0.156	0.184	0.221
	(0.042)**	(0.124)	(0.018)**	(0.043)**
CLNY	0.667	0.532	0.498	0.588
	(0.098)**	(0.256)*	(0.039)**	(0.080)**
FTA	0.637	0.437	0.379	0.272
	(0.041)**	(0.119)**	(0.020)**	(0.044)**
SMCTRY	5.984	4.883	4.283	7.493
	(3.155)	(2.793)	(2.144)	(2.983)
N	2685599	577963	34618475	632373
C. On the Importance of the Multilateral Resistances				
DIST	-0.742	-1.217	-0.706	-0.688
	(0.053)**	(0.100)**	(0.024)**	(0.034)**
CNTG	0.445	0.412	0.380	0.176
	(0.105)**	(0.198)*	(0.043)**	(0.063)**
LANG	0.215	0.131	0.229	0.254
	(0.094)*	(0.172)	(0.043)**	(0.060)**
CLNY	0.640	0.505	0.398	0.530
	(0.178)**	(0.284)+	(0.077)**	(0.103)**
FTA	0.667	0.253	0.400	0.364
	(0.070)**	(0.142)+	(0.037)**	(0.076)**
SMCTRY	4.132	4.554	3.619	7.004
	(2.109)	(2.634)	(1.77)	(2.801)
N	2685696	578166	34619387	634190

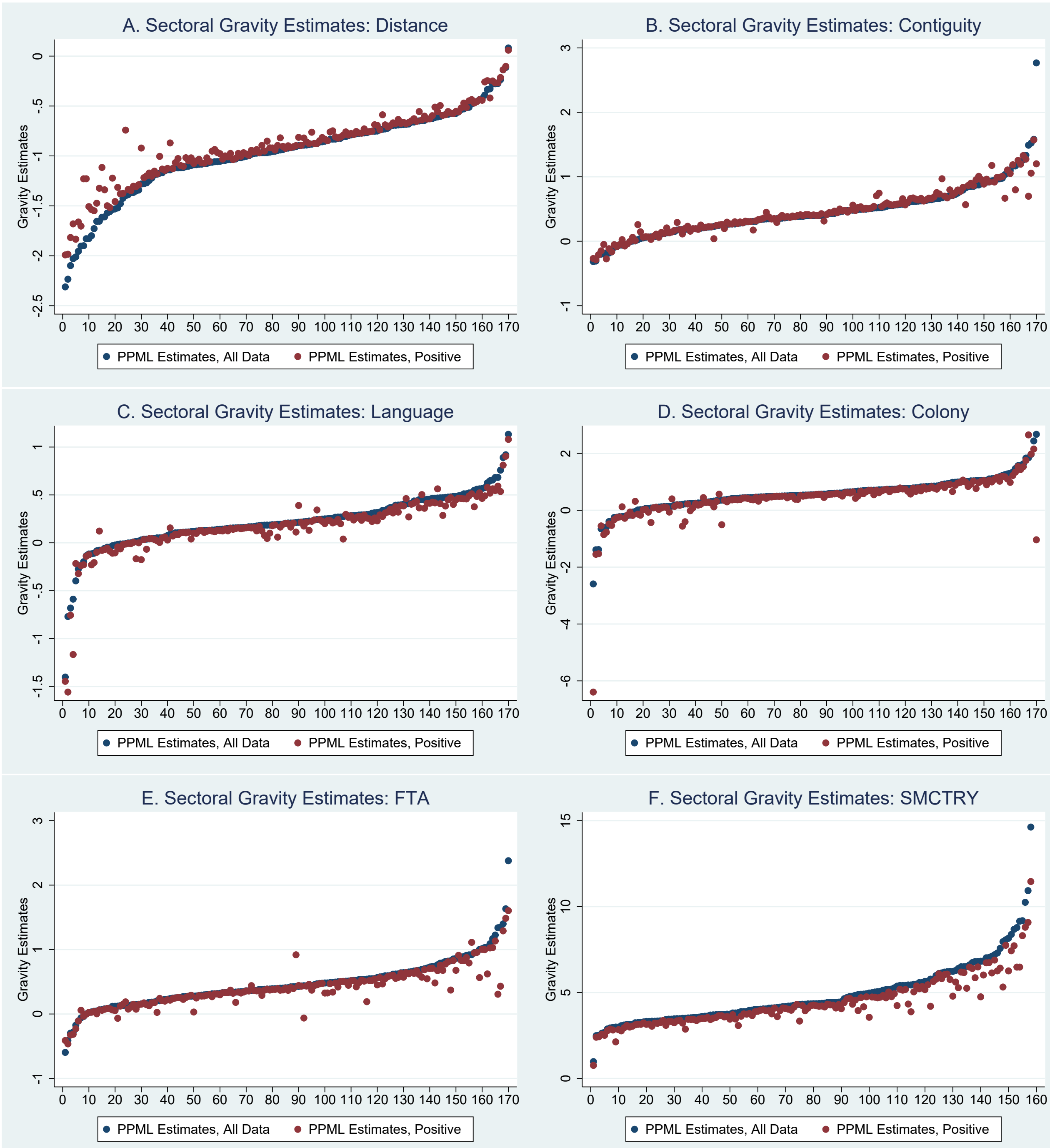
Notes: This table reports gravity estimates for the four main industries (Agriculture, Mining and Energy, Manufacturing, and Services) in the ITPD-E. Panel A reports simple averages and standard deviations (in parentheses), which are obtained from the industry-level estimates in Table 1 for each of the main ITPD-E broad sectors. Panel B reports gravity estimates for the same four broad sectors, which are obtained with pooled industry-level data and from panel PPML gravity specifications that correspond to equation (2), and where, in accordance with theory, the set of fixed effects are exporter-sector-time and importer-sector-time. The SMCTRY indexes in panel B are averages across the underlying country-specific SMCTRY estimates and their standard errors are reported in parentheses. Finally, panel C reproduces the results from panel B but with exporter-time and importer-time fixed effects instead of with exporter-sector-time and importer-sector-time fixed effects. Once again, the SMCTRY indexes in panel C are averages across the underlying country-specific SMCTRY estimates and their standard errors are reported in parentheses. The levels of significance of the estimates in panel B are denoted as follows + $p < 0.10$, * $p < .05$, ** $p < .01$. See text for further details.

Figure 1: Structural Gravity Estimates, PPML vs. OLS



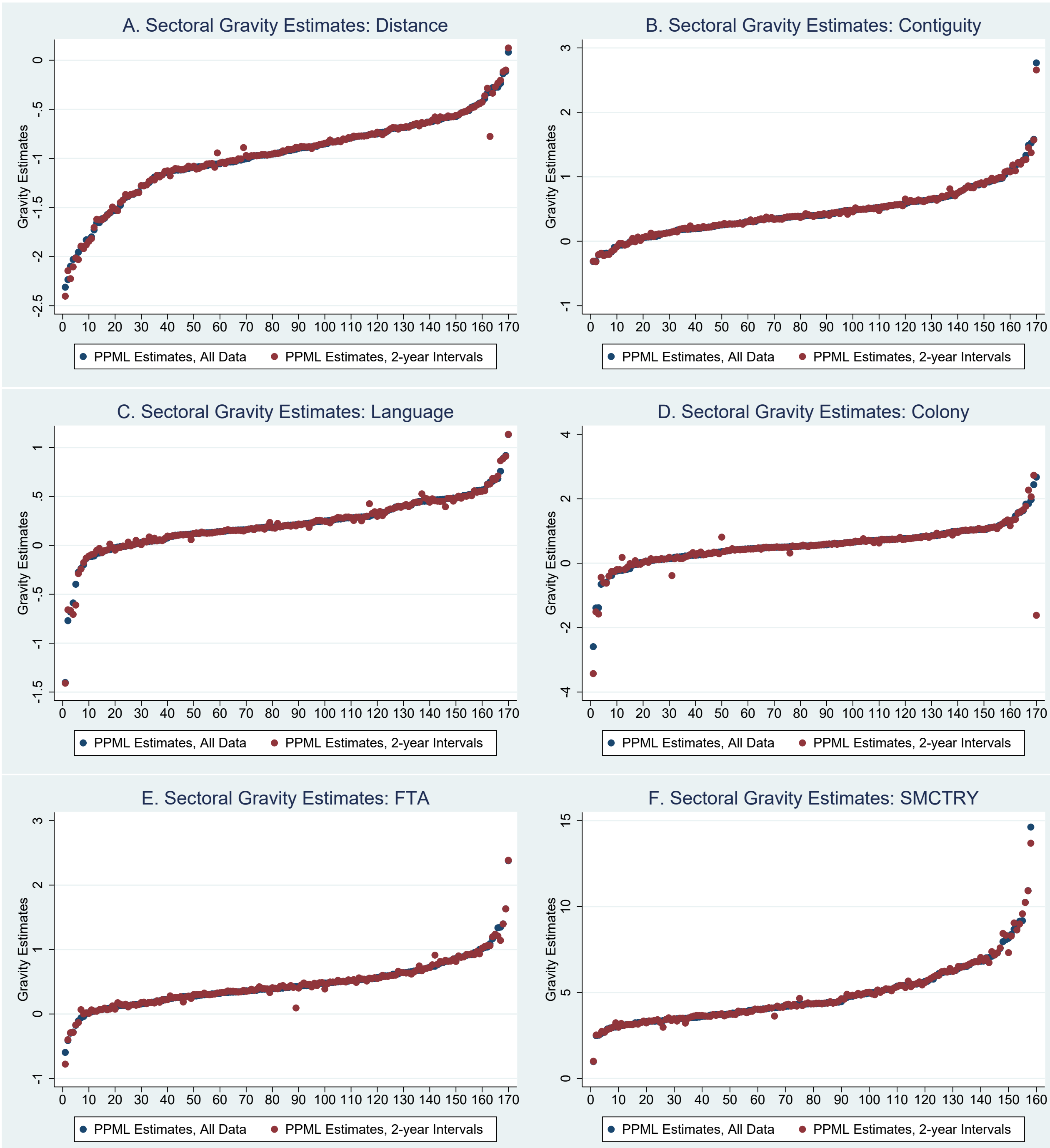
Notes: Each panel of this figure compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. Both sets of estimates are based on the complete ITPD-E. The first set of estimates is obtained with the PPML estimator while the second set is obtained with OLS. Each dot in each panel represents an estimate for a particular industry and the estimates are ordered from the smallest to the largest PPML estimate with all data. See text for further details.

Figure 2: Structural Gravity Estimates, PPML: All Data vs. Positive Values Only



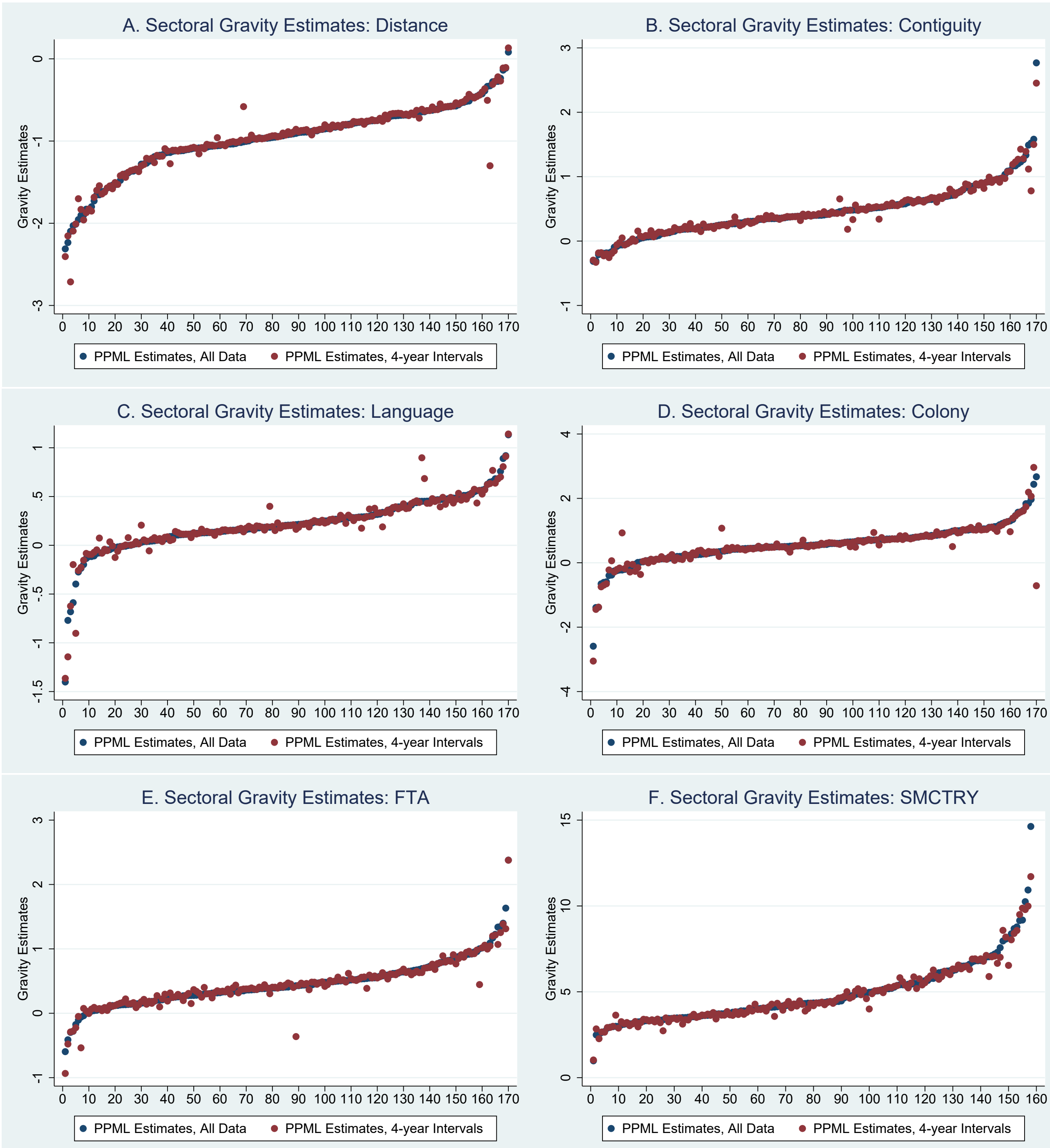
Notes: Each panel of this figure compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. The first set of estimates is obtained with the PPML estimator and with all data from the ITPD-E, while the second set of results is obtained also with the PPML estimator but data on positive trade flows only. Each dot in each panel represents an estimate for a particular industry and the estimates are ordered from the smallest to the largest PPML estimate with all data. See text for further details.

Figure 3: Structural Gravity Estimates, PPML: All Data vs. 2-year Intervals



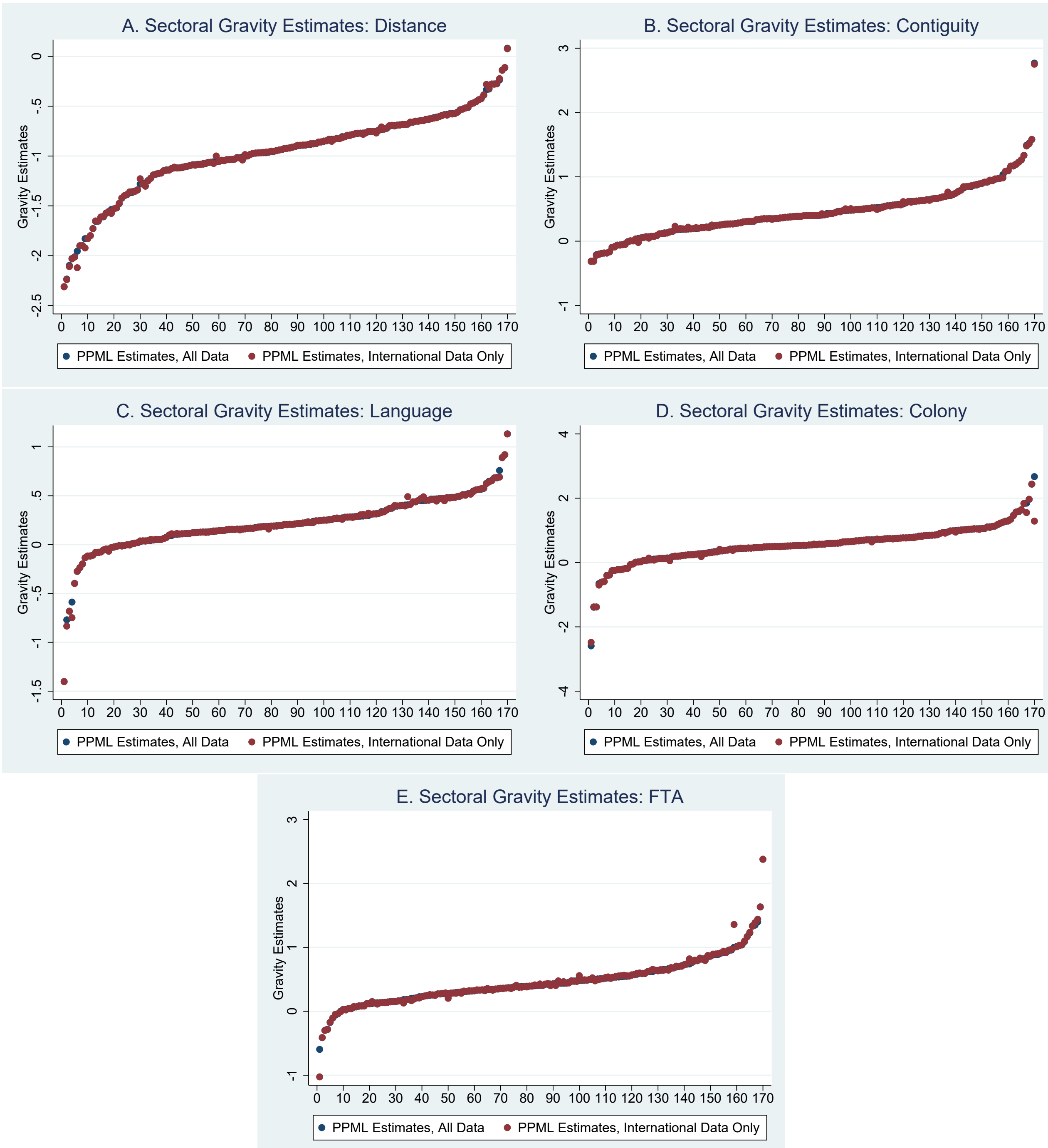
Notes: Each panel of this figure compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. The first set of estimates is obtained with the PPML estimator and with all data from the ITPD-E, while the second set of results is obtained also with the PPML estimator but with 2-year interval data. Each dot in each panel represents an estimate for a particular industry and the estimates are ordered from the smallest to the largest PPML estimate with all data. See text for further details.

Figure 4: Structural Gravity Estimates, PPML: All Data vs. 4-year Intervals



Notes: Each panel of this figure compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. The first set of estimates is obtained with the PPML estimator and with all data from the ITPD-E, while the second set of results is obtained also with the PPML estimator but with 4-year interval data. Each dot in each panel represents an estimate for a particular industry and the estimates are ordered from the smallest to the largest PPML estimate with all data. See text for further details.

Figure 5: Structural Gravity Estimates, PPML: Domestic vs. International Data



Notes: Each panel of this figure compares two corresponding sets of structural gravity estimates for a given standard gravity covariate. The first set of estimates is obtained with the PPML estimator and with all data from the ITPD-E, while the second set of results is obtained also with the PPML estimator but only using the ITPD-E data on international trade flows, i.e., without domestic trade. Each dot in each panel represents an estimate for a particular industry and the estimates are ordered from the smallest to the largest PPML estimate with all data. See text for further details.