

Working Paper Series

No. 11-2024

Ingo Borchert

Mario Larch

Serge Shikher

Yoto V. Yotov

Abstitr 3202 34 0 26.m3 (V) Lie1 (02.3 -1.nt) 2102.duc I (1.nt) E2 (6at) B -1. (02.n **Estimation**

JEL codes: F10, F14, F63, O11, O19

Key words: Trade dataset, Domestic trade, Globalization, Inequality, Disaggregated simulation analysis, Gravity

Globalization, Trade, and Inequality: Evidence from a New Database

Ingo Borchert Mario Larch[.] Serge Shikher Yoto V. Yotov[¶]

May 15, 2024

Abstract

We introduce the International Trade and Production Database for Simulation (ITPD-S) and use it, in combination with the International Trade and Production Database for Estimation (ITPD-E), to quantify the impact of globalization on bilateral trade, real income, and inequality in the world at the industry level in 1990-2019. To perform the analysis we rely on a new quantitative trade model, which enables us to estimate the magnitude of globalization and then perform a counterfactual analysis of the impact of globalization on real output within the same framework. Our estimates reveal that, on average, bilateral globalization forces have led to a remarkable increase in international trade of about 570%, between 1990 and 2019, with very wide but intuitive variation across industries. Our counterfactual analysis reveals that globalization has beneted most countries but relatively more so smaller and more open economies, which are typically developing countries. As a result, this `catch-up' implies less cross-country income inequality.

JEL codes: F10, F14, F63, O11, O19

Keywords: Trade dataset, Domestic trade, Globalization, Inequality, Disaggregated simulation analysis, Gravity

The data that support the ndings of this study are openly available from the USITC Gravity Portal at [https://gravity.usitc.gov.](https://gravity.usitc.gov)

[&]quot; University of Sussex Business School, CITP, UKTPO.

[…]University of Bayreuth, CEPII, CESifo, GEP, ifo Institute.

 \tilde{Y} United States International Trade Commission. The views expressed in this paper are strictly those of the authors and do not represent the opinions of the U.S. International Trade Commission or of any of its Commissioners.

[¶] Drexel University, ifo Institute, CESifo.

1 Introduction

Trade policy is best analyzed in a model that takes into account trade diversion and price

a combination of two datasets that are perfectly compatible by design, one of which can be used for estimation and the other for simulation.

Third, we deploy the two ITPD databases to study the impact of globalization. We dene

to the best-performing methods when lling in missing trade values. Using these methods we can create a complete set of domestic trade observations for 198 countries, plus an additional 67 countries for which some, but not all, domestic trade observations are available.

To perform the empirical analysis, we rely on a well-established new quantitative trade model, which, as demonstrated [by Arkolakis et a](#page-40-0)l[. \(201](#page-40-0)2), is representative of a wide family of trade models. New quantitative trade models use structural gravity to explain trade. In addition to its intuitive appeal and solid theoretical foundations, an attractive feature of the gravity system is that it nests the theoretical foundation for the estimating gravity equation, which delivers our partial equilibrium estimates of the eects of globalization.

In addition to the ITPD-E database, which we use for our estimation analysis, and the ITPD-S database, which we use for the counterfactual analysis, we utilize several other datasets, including the Regional Trade Agreement dataset [of Egger and Larch \(20](#page-41-0)08) for data on RTAs, the Dynamic Gravity Database of the United States International Trade Commission [\(Gurevich and Herman, 2018](#page-41-1)a) for data on WTO and EU membership, the Global Sanctions Database (Felbermayr et al., 2020; Syropoulos et al., 2023) for data on complete and partial trade sanctions, and the classication of countries by income level of the World Bank (year 2000).

Several noteworthy ndings stand out from our estimation results. First, overall, we obtain very large, positive, and statistically signi cant estimates of the e ects of globalization on trade. Speci cally, only 5 of our estimates are negative, while 93% of the positive estimates are also statistically signi cant. In terms of magnitude, our estimates imply that, on average, bilateral globalization forces (other than trade agreements, WTO membership, and EU membership) have led to a remarkable increase of 570% in international trade relative to domestic sales over the period 1990-2019.

Second, the globalization estimates that we obtain manifest in a very heterogeneous way across broad ITPD sectors. Our estimates suggest that the services sector has experienced the largest impact of globalization, followed by manufacturing, and then agriculture. We

3

also see signi cant heterogeneity of globalization e ects across industries within each broad sector. Thus, for example, the services categories that have experienced the largest e ects are `Health' services and `Travel' services, while the smallest eects are for `Transport' services and `Trade-related' services. Finally, out of ve negative globalization estimates the only statistically signi cant negative e ect is for `Cutting shaping and nishing of stone' while the largest negative estimate that we obtain is for the industry `Publishing of newspapers journals etc.' We nd the latter result intuitive.

We also o er a preliminary investigation for possible heterogeneous e ects of globalization depending on the country income group. To this end, we rely on the 2000 income classication of the World Bank to identify the `High Income' countries in our sample, and we obtain estimates of the e ects of globalization for that subsample of rich economies only. Overall, we do not observe systematic dierences in the eects of globalization for the rich countries. One possible explanation for this result is mechanical; i.e., due to the disproportional size and trade shares of these large countries, our average results may be driven by th[es](#page-5-0)e large high-income countries. Moreover, we see some intuitive variation across the four Heale e437(fee)-ll7(osi27(en)ld3o2(sh6esour))27(e)h6esntries. tries. 3(b-407esour)counsh6e e ects, both across the countries and across the industries in our sample. Second, we observe substantial heterogeneity in our estimates in both dimensions, with the heterogeneity across countries even more pronounced. Third, a closer look at the heterogeneous e ects across countries reveals that developing, smaller countries seem to have gained relatively more from opening up to international trade. This is also supported by a calculation of the Gini index over current cross-country output and the counterfactual output without globalization, suggesting that world inequality has decreased due to the globalization forces that we identify. The fact that some small, open economies have gained the most in percentage terms implies that inequality has risen within that group of economies as some have bene tted more than others, but importantly `the pie has grown' overall and the advances of smaller countries mean that global inequality has fallen overall across all countries.

The rest of the paper is organized as follows. Section 2 describes the methods that we use to construct the International Trade and Production Database for Simulation (ITPD-S), showcases its main features, and discusses potential caveats with its use. Section 3 oers a brief review of the new quantitative trade model, which we rely upon to obtain our partial equilibrium estimates and for the counterfactual analysis as well. Section 4 presents our partial equilibrium estimates of the e ects of globalization, translates them into real output e ects, and discusses our main ndings. Section 5 summarizes our main contributions and oers directions for future work. The Appendix includes more detailed descriptions of the procedures that were used to construct the ITPD-S, and includes some additional estimates and results.

2 The ITPD-S

The International Trade and Production Database for Simulation, or ITPD-S, that underpins this paper is based on the International Trade and Production Database for Estimation, or ITPD-E (Borchert et al., 2021, 2022). ITPD-E uses only raw administrative data, which

5

mak[e](#page-7-0)s it suitable for estimation, but also means that it has many missing values Yet simulation of new quantitative trade models requires data that is complete i.e., non-missing in all relevant dimensions. The ITPD-S meets these requirements and thus allows researchers to perform simulations with a variety of partial equilibrium (PE) and general equilibrium (GE) models, including the the structural gravity model. In combination, ITPD-E and ITPD-S provide researchers and policy analysts with mutually consistent databases for estimation and simulation; in particular, estimation of simulation parameters can be done with ITPD-E with comparable o cial data that exhibits the same dimensionality as the envisaged simulation exercise.

By taking the latest version of ITPD-E as a starting point, ITPD-S inherits its high granularity. It includes international and domestic trade data for 265 countries, 170 industries across all broad sectors (agriculture, mining and energy, manufacturing, and services), and 34 years³. In addition, ITPD-S Ils in most domestic trade values that are missing in ITPD-E using the methodology explained in the next section.

2.1 ITPD-S Methodology

The rst step in the construction of ITPD-S^{[4](#page-7-2)} is the creation of a blank database with 265 exporters and importers, 270 industries, and the year dimensions matching ITPD-E-R02: 1986-2019 for agriculture, 1988-2019 for mining and energy, 1988-2019 for manufacturing, and 2000-2019 for services. The resulting blank database is square in the exporter and importer dimension for each industry and year. Following ITPD-E, countries in ITPD-S are dened by the USITC's Dynamic Gravity Dataset [\(Gurevich and Herman, 2018](#page-41-1)a). Industries in ITPD-S R01 follow the de nitions in ITPD-S R02.

The blank database is then populated by international trade entries from ITPD-E-R02.

²The current version of ITPD-E (Release R02) covers 265 countries, 170 sectors, and over 30 years for

Similar to ITPD-E, ITPD-S contains a ag variable (flag _ zero) that is equal to `r' for observations with zeroes coming from original data sources, `p' for observations with positive trade ows, and `u' for observations lled with zeroes. As in ITPD-E, all trade observations that are not reported by either importer or exporter are assumed to be zero and denoted by the appropriate ag. Considering that reported international trade ow statistics as taken from ITPD-E are quite comprehensive, we believe that this is a plausible assumpt[io](#page-8-0)nt is well known, after all, that the international trade ow matrix in its entirety is indeed sparse.

While ITPD-E includes many missing domestic trade observations, considerable e ort is made in ITPD-S to ll those missing observations. Missing domestic trade ows are more

Next, we use forward and backward ll methods, i.e. we carry forward in time the last observed value and backward in time the rst observed value for a maximum of seven years to avoid over-reliance on individual data points. Domestic trade values obtained using forward and backward ll are denoted with ags 4 and 5, respectively.

We aim to recover domestic trade ows that are still missing at this stage by deploying estimation and projection methods as described in the next section.

2.1.2 Econometric methods

This set of methods for predicting missing domestic trade ows relies on state-of-the-art structural gravity models [\(Anderson and van Wincoop, 200](#page-40-1)3; Yotov et al., 2016, see). This estimate and exploit for prediction. Thus, we devise two approaches for recovering estimates of domestic trade costs, or equivalently, border frictions: i) we proxy for trade costs using observables, and ii) we rely on the panel structure in ITPD-E to proxy for trade costs using time-invariant bilateral xed e ects in combination with utilizing di erent aggregations and common xed e ects for these aggregates.

In the rst approach, structural gravity estimation is deployed at the level of each individual industry in ITPD, whereby in addition to the xed e ects structure trade costs are proxied by a set of 10 bilateral time-varying observables that comprehensively cover geogra-phy, policy and institutions.^{[7](#page-10-0)} Within that baseline framework, we estimate ve alternative speci cations for modelling domestic trade costs:

- 1. one common, time-invariant border e ect for all countries;
- 2. time-varying border e ects common for all countries;
- 3. time-invariant but country-speci c border e ects;
- 4. a border e ect that is allowed to vary with observable country characteristi[cs](#page-10-1);
- 5. a time-varying border eect that is allowed to vary with observable country characteristics.

To see the intuition for these modeling choices, consider the following example: if there were no `gross output' statistics for ITPD industry 148 (Furniture) for Bolivia, domestic trade for that industry will be missing across all years. Yet, if it were possible to estimate a border eect and its variation with internal distance, market size etc. based upon data for countries that do report gross output for Furniture (specication 4 above), then we can use that coecient, combined with values for observable characteristics such as distance, GDP, etc. for Bolivia, to predict Bolivian domestic trade in industry 148, suitably adjusted

⁷The complete list of observables entails bilateral distance, contiguity, common language, common legal origin, common religion, common colonial past, joint EU membership, joint WTO membership, and whether countries are signatories to a customs union or any kind of preferential trade agreement at time respectively. The data come from USITC's Dynamic Gravity Dataset and from CEPII.

⁸We employ four proxies for capturing domestic trade costs: internal distance, degree of religious homogeneity within a country, log GDP for market size, and log GDP per capita for stage of development.

- ˆ Flag=13: Extends the nal data by lling in the remaining missing observations by interpolation.
- ˆ Flag=14: Extends the nal data by lling in remaining missing observations by forward ll.
- ˆ Flag=15: Extends the nal data by lling in the remaining missing observations by backward ll.
- 2. Cross-sectional estimation methods.
	- ˆ Flag=21: Time-unvarying common border eect for all countries (model 1)
	- ˆ Flag=22: Time-varying common border eect for all countries (model 2)
	- ˆ Flag=23: Time-unvarying country-specic border eect (model 3)
	- ˆ Flag=24: Border eect proxied by country characteristics (model 4)
	- ˆ Flag=25: Border eect proxied by country characteristics interacted with year xed eects (model 5)
- 3. Panel estimation methods.
	- ˆ Flag=31: 170 industries (level 1)
	- ˆ Flag=32: 26 industry groups (level 2)
	- ˆ Flag=33: 15 industry groups (level 3)
	- ˆ Flag=34: 11 industry groups (level 4)
	- ˆ Flag=35: 4 broad sectors (level 5)
	- ˆ Flag=36: 1 economy, i.e. all industries combined (level 6)

2.2 Evaluation of estimation methods and our procedure

Since various simple and econometric methods described above can be used to ll in missing domestic trade observations, it is imperative to evaluate these methods to determine how

6. Set outliers to missing

7. Simple methods with ags 13-15

2.3 Summary of estimated domestic trade observations

This section summarizes the provenance of domestic trade observations in ITPD-S. Two sets of results are presented. The rst includes all countries in ITPD-E. The methods used to ll in missing domestic trade data in ITPD-S cannot estimate all missing observations. Some missing observations cannot be estimated because not enough information is available. However, there are 189 countries for which all missing observations are estimated. The second set of results focuses on just those countries.

Table 1 shows the results of lling in missing observations in all countries, industries, and years of ITPD-E. There are 1,395,530 domestic trade observations in ITPD-E. Of those, 162,865 have data and 1,232,665 are missing. Of all missing observations, 1,198,129 are estimated and 34,543 cannot be estimated.

Table 2 shows the results of lling in missing observations in 189 countries with the complete set of observations. There are 1,048,900 domestic trade observations in those countries. Of those observations, 162,511 have data and 886,389 are missing and estimated. Simple methods with ags 2-5 provide 525,972 estimates, gravity models provide 323,366 estimates, and post-estimation simple methods with ags 13-15 provide another 37,045 estimates. The online appendix shows the list of 69 countries with missing observations that could not be estimated.

2.4 List of Variables

The variables included in ITPD-S are shown in Table 3. Most of the variables are carried over from ITPD-E. The only addition is flag _ itpds which shows the provenance of domestic trade values.

13

| category or ag | count |
|-------------------|----------------|
| all observations | 1,048,900 |
| missing | 886,389 |
| estimated | 886,389 |
| not estimated | 0 |
| 1 | 162,511 |
| $\overline{2}$ | 200,749 |
| 3 | 187,255 |
| $\overline{4}$ | 119,444 |
| 5 | 18,524 |
| 31 | 23,299 |
| 23 | 33 |
| 32 | 103,159 |
| 33 | 21,071 |
| 34 | 9,446 |
| 35 | 34,906 |
| 36 | 131,452 |
| 24 | 0 |
| 21 | $\mathbf{0}$ |
| 22 | $\overline{0}$ |
| 25 | $\overline{0}$ |
| 13 | 6 |
| 14 | $\overline{0}$ |
| 15 | 37,045 |
| | |

Table 2: Domestic trade observations in 189 countries with a full set of observations

Table 3: Variables in ITPD-S-R01

| Variable description |
|--|
| ISO 3-letter alpha code of the exporter |
| Name of the exporter |
| ISO 3-letter alpha code of the importer |
| Name of the importer |
| exporter_dynamic_code Dynamic alpha code of the exporter based on DGD |
| Dynamic alpha code of the importer based on DGD |
| Year |
| ITPD industry code |
| ITPD industry description |
| Broad sector description |
| Trade ows in million of current US dollars |
| |

framework is that it nests the theoretical foundation for the estimating gravity model, which will deliver our estimates of the e ects of globalization.

Capitalizing on the power and representativeness of the gravity model and given the characteristics of our data (e.g., we do not have input-output linkages at such disaggregated level), we present its theoretical foundations in a simple one-sector endowment-economy setting with CES preferences as, for example, summarize[d in Costinot and Rodríguez-C](#page-41-2)lare ([2014](#page-41-2)) and Yotov et al. (2016)¹¹

There areN countries with a xed stock Q_i of endowment with a unique variety [\(Arm](#page-40-2)[ington, 1969](#page-40-2)). Varieties are traded internationally. The value of total output is given by $Y_i = p_i Q_i$, with p_i denoting the product price in the exporting countryi.^{[12](#page-18-1)} Preferences are

with P_j denoting the price index, which is given by:

$$
P_{j}^{1} = \frac{X}{i} (i p_{i} t_{ij})^{1} : \qquad (3)
$$

Noting that X_{ij} given in Equation [\(2](#page-18-2)) also gives the trade ows from countryi to country j , and utilizing P_i^1 j^1 , bilateral trade ows between any two countries and j can be stated as follows:

$$
X_{ij} = P \frac{(i p_i t_{ij})^1}{(i p_i t_{ij})^1} E_j:
$$
 (4)

Dividing both sides of this equation by total spending from country , E_j , we obtain the share of total spending of imports from country in country j , $_{ij}$:

$$
{ij} = \frac{X{ij}}{E_j} = P \frac{(p_i t_{ij})^1}{(p_i t_{ij})^1}.
$$
 (5)

The share of spending is a function of prices and trade costs. From our partial estimates, we obtain estimates for the e ects of international borders, but not trade costs in levels. Assuming that the CES preference parameters's stay constant in the counterfactual analysis, [Deklp](#page-41-3)

equation [\(4](#page-19-0)) a[n](#page-19-1)d trade shares as given in equation (5) , can be written as:

$$
Y_{i} = \sum_{j}^{K} P\left(\frac{p_{i}t_{ij}}{(p_{i}t_{ij})^{1}}\right)^{T} - E_{j} = \sum_{j}^{K} E_{j}:
$$
 (7)

This expression forY_i holds in the baseline and the counterfactual, i.e.Y_i^b = P _{j ij} E_j and $Y_i^c =$ P $\int_{j}^{\text{c}} E_{j}^{\text{c}}$, respectively.

As we perform our quanti cation of the real expenditure e ects industry-by-industry, we observe trade imbalances in the data, i.e., total expenditures of a country in an industry will not equal total output of that industry. We take the observed trade imbalances, denoted by TI_i, as exogenous and constant between baseline and counterfactual, T.b, = $E_i - Y_i$.

To solve for the change ol interply we can use equation[s \(](#page-19-2)6) $Y_i^c =$ P ${}_{j}^{c} E_{j}^{c}$, $E_{i} = Y_{i} + T I_{i}$, $\mathbf{\nabla}_{i} = \mathbf{p}_{i}$, and $\mathbf{\dot{P}}_{i} = (\mathbf{Y}_{i}^{b} \mathbf{\dot{P}}_{i} + \mathbf{T} \mathbf{I}_{i}) = E_{i}^{b}$, to end up with:

$$
Y_i^{b}\boldsymbol{\varphi}_i = \begin{matrix} X & \frac{b}{ij} & \boldsymbol{\varphi}_i \boldsymbol{\varphi}_j \\ P & \frac{b}{i} & \boldsymbol{\varphi}_i \boldsymbol{\varphi}_j \end{matrix} \begin{matrix} \boldsymbol{\varphi}_i \\ \boldsymbol{\varphi}_i \end{matrix} \begin{matrix} \boldsymbol{\varphi}_i \\ \boldsymbol{\varphi}_i \end{matrix} + \boldsymbol{T} \boldsymbol{I}_j : \qquad (8)
$$

Hence, only data on trade shares in the baseline $_{ij}^{b}$ () and knowledge about, are needed to solve for \oint_i . Output is calculated from the observed trade ows, i.e., Y_i^b = P j X b ij . Trade imbalances,TI_j = E_j^b Y_j^b, are calculated using output in the baseline and expenditure in the baseline, calculated a $\mathbf{\bar{g}}_{j}^{\text{b}}$ = P i_{ij} X $_{ij}^{\text{b}}$. We set equal to 5 in line with the median value of 3:78 of the price elasticities(1) for structural gravity estimates reported in Table 3.5 in Head and Mayer (2014). Ideally, would be estimated using the econometric specication presented in 3.2, e.g., asi[n Fontagné et al. \(20](#page-41-4)22). This would require additional data on taris at the ITPD industry level. Such a database is currently under construction and, when completed, can be used to estimate industry-speci c elasticities.

The change in trade cost $\hat{\mathbf{s}}_i$ is de ned by our counterfactual experiment. Speci cally, we use the point estimates for international borders of the year 2019 for each industry. Since the international border coe cient for the rst year in our dataset for each industry is dropped,

it is convenient to take the border coecients of the last year in the dataset as measures of the globalization e ect for each industry for the respective period¹ We obtain border estimates for the year 2019 for 148 out of the 170 industr[ies](#page-21-1). As we want to quantify the e ects of globalization and use the latest year in our dataset, 2019, for the quanti cation, we perform an ex-post analysis. The observed values therefore are the baseline values in 2019, whereas the calculated counterfactual values are the values when globalization would not have taken place. The point estimates are therefore translated into changes of trade costs, $\mathfrak{k}_{\mathfrak{j}}$, in the following way: $\mathfrak{k}_{\mathfrak{j}}$ $=$ [1=exp($_{\textsf{T}}$)] $^{1=(1-)}$ for all i \mathfrak{s} j; T $=$ 2019 and $\mathfrak{k}_{\mathfrak{j}}$ $=$ 1 for $i = j$, in which T denotes the international border coe cient estimate in the nal period (2019) as per equation [\(1](#page-22-0)5) below. Note that the theory section abstracts from the sectoral dimension, which is present in the estimable equation and indicated with asuperscript on the time-varying border coe cients.

With solved values for changes ol, ϕ , the changes for expenditures \dot{R}_i), producer prices (p_j) , consumer prices (p_j) , trade shares b_{ij}), and nominal trade ows (\dot{x}_{ij}) can be calculated as follows:

$$
\dot{\mathbf{E}}_{j} = \frac{Y_{j}^{b}\dot{\mathbf{V}}_{j} + T\mathbf{I}_{j}}{\mathbf{E}_{j}^{b}};
$$
\n(9)

 $\mathbf{p}_j = \mathbf{\varphi}_j$; (10) $\mathbf{p}_j =$ X l $\frac{b}{b}$ p $\frac{b}{b}$ $\frac{1}{b}$ $! \frac{1}{11}$

Note that these changes give the values for the counterfactual when no globalization would haven taken place. For our variable of interest, real output change $\mathbf{\hat{s}}_i$, we calculated the e ect of globalization as follows:

$$
\hat{\mathbf{Y}}_j = \frac{\mathbf{p}_j}{\mathbf{p}_j} = \frac{1}{b_{jj}}^{-\frac{1}{1-\epsilon}};
$$
\n(14)

where we report the change from the solved, counterfactual values to the observed ones to get a quanti cation of globalization.^{[15](#page-22-1)} The last expression was derived [by Arkolakis et a](#page-40-0)l. ([2012](#page-40-0)), holding when $\mathbf{P}_{\text{ii}} = 1$ for all j, as is the case in our counterfactual scenari δ .

3.2 Econometric Speci cation

Based on Equation [\(2](#page-18-2)), and capitalizing on many developments from the empirical gravity literature, we specify the following econometric model, which will deliver the estimates of the e ects of globalization for each industryk from the ITPD-E database:

$$
X_{ij;t}^{k} = exp \n\begin{array}{ccc}\n hX & k\text{BRDR}_{ij;t} + {}_{1}RTA_{ij;t} + {}_{2}WTO_{ij;t} + {}_{3}EU_{ij;t} \\
h & \text{exp} & {}_{4}SANCT_COMPL_{ij;t} + {}_{5}SANCT_PARTL_{ij;t} \\
h & \text{exp} & \frac{k}{ijt} + \frac{k}{jjt} + \frac{k}{ij} & \frac{k}{ij;t} \\
\end{array} \n\tag{15}
$$

Here, $X_{ij;t}^k$ denote bilateral trade ows in levels in industryk from exporter i to importer j at time t. As discussed in the theory subsection, due to the separability property of the structural gravity model, equation [\(15](#page-22-0)) can be estimated at any desired level of aggregation

 15 As de ned in equation [\(14,](#page-22-3) changes in real output are positive when moving from the no-globalization scenario to the globalization scenario.

 16 As the equation system [\(8](#page-20-0)) is homogeneous of degree zero in prices, we chose producer prices in Canada as our numéraire.

 $(e.g., at the product, sector, industry, and/or aggregate level's).$ This is particularly important for us, as we will obtain estimates of the globalization eects for each of the ITPD-E industries in our sample. Consistent with gravity theory,X $_{\mathsf{i}\mathsf{j};\mathsf{t}}^{\mathsf{k}}$ includes domestic trade ows, cf. Yotov (2022). Domestic trade ows are important because they allow for trade diversion or import substitution with the domestic market, depending on the policy or trade shock being analyzed. Most important for our purposes, the fact that ITPD-E includes domestic trade ows will enable us to identify the e ects of globalization that we are after. Finally, following the recommendations of Egger et al. (2022 $\chi^{\,\rm k}_{\rm ij;t}$ includes data for all years in the sample.¹⁸

We will estimate Equation 15 for each industry with the Poisson Pseudo Maximum Likelihood (PPML) estimator, which, owing to Santos Silva and Tenreyro (2006, 2011), has two main advantages for gravity estimations. First, PPML addresses the problem that, due to heteroskedasticity, the OLS gravity estimates are inconsistent. Second, due to its the inclusion of these covariates and Bergstrand et al. (2015) demonstrate that the estimates of trade agreements in gravity regressions may be biased upward because they potentially capture common globalization trends. Relying on a comprehensive set of dummy variables to capture the eects of globalization has two advantages for our purposes. First, these covariates are exogenous by construction. Second, they would account for all possible globalization forces shaping trade, in addition to the policy covariates that will be included explicitly in our model. The large and signicant estimates that we will obtain reinforce our choice for the econometric treatment of globalization with time-varying border dummie $\frac{49}{2}$.

In addition to the time-varying globalization e ects, we also control for time-varying policy variables. Speci cally, we use indicator variables for the presence of regional trade agreements (RTAs) betweeni and j at time t, RTA_{lit} . The data on RTAs come from Egger and Larch (2008). We also control for whether the two trading partners are members of the World Trade Organization (WTO), WTO $_{ij;t}$, or of the European Union (EU),EU $_{ij;t}$. Data on memberships in the EU and the WTO come from the Dynamic Gravity Database of the United States International Trade Commission (US ITC), (Gurevich and Herman, 2018b). Finally, we control for the presence of complete and partial trade sanction $\text{SANCT}_\text{COMPL}_{\text{i}}$ and $SANCT$ PARTL_{iit}, respectively. Data on sanctions come from the latest edition of the Global Sanctions Database (Felbermayr et al., 2020; Syropoulos et al., 2023).

Equation (15) includes three sets of xed e ects. $k_{i;t}$ and $k_{j;t}$ are exporter-industry-time and importer-industry-time xed e ects. The theoretical motivation for including these xed e ects in gravity regressions is that they fully control for the unobservable multilateral resistance terms of Anderson and van Wincoop (2003) or, alternatively, for consumer and producer prices. In addition to controlling for the structural MRs, the exporter-industry-time and the importer-industry-time xed e ects will also absorb size variables (e.g., per capita

¹⁹Ideally, one would like to capture the impact of globalization by including only observable variables. To this end, we do include a set of policy variables that are conventionally used in gravity models. However, we still obtain very large additional e ects of globalization, which suggests the presence of many omitted factors for which data may not be available. From that perspective, our industry-time-varying estimates may be interpreted as all-inclusive measures of the e ects of globalization on trade.

income) and control for any other country-industry-speci c characteristics on the exporter and on the importer side that may a ect bilateral trade ows.

 $_{ij}^k$ denotes the set of directional country-pair-industry xed e ects. The motivation for $_{ij}^k$ is twofold. First, the country-pair-industry xed e ects will control for and absorb all possible time-invariant bilateral determinants of trade ows. This is potentially important in light of the ndings from Egger and Nigai (2015) and Agnosteva et al. (2019) who show that the standard gravity variables (e.g., distance, colonial relationships, etc.) are poor proxies for bilateral trade costs. Second, on a related note, as famously demonstrated by Baier and Bergstrand (2007), the use of country-pair xed eects mitigates potential endogeneity concerns in relation to bilateral trade policies by absorbing much of the unobserved/unmodeled correlation between the endogenous policy variables and the error term.

4 Results and Discussion

This section presents our disaggregated partial equilibrium estimates of the e ects of globalization on trade (in Subsection 4.1) and translates them into the impacts on real output using simulation (in Subsection 4.2).

4.1 Partial Equilibrium E ects of Globalization on Trade

Equation 15 delivers a sequence of globalization estimates for each of the ITPD-E industries. Due to di erent time-coverages for goods and services in the original data, we obtain globalization e ects over di erent periods for goods vs. services. Speci cally, for Agriculture, Mining and Energy, and Manufacturing, we use the period 1990-2019, while for Services, it is 2000-2019. We also note that, by construction, the estimates for the last year in our sample, i.e., 2019, capture the cumulated globalization e ects over the whole period of investigation. Therefore, for expositional simplicity, we report and discuss the estimates for 2019.

Due to the large number of industries in our sample, we visualize our results in Figure

24

1. The top panel of the gure reports all estimates, and the bottom panel removes the top and bottom 5% of outliers. Whereas conventional border e ects would ordinarily yield negative coe cient estimates as they re ect the border friction, the coe cients depicted in Figure 1 are positive since these border e ects in 2019 are ative to the initial (unidentied) year. This setup implies that, if the border became less important over time, this e ect of a relatively lower border friction then manifests as a positive coe cient. All estimates also appear in Table 4.[20](#page-26-0)

We draw two main conclusions based on the estimates in Figure 1 and Table 4. First, we note that, even after explicitly controlling for the impact of the WTO, RTAs, and EU membership, the e ects of globalization on international trade have been very large and signi cant. The average across all industries globalization e ect that we obtain is 1.9 , which implies a remarkable increase in global trade of about 570% over the period of investigation. By contrast, only ve industries have negative estimated globalization e ects and only one of them, `Cutting shaping and nishing of stone', is statistically signicant,

Figure 1: Industry-level Globalization Estimates

Notes: This gure plots the PPML gravity estimates and corresponding condence intervals for the eects of globalization on industry-level trade. The dependent variable is nominal trade in levels from ITPD-E. All estimates are obtained with exportertime xed e ects, importer-time xed e ects, pair xed e ects, and time-varying policy variables (e.g., WTO membership, EU membership, RTAs, and Sanctions). The globalization estimates are those for 2019, thus capturing the cumulated eects from the rst year of the sample for each industry. For Agriculture, Mining and Energy, and Manufacturing the omittted/reference year is 1991. For Services, it is 2000. Standard errors are clustered by country pair. The full set of estimates appears in Table 4. The top panel reports all estimates. The bottom panel removes the top and bottom 5% of outliers.

Figure 2: Globalization Estimates, Broad Sectors

Notes: This gure plots the statistically signi cant PPML gravity estimates and corresponding con dence intervals for the e ects of globalization on industry-level trade. The dependent variable is nominal trade8(is)-37sTFduee

The ve industries with negative estimated globalization e ects may seem surprising at rst, but we believe that these results should not be discarded as they may yield some insights. Specically, even though not statistically signicant, the largest negative estimate that we obtained is for the industry `Publishing of newspapers journals etc.' Our estimates suggest that, due to the globalization forces that are captured by our time-varying border dummy variables, international trade of newspapers and journals has decreased relative to domestic sales. We nd this result intuitive, and a natural explanation for it is a combination of relatively high transportation costs for such media on the one hand and the rapid advancements in online and social media on the other hand. The other three industries for which we obtain negative estimates are `Aircraft and spacecraft', `Construction', and `Electric motors generators and transformers'.

We also o er a preliminary investigation for possible heterogeneous e ects of globalization depending on country level of development. To this end, we use the 2000 classication of the World Bank to identify the `High Income' countries in our sample, and we obtain estimates of the eects of globalization for the subsample of rich countries only. Our estimates for 2019 are included in Table 5 and we visualize them, together with the average industry estimates, in Figure 3. The main conclusion that we draw based on these results is that there are no systematic dierences in the eects of globalization for the rich countries.

However, it may also be possible that, due to their disproportional size, our average results are driven by large and rich countries. Therefore, we plan to investigate the e ects of globalization across four groups, including exports from rich to rich countries, from rich to poor countries, from poor to rich countries, and from poor to poor countries. Moreover, we do see some intuitive variation when we compare the average globalization estimates with those for the rich countries across the four broad sectors in our sample. Specically, we nd the the globalization e ects are smaller for the rich countries in Agriculture, but larger for the rich countries in Services, while for manufacturing, the two estimates are almost identical with a slightly larger estimate for the rich countries. Finally, we noted that it is very likely

28

that even if the e ects of globalization are uniform across the countries in our sample within each sector, they can generate very heterogeneous welfare e ects across the countries.

4.2 On The Real Output E ects of Globalization

This section presents and discusses the real output eects that correspond to our partial estimates, and which we obtained from the model that we described in Section 3.1. While we use ITPD-E for our estimations, as it relies only on reported data, ITPD-E is highly unbalanced and thus not suitable for the quanti cation of the real output e ects. With ITPD-S, researchers now have a dataset that is consistent with ITPD-E in terms of countries, industries, and years, and which is balanced, such that it is suitable for quantitative, counterfactual analysis.

To obtain the impact of globalization on real output, we use the model presented in Section 3.1 to simulate a counterfactual scenario in which globalization had not occurred. Our base year is the average of the last three years in ITPD-S, 2017-2019. This averaging increases the number of non-zero observations by 17%. The change in real output due to globalization is calculated as the real output in the base year with globalization relative to the real output in the base year without globalization.

To simplify computation, we aggregate some countries into the rest of the world (ROW) in each industry. All EU countries, the United States (USA), China (CHN), Russia (RUS), Canada (CAN), as well as the largest 70 other exporters in an industry are modeled individually. The rest of the countries are aggregated into the ROW. Therefore, each industry has around 100 countries instead of the 265 in ITPD-S. These countries cover, on average, 99:93%of total trade (with a minimum across industries of98:72%and a maximum of100%).

We obtain real output e ects for each of the 148 industries for all countries available in this industry. As these are far too many numbers to report and digest, we provide two gures. The rst, Figure 4, reports output-weighted averages of real output changes over all countries within an industry. The x-axis ranks the industries according to the size of the real

Figure 3: Globalization Gravity Estimates, Rich vs. All

Notes: This gure plots the PPML gravity estimates and corresponding con dence intervals for the eects of globalization on industry-level trade. In addition, as dots, the gure plots the corresponding globalization estimates for the rich countries in our sample, as classied according to the 2000 World Bank income group classication. In each case, the dependent variable is nominal trade in levels from ITPD-E. All estimates are obtained with exporter-time xed eects, importer-time xed eects, pair xed eects, and time-varying policy variables (e.g., WTO membership, EU membership, RTAs, and Sanctions). The globalization estimates are those for 2019, thus capturing the cumulated eects from the rst year of the sample for each industry. For Agriculture, Mining and Energy, and Manufacturing the (omitted) reference year is 1991. For Services, it is 2000. Standard errors are clustered by country pair. The full set of estimates appears in Table 5. The top panel reports all estimates. The bottom panel keeps only the statistically signi cant estimates.

output e ects. For all gures, we cut all observations below the 5th percentile and above the 95th percentile for better readability. The gure reveals two important insights. First, for all industries besides one (`Mining of iron ores') we nd positive eects of globalization in terms of real output e ects. Second, there is substantial heterogeneity across sectors. The real output e ects range from 1:

with slight negative e ects). Also across countries, we nd substantial heterogeneity, ranging from :08% to 117:48%. Hence, it seems that the eects across countries vary more than across industries. The results for the simple averages are reported in the Appendix in Figure [11](#page-48-0). We also provide a comparison of the weighted and simple averages in Fi[gure](#page-48-1) 12, which highlights again that the results are qualitatively similar. However, across countries, also the magnitudes of the weighted and simple averages are quite similar.

4.3 Distributional Implications

So far, we have seen that there is substantial heterogeneity both across industries and across countries, with the latter even more pronounced. We, therefore, next investigate the relationship between the real output gains from globalization relative to the size of countries, measured by average output over the years 2017-2019 without the impact of globalization as a baseline, which is also used in the quanti cation.

Figure 6 plots real output e ects (using output as weights to aggregate) against the log of 2019 no-globalization output; as such, this is a type of `convergence graph' that reveals how gains are distributed across pre-globalisation country size. Notwithstanding considerable heterogeneity, a negative overall relationship is apparent. Specically, small economies which are almost always poorer developing countries such as the Central African

coe cient is strongly positive $(+0.73)$. In part, this relationship simply reects the nature of our counterfactual exercise, as the impact on real output of a counterfactual trade cost change is bound to be larger for countries that trade a larger share of their output. The second insight, though, which is perhaps less obvious, is that an economy can (will) bene t from trade openness no matter at which stage of development it is. Put dierently, the positive correlation isnot driven by any particular color group, which denotes a country's income per capita in 201 95 rather, countries from each income bracket are scattered around the tted line in roughly the same manner. This implies that the bene ts of more trade are in principle open to any economy whether it is poor or rich.

²⁵ Low income countries are green, lower-middle income is orange, upper-middle income is lavender, and high income countries are navy blue.

Figure 8: Distribution of Real Output Gains

5 Conclusions

In this paper, we introduce theInternational Trade and Production Database for Simulation (ITPD-S), which is a fully balanced database that covers 170 industries and 265 countries during the period 1990-2019. As such, the ITPD-S is the most disaggregated dataset that is currently available for performing counterfactual simulations for trade policy analysis. To highlight these possibilities, we combine the ITPD-S with the International Trade and Production Database for Estimation(ITPD-E), which is of the same dimensions and can be used for estimation, and we quantify the impact of globalization on trade and welfare in the world over the period 1990-2019. To perform the analysis, we rely on well-established methods and we complement the ITPD datasets with several additional standard databases.

We start by obtaining partial equilibrium estimates of the e ects of globalization at the industry level. To this end, we capitalize on the fact that the ITPD-E includes domestic trade ows. Several ndings stand out. Most importantly, we obtain large, positive and statistically signicant estimates of the eects of globalization on trade, which imply that, on average, bilateral globalization forces (other than trade agreements, WTO membership, and EU membership, which we control for in our analysis) have led to a remarkable increase of 570% in international trade relative to domestic sales over the period 1990-2019. In addition, we nd that the globalization estimates that we obtain are very heterogeneous across the ITPD sectors, with larger e ects for `Services' and smaller e ects for `Agriculture'. Finally, even though we do not observe signi cant di erences between the e ects of globalization for the rich countries in our sample, we do see some intuitive variation across the broad sectors with stronger globalization bene ts for the rich countries in `Services' but smaller than the average e ects in `Agriculture'.

Our analysis reveals that the gains from globalization in terms of real industry output have been signi cant for most countries. At the same time, we also document substantial heterogeneity in these e ects, which appear to be more pronounced across countries than across industries. Speci cally, developing and smaller countries seem to have proted the

37

most from increasing international trade. This nding is then re ected in decreasing global inequality measured by comparing Gini coe cients from current, observed output levels and output levels predicted if globalization had not taken place.

While these results are encouraging, they leave as yet unanswered deeper questions about the driving forces behind those globalization e ects that manifest even after explicitly accounting for a variety of factors including regional integration agreements. For instance, globalization, as we dene it, could be driven bypolicy interventions such as unilateral tari and NTM reductions, or by secular trends

References

- Delina E. Agnosteva, James E. Anderson, and Yoto V. Yotov. Intra-national trade costs: Assaying regional frictions. European Economic Review 112:32 50, 2019. doi: 10.1016/j. euroecorev.2018.11.008.
- James E. Anderson. A theoretical foundation for the gravity equation.American Economic Review, 69(1):106116, 1979.
- James E. Anderson and Eric van Wincoop. Gravity with gravitas: A solution to the border puzzle. American Economic Review 93(1): 170 192, 2003.
- James E. Anderson and Eric van Wincoop. Trade costs burnal of Economic Literature, 42 (3):691751, 2004.
- James E. Anderson and Yoto V. Yotov. Short run gravity. Journal of International Economics, 126:103341, 2020. doi: 10.1016/j.jinteco.2020.103341.
- James E. Anderson, Ingo Borchert, Aaditya Mattoo, and Yoto V. Yotov. Dark costs, missing data: Shedding some light on services trade European Economic Review105:193 214, 2018.
- James E. Anderson, Mario Larch, and Yoto V. Yotov. Transitional growth and trade with frictions: A structural estimation framework. The Economic Journal, 130(630):1583 1607, 2020. doi: 10.1093/ej/ueaa020.
- Costas Arkolakis, Arnaud Costinot, and Andrés Rodriguez-Clare. New trade models, same old gains? American Economic Review, 102(1):94130, 2012.
- Paul S. Armington. A theory of demand for products distinguished by place of production. IMF Sta Papers, 16:159176, 1969.
- Scott L. Baier and Je rey H. Bergstrand. Do free o8rope8Delina E. Agnostevo8rop Emem-tructural56
- Arnaud Costinot and Andrés Rodríguez-Clare. 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. Rogo, Elsevier Ltd., Oxford, pages 197261, 2014.
- Arnaud Costinot, Dave Donaldson, and Ivana Komunjer. What goods do countries trade? a quantitative exploration of ricardo's ideas.Review of Economic Studies79(2):581 608, 2012.
- Robert Dekle, Jonathan Eaton, and Samuel Kortum. Unbalanced trade@merican Economic Review: Papers and Proceeding 97(2): 351 355, 2007.
- Robert Dekle, Jonathan Eaton, and Samuel Kortum. Global Rebalancing with Gravity: Measuring the Burden of Adjustment. IMF Sta Papers, 55(3):511 540, June 2008. ISSN 1020-7635. doi: 10.1057/imfsp.2008.17. URtttp://www.palgrave-journals. com/doifinder/10.1057/imfsp.2008.17 .
- Jonathan Eaton and Samuel Kortum. Technology, geography and trade conometrica 70 (5):1741 1779, 2002.
- Peter Egger and Mario Larch. Interdependent preferential trade agreement memberships: An empirical analysis. Journal of International Economics, 76(2):384399, 2008.
- Peter H. Egger and Sergey Nigai. Structural gravity with dummies only: Constrained ANOVA-type estimation of gravity models. Journal of International Economics, 97:8699, 2015.
- Peter H. Egger and Filip Tarlea. Multi-way clustering estimation of standard errors in gravity models. Economics Letters 134:144 147, 2015.
- Peter H. Egger, Mario Larch, and Yoto V. Yotov. Gravity-model estimation with timeinterval data: Revisiting the impact of free trade agreements Economica 89:44 61, 2022.
- Gabriel Felbermayr, Aleksandra Kirilakha, Constantinos Syropoulos, Erdal Yalcin, and Yoto V. Yotov. The global sanctions data base. European Economic Review 129(C), 2020.
- Lionel Fontagné, Houssein Guimbard, and Gianluca Orece. Tari-based product-level trade elasticities. Journal of International Economics, 137(C), 2022.
- Tamara Gurevich and Peter Herman. The dynamic gravity dataset: 1948-2016USITC Working Paper 2018-02-A, 2018a.
- Tamara Gurevich and Peter Herman. The dynamic gravity dataset: 1948-2016USITC Working Paper 2018-02-A, 2018b.
- Keith Head and Thierry Mayer. Gravity equations: Workhorse, toolkit, and cookbook. Chapter 3 in the Handbook of International Economics Vol. 4, eds. Gita Gopinath, Elhanan Helpman, and Kenneth S. Rogo, Elsevier Ltd., Oxford, pages 131 195, 2014.

Table 4: Industry-Level Globalization Estimates

Notes: This table reports PPML gravity estimates of the e ects of globalization on industry-level trade.
The dependent variable is nominal trade in levels from ITPD-E. All estimates are obtained with exporter-
time, import

Table 5: Industry-Level Globalization Estimates, Rich Countries

Notes: This table reports PPML gravity estimates of the elects of globalization on industry-level trade. The dependent variable is nominal
trade in levels from ITPD-E. All estimates are obtained with exporter-time, importe

Figure 9: The E ects of Globalization - Industry Results (simple average)

Figure 10: The Eects of Globalization - Industry Results; Comparison of Simple and Weighted Average

Figure 11: The E ects of Globalization - Country Results (simple average)

Figure 12: The Eects of Globalization - Country Results; Comparison of Simple and Weighted Average

Figure 13: The E ects of Globalization - The Role of Country Size

Figure 14: The E ects of Globalization - The Role of Country Size (simple average)

Figure 15: The E ects of Globalization - The Role of Country Size (simple average, without LCA)