CGPA WORKING PAPER
2025-01

Estimating Gravity Equations: Theory Implications, Econometric Developments, and Practical Recommendations

Mario Larch
Serge Shikher
Yoto V. Yotov

CENTER FOR GLOBAL POLICY ANALYSIS
LeBow College of Business
Drexel University

Estimating Gravity Equations: Theory Implications, Econometric Developments, and Practical Recommendations*

Mario Larch† Serge Shikher‡ Yoto V. Yotov§

January 3, 2025

Abstract

We trace the developments in the empirical trade literature to make fifteen recommendations for estimating gravity equations, which are structured in three categories: data, estimating equation, and heterogeneity. We also offer practical tips and identify areas where further research is needed. Based on these recommendations, we specify a comprehensive estimating model, which can serve as a benchmark for gravity estimations even when it is not possible to implement all of our recommendations. The proposed methods should be useful for gravity estimations beyond international trade, e.g., migration, foreign investment, cross-border patenting, and other flows.

JEL codes: F10, F14, F63.
Keywords: Gravity equation, Estimation, Recommendations, Practical tips.

*We thank Jim Anderson, Scott Baier, Jeff Bergstrand, Peter Egger, Tom Zylkin, and two anonymous referees for many helpful comments and suggestions. 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 any of its Commissioners, nor the opinions of any of the other organizations that the authors are affiliated with. We dedicate this paper to Professor Jeffrey Bergstrand. Each of us has been deeply impacted by Jeff and his work, and we are very grateful for his tremendous contributions to the trade gravity literature and beyond.

†Department of Law, Business Administration & Economics, University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany; Centre d'Etudes Prospectives et d'Informations Internationales (CEPII), Paris, France; ifo Institute and CESifo Research Network, Munich, Germany; Austrian Institute of Economic Research (WIFO), Vienna, Austria; Nottingham Centre for Research on Globalisation and Economic Policy (GEP), Nottingham, United Kingdom. E-mail: mario.larch@uni-bayreuth.de.

‡Corresponding author. United States International Trade Commission. Email: serge.shikher@usitc.gov.

§Drexel University, ifo Institute, CESifo Research Network, Bulgarian Council for Economic Analysis. E-mail: yvy23@drexel.edu.

1 Introduction

The gravity equation of trade is the most successful empirical model in (international) economics, and hundreds of thousands of academic papers and policy reports have used the gravity model to estimate the effects of various determinants of trade flows.[1] Proper estimation of the gravity equation will lead to consistent and unbiased estimates of the trade elasticities for various policies. Moreover, a correctly specified gravity equation will deliver a good fit of observed trade flows and, therefore, properly predict trade costs and flows. This is important if one utilizes estimated trade costs to construct trade flows for counterfactual analysis (Anderson et al., 2018b).

Against this backdrop, the motivation for this paper is threefold. First, a decade has passed already since the last systematic analyses and recommendations for estimating trade gravity equations (Head and Mayer, 2014; Yotov et al., 2016).[2] In the meantime, the trade profession has made and witnessed many new and important developments, which have not been documented in the existing gravity surveys and monographs. Second, despite having some well-documented recommendations for estimating gravity models and the great progress that has been made to do theory-consistent applied gravity work, many policy reports and academic papers, including recent papers in top journals, do not follow, or even acknowledge, some of the most widely accepted and well-established practices for estimating gravity models (e.g., controlling for the multilateral resistances or estimating gravity in multiplicative form). Third, there are many recent gravity applications, including theoretical foundations, which go beyond trade.[3] Based on some of our work and discussions with colleagues from different fields, we believe that (apart from the other data) most, if not all, of the developments in the estimating trade gravity literature, are directly applicable to gravity settings beyond international trade, e.g., commuting, migration, FDI, financial-asset flows, cross-border patents, mergers, and acquisitions, etc., as summarized in Footnote 3.

As a result, this paper aims to review the related literature to identify the best current practices and practical tips for estimating gravity equations and synthesize these practices and tips into recom-

[1] A *Google Scholar* search on "trade gravity" on November 28, 2024, delivered 1,710,000 results.

[2] See Baldwin and Taglioni (2006) and De Benedictis and Taglioni (2011) for earlier surveys.

[3] See, for example, Persyn and Torfs (2015) and Dingel and Tintelnot (2023) for gravity for commuting, Poot et al. (2016) and Beverelli (2021) for gravity for migration, Eaton and Tamura (1994), Benassy-Quere et al. (2005), Bergstrand and Egger (2007), Paniagua (2015), Kox and Rojas-Romagosa (2020) and Larch and Yotov (2024a) for gravity for foreign direct investment (FDI), Mercado (2022) for gravity for financial-asset flows, Carril-Caccia et al. (2023) for gravity for greenfield investment, LaBelle et al. (2023) for gravity for cross-border patents, and Bradley et al. (2023) for gravity for mergers and acquisitions.

mendations for estimating gravity models. To fulfill these objectives, we proceed in three steps. First, in Section 2, we introduce a representative theoretical gravity model for two purposes. Consistent with the seminal work of Arkolakis et al. (2012), it demonstrates that many different theoretical microeconomic foundations converge to the same gravity representation. Second, guided by the economic gravity theories and capitalizing on many developments in the empirical gravity literature and the econometrics literature more broadly, in Section 3, we offer fifteen recommendations for estimating trade gravity models. In addition, we also identify opportunities for new work needed to address certain challenges in the existing literature.

To facilitate the exposition and structure of the analysis, we group our recommendations into three broad categories – Data, Estimation, and Heterogeneity. Moreover, we have attempted to follow the same structure for the presentation of each recommendation. Specifically, after stating the recommendation, we first discuss the motivation (including related challenges with gravity estimations) and the advantages of implementing this recommendation. Then, we discuss potential caveats, limitations, and alternative approaches for instances when it is not possible to follow our recommendation. Finally, to the extent possible, we have ordered the recommendations within each of the three broad categories in terms of importance.

We make the following five recommendations within the 'Data' category: (i) Use data on bilateral trade flows for all possible countries; (ii) Use administrative data, on nominal trade flows in common currency, at delivered prices; (iii) Use disaggregated data; (iv) Use panel data for consecutive years; and (v) Include domestic trade data. We also make six additional recommendations on the 'Estimation' of the gravity model: (vi) Estimate gravity in its multiplicative form using PPML; (vii) Use exporter-time and importer-time fixed effects; (viii) Employ asymmetric country pair fixed effects; (ix) Carefully model bilateral trade costs; (x) Allow non-discriminatory trade costs; and (xi) Cluster standard errors. Finally, we make four recommendations regarding the 'Heterogeneous' trade costs and policy effects: (xii) Obtain disaggregated policy estimates; (xiii) Allow for dynamic adjustments in the trade costs; (xiv) Consider other types and sources of heterogeneity; and (xv) Consider using heterogeneity-robust DiD methods.

In the third and final step, in Section 4, we synthesize our recommendations to specify a comprehensive estimating gravity model. While we believe that the gravity model that we propose encompasses the leading methods for gravity estimations and reflects the current state of the literature, we are aware that it may not be possible to implement and apply all of our recommendations at all times, e.g., due to data limitations. However, we also believe that it is useful to have a benchmark estimating model and recognize the benefits of applying our recommendations and staying close to the proposed specification as well as to be aware of the potential caveats of not being able to do so. Finally, as noted earlier, we think that our recommendations can benefit and be applied to gravity estimations beyond trade, e.g., migration, FDI, cross-border technology transfers, commuting, financial-asset flows, greenfield investment, M&As, etc.

Before we continue, we note that this paper is not intended to be an exhaustive survey of the empirical gravity literature. As can be seen from our bibliography, we cite a large body of related literature. However, we may have missed some relevant contributions. Our approach to reviewing the literature included four main criteria: (i) To cover the most famous and impactful gravity papers. The motivation for this choice is obvious – these papers have shaped the gravity literature; (ii) To acknowledge the most recent contributions, which we believe are relevant to our recommendations. The motivation for this choice is to bring the reader to the frontier of the literature; (iii) To review the papers, that we believe are most closely related to our recommendations and objectives; and (iv) Finally, even though we are aware that self-citations may be viewed as self-serving, we cite our papers whenever they are relevant to a particular recommendation and, especially, when they were written specifically to address challenges with the estimation of gravity models.

2 Theoretical foundations and implications for gravity estimations

By analogy with physics, most bilateral relationships (X_{ij}) in economics, e.g., trade, migration, and foreign direct investment (FDI), can be represented as combinations of (i) origin-specific characteristics (O_i), e.g., value of output, (ii) destination-specific characteristics (D_i), e.g., expenditure, and (iii) bilateral characteristics (T_{ij}), e.g., bilateral distance, subject to the following multiplicative equation:[4]

$$X_{ij} = O_i \times D_j \times T_{ij}. \tag{1}$$

[4]The first bilateral relationship of this form was estimated for migration flows by Ravenstein (1885, 1889), while Tinbergen (1962) is credited for the first gravity application to trade. The gravity relationship holds for many other flows, such as capital, knowledge, and foreign direct investments. The focus of our discussion is on trade flows, however, most of the empirical recommendations we will make in the next section apply more broadly to other bilateral flows.

While the generic equation (1) can represent any bilateral economic relationship, the most influential developments in the economic theory of gravity come from the field of international trade. Following Anderson (1979), who derived the first theoretical gravity model of trade, other early theoretical foundations for the trade gravity equation were provided by Bergstrand (1985), and Bergstrand (1989). Since then, many authors have built gravity models that not only provide structural foundations for the three terms in equation (1), e.g., Anderson and van Wincoop (2003) and Eaton and Kortum (2002), but also extend the gravity model to the sectoral level, e.g., Anderson and van Wincoop (2004); Costinot et al. (2012a); Shikher (2012a,b), and to derive dynamic gravity models, e.g., Olivero and Yotov (2012); Eaton et al. (2016); Anderson et al. (2020). Arkolakis et al. (2012) demonstrate that the structural gravity equation is representative of many alternative theoretical foundations, while Allen et al. (2020) prove the uniqueness and existence of the 'universal' gravity model. The following structural gravity equation synthesizes some of the major developments in the theoretical gravity literature:[5]

$$X_{ij,t}^k = \underbrace{(Y_{i,t}^k/Y_t^k)(\Pi_{i,t}^k)^{\theta^k}}_{O_{i,t}^k} \times \underbrace{E_{j,t}^k(P_{j,t}^k)^{\theta^k}}_{D_{j,t}^k} \times \underbrace{({t_{ij,t}}^k)^{-\theta^k}}_{T_{ij,t}^k}. \tag{2}$$

In what follows, we will define each of the terms in equation (2), while using some of the underlying gravity theories to motivate many of our recommendations for gravity estimations. $X_{ij,t}^k$ denotes exports in sector k from origin/exporter i to destination/importer j at time t at delivered prices, i.e., inclusive of trade costs. The empirical implication of the latter is that the theory-consistent dependent variable in the estimating gravity model should include trade costs (e.g., shipping costs). The sectoral dimension of $X_{ij,t}^k$ reflects the so-called 'separability' property of the theoretical gravity model. The empirical implication of this theoretical property is that gravity can be estimated at any level of aggregation. The time dimension of $X_{ij,t}^k$ is motivated by the fact that the value of output, expenditure, and some bilateral trade costs, e.g., sanctions and tariffs, vary over time.[6] This has implications for estimations with panel data. Finally, gravity theory implies that $X_{ij,t}^k$ should include not only international but also domestic trade flows, and, as we argue in the next section, the use of

[5]Our objective is not to provide a thorough review of the theoretical foundations of the gravity models. Instead, we focus on the gravity theories that will motivate our estimation recommendations. We refer the reader to Yotov et al. (2016) for alternative derivations of equation (2) with the same notation as in the current paper, which follows Anderson and van Wincoop (2003), and to Costinot and Rodriguez-Clare (2014) for gravity presentations with alternative notation, which follow Eaton and Kortum (2002).

[6]From a theoretical perspective, the time dimension of $X_{ij,t}^k$ is motivated by gravity models with country-specific dynamics, e.g., Olivero and Yotov (2012), Eaton et al. (2016), Anderson et al. (2020), or bilateral dynamics, e.g., Arkolakis (2010) and Anderson and Yotov (2022).

domestic trade flows has many important implications for gravity estimations.

The structural elements on the right-hand side of equation (2) can be grouped into three broad terms, which correspond to the three terms in (1). The first broad term ($O^k_{i,t}$) varies at the exporter-sector-time dimension and includes three structural elements. $Y^k_{i,t}/Y^k_t$ is the value of gross output in sector k in exporter i at time t ($Y^k_{i,t}$) as a fraction of the value of the total world output in sector k at time t (Y^k_t). $\Pi^k_{i,t}$ is coined by Anderson and van Wincoop (2003) as the *Outward Multilateral Resistance* (OMR), and the intuition behind this term is that it captures the fact that the further the producers in sector k in exporter i at time t are from the rest of the world, the more they will ship to a particular destination. Alternatively, the exporter-sector-time dimension can be represented by producers prices.[7] Finally, $\theta^k > 0$ is the trade elasticity, which measures the response of bilateral trade flows with respect to changes in bilateral trade costs. The sectoral dimension of the trade elasticity is motivated by disaggregated gravity theories, e.g., Anderson and van Wincoop (2004); Costinot et al. (2012a); Shikher (2012a,b). Individually, or in combination, the structural terms in $O^k_{i,t}$ motivate using exporter-sector-time fixed effects when gravity is estimated with pooled sectoral panel data. When the gravity model is derived at the aggregate level, the corresponding fixed effects should vary at the exporter-time dimension. Finally, exporter fixed effects should be used if gravity is derived in the cross-section and without sectors.

The second broad term ($D^k_{j,t}$) in equation (2) varies at the importer-sector-time dimension and includes two new structural elements. $E^k_{j,t}$ denotes expenditure in sector k in importer i at time t, and $P^k_{j,t}$ is the *Inward Multilateral Resistance* (IMR) (Anderson and van Wincoop, 2003). The intuition behind the IMR is that it captures the fact that the further the consumers of goods from sector k in importer i at time t are from the rest of the world, the more they will buy from a particular exporter. Another intuitive (and theory-consistent) interpretation of $P^k_{j,t}$ is as an ideal consumer price index. The structural terms in $D^k_{j,t}$ motivate the use of importer-sector-time fixed effects when the gravity model is estimated with pooled sectoral panel data. When the gravity model is derived at the aggregate level, the corresponding fixed effects should vary at the importer-time dimension. Finally, if gravity is derived in the cross-section and without sectors, importer fixed effects should be used.

[7] The OMR is an intuitive theoretical construct without a direct equivalent in the data. However, through the market clearing condition, the OMRs from equation (2) can be mapped one-to-one to the factory gate prices for the producers in the gravity model. Specifically, $\left(\beta^k_i p_{i,t}\right)^{-\theta^k} = (\Pi^k_{i,t})^{\theta^k}(Y^k_{i,t}/Y^k_t)$, where β^k_i is a country- and sector-specific (preference or technology) parameter. Thus, the OMRs can be interpreted in terms of the producer prices. In this case, the first term in equation (2) becomes $O^k_{i,t} = (\beta^k_i p_{i,t})^{-\theta^k}$.

The last term in equation (2) includes the vector of bilateral trade costs ($t^k_{ij,t}$). The assumption of a positive trade elasticity ($\theta^k > 0$) implies an intuitive inverse relationship between the bilateral trade flows and the bilateral trade costs. Disaggregated gravity theories, and the fact that many trade policies are implemented at the sectoral level, motivate the sectoral dimension of $t^k_{ij,t}$. The time dimension of $t^k_{ij,t}$ is motivated by the fact that some bilateral trade policies are time-varying.[8] Gravity theories with heterogeneous firms, e.g., Chaney (2008) and Egger et al. (2021), imply that the vector of bilateral trade costs consists of variable and fixed costs. Tariffs should also feature in the vector of bilateral trade costs, and an important implication of gravity theories with tariffs, e.g., Anderson and van Wincoop (2001), is that the estimate of the coefficient on tariffs can be used to recover the trade elasticity parameter. Finally, the bilateral trade costs can be time-varying and/or time-invariant. The latter motivates the use of country-pair fixed effects in gravity estimations.

We conclude this section with two potential caveats. First, while many theoretical foundations lead to a gravity equation as stated in equation (2), some frameworks do not. For example, the trans-log gravity framework from Novy (2013b), the non-homothetic gravity framework of Fajgelbaum and Khandelwal (2016), the "non-Pareto"-framework of Head et al. (2014) and Melitz and Redding (2015), and frameworks with persistence resulting from sunk market entry as in Egger et al. (2023) or sticky prices as in Larch et al. (2024b). Second, while we believe that our recommendations for estimating the gravity model of trade apply in most bilateral settings (even beyond trade), we emphasize that the exact estimating equation should be guided by the research question and the specific underlying theoretical foundation.

3 Recommendations for gravity estimations

Guided by theory, and capitalizing on many developments in the empirical gravity literature and the econometrics literature more broadly, we make fifteen recommendations for estimating trade gravity models in this section.[9] To facilitate the exposition and add structure to the analysis, we group our recommendations into three broad categories – (i) Data, (ii) Estimation, and (iii) Heterogeneity, which correspond to the three subsections of this section. Moreover, we have attempted to follow the same

[8] An additional motivation can be bilateral dynamics, e.g., Arkolakis (2010) and Anderson and Yotov (2022).

[9] While we recognize that there may be overlap between some of the recommendations (and we explicitly acknowledge such overlaps whenever appropriate), after numerous discussions, we are convinced that each of the fifteen recommendations that we make is sufficiently different and important to stand as a separate item on our list.

structure for each recommendation. Specifically, after stating the recommendation, we first discuss the motivation (including related challenges with gravity estimations) and the advantages of implementing this recommendation. Then, we discuss potential caveats, limitations, and alternative approaches for instances when it is not possible to follow our recommendation.

3.1 Data on the dependent variable in gravity estimations

Most of the recommendations in this section are guided by theory, though some of them are motivated by practical considerations. Our focus is exclusively on the dependent gravity variable, but in the next section we offer some discussion on the gravity covariates and the corresponding data.

Recommendation 1: Use data on bilateral trade flows for all possible countries.

By definition, the gravity model is bilateral. Therefore, to perform gravity estimations, one needs at minimum trade data for multiple pairs of countries. Over time, the datasets on bilateral trade flows have improved significantly, both in terms of reliability and in terms of coverage across countries, products, and time. Thus, in most cases, it should be possible to access reliable data for many countries, and our recommendation is to estimate gravity with trade data for all possible countries. In addition to enhancing the possibility of obtaining certain country-specific gravity estimates, the inclusion of more countries will also improve estimation efficiency, while not considering all possible countries and bilateral flows may lead to a sample selection bias and inconsistent and biased gravity estimates. Sometimes, however, two practical challenges arise.

First, we have often been asked, especially by government agencies that have access to unique administrative trade data for their respective countries,[10] whether it is possible and valid to estimate gravity for a single country. The short answer to this question is 'Yes'. However, one important challenge in such settings is to properly account for the multilateral resistances, which can no longer be controlled for with country-specific fixed effects (as we will argue in the next subsection). An easy solution is to construct GDP-weighed 'remoteness' indexes based on distance (e.g., Wei, 1996; Davis and Weinstein, 2002) or based on bilateral trade costs that are defined more broadly (e.g., MacDonald et al., 2020). Proxying for the MRs with GDP-weighted remoteness indexes is better than not controlling for the MRs at all, though Felbermayr and Yotov (2021) demonstrate that the

[10]For example, this may be the case for country-specific analyses that are done at the tariff-line level. In the United States, the tariff-line level is the 10-digit level of the U.S. Harmonized Tariff System (HTS).

a-theoretical remoteness indexes do not perform well in predicting bilateral trade flows. Therefore, our recommendation is to attempt to construct structural MRs that are consistent with the theoretical gravity model, and we refer the reader to Baier and Bergstrand (2009, 2010) who provide a first-order log-linear Taylor series expansion of the MRs that are GDP-share-weighted (geometric) averages of the gross trade costs and Lampe et al. (2023), who propose and implement a procedure to control for the structural MRs in a single-country gravity setting.[11]

Another practical concern that has been brought to our attention quite often is related to the number of countries to be included in the gravity model. Convergence issues have hampered some gravity estimations with large data and many fixed effects in the past (e.g., Glick and Rose, 2016). However, this should no longer be an issue due to the emergence of fast estimation commands that can handle many fixed effects in linear (Correia, 2016) and non-linear (Larch et al., 2019; Correia et al., 2020) gravity models. Thus, our recommendation is to estimate gravity by including all possible countries, even though for obtaining consistent elasticity estimates this is not necessary when the proper set of fixed effects are employed.[12] On a related note, some authors have estimated gravity by splitting the sample according to certain country-specific criteria, e.g., developed vs. developing countries or EU vs. non-EU countries. Pfaffermayr and Yotov (2024, Drexel University) argue that such practices are not consistent with gravity theory and may lead to biased gravity estimates, especially if the selection of observed countries is not purely random. Therefore, following Pfaffermayr and Yotov (2024, Drexel University), our recommendation is to estimate the desired effects across country groups by introducing interactions and using the full estimating sample instead of splitting it.

If it is really necessary to limit the number of countries, e.g., for presentation or computational reasons (which, due to recent advances, hopefully, is seldom the case), then our recommendation is to select the countries to be dropped from the estimating sample very carefully. Naturally, the selection of countries should be guided by the question of interest. In general, we would not recommend dropping countries that are large trade partners for some of the countries that remain in the estimating sample. Similarly, we would not drop countries that have trade agreements or any other time-varying economic

[11] Aytun et al. (2024) propose methods to control for the multilateral resistances in firm-level gravity regressions. Most empirical firm-level gravity analyses are performed at the country level. However, due to the availability of better (more reliable and with wider coverage) firm-level data, we expect that multi-country gravity regressions with firm-level data will become more prevalent.

[12] Note that, as argued in Recommendations 7 and 8 of the next subsection, the use of exporter-time, importer-time, and country-pair fixed effects in the gravity model will greatly limit the data collection efforts for gravity estimations. Thus, in most cases, the only data that would be needed to include a country in the gravity estimation, in addition to trade data, would be data on the bilateral time-varying/policy variables for this country (e.g., its RTAs, sanctions, etc.).

relationships or political ties with the countries in the main estimating sample. Moreover, instead of dropping some countries from the estimating sample, we would consider aggregating them into a rest-of-the-world (ROW) region or experimenting with both alternatives. Including a ROW region in the gravity analysis may pose challenges when modeling ROW's bilateral trade costs. However, (i) as we argue in Recommendation 8 of the next subsection, the use of pair fixed effects may mitigate such concerns, and (ii) if some of the countries in the ROW region have agreements, sanctions, etc. with (many of) the countries in the main estimating sample, then they should probably not have been aggregated in the ROW region to begin with.

Gravity theory extends to sub-national trade flows, e.g., to trade between the U.S. states, and we are aware of studies that have estimated gravity for sub-national trade flows (e.g., McCallum (1995), Anderson and van Wincoop (2003), Anderson and Yotov (2010), Adam et al. (2024a), and Adam and Larch (2024)). In principle, our recommendations for gravity estimations also apply to sub-national trade flows, however, two related practical challenges arise. First, consistent with our recommendation to use bilateral data for all possible countries, to properly nest sub-national trade within the gravity model, one needs data not only for the trade flows between the sub-national units (e.g., trade between Ohio and Nebraska) but also data on the international trade flows between each sub-national unit and each foreign country too (e.g., trade between Ohio and Bulgaria). Second, combining sub-national and international trade flows is subject to the assumption that sub-national trade is driven by the same data-generating process as international trade. Thus, special care should be devoted to modeling sub-national trade costs, which are probably lower than the international trade costs but may still be substantial. Some of the recommendations that we make below, e.g., the use of bilateral fixed effects, may help with the modeling of sub-national trade costs. For example, one may allow border effects to vary according to whether trade flows are crossing international borders or only regional borders as in Adam et al. (2024a). However, we are not aware of datasets that include consistently constructed proxies and policies variables for sub-national trade costs. Thus, while we see significant benefits and potential from integrating sub-national trade within the international gravity model and we believe that our recommendations apply to sub-national trade flows, the focus of the remainder of the paper will be on country-to-country trade.

Recommendation 2: Use administrative data on nominal trade flows measured in common currency at delivered prices.

This recommendation includes three elements. First, the trade data used for the estimation should not have been constructed using statistical inference. Despite being seemingly obvious, we explicitly emphasize the use of raw, administrative data, which has not been subjected to statistical manipulation, for gravity estimations because we are aware of papers that rely on data that has been constructed using economics and statistical methods (e.g., GTAP and WIOD), including gravity models, and are, therefore, not appropriate for gravity estimations.

Second, gravity models are usually estimated with trade data measured in nominal terms in a single currency (usually, but not necessarily the U.S. dollar). This is consistent with the theoretical gravity model and it is also convenient from a practical point of view since most trade data that are consistent for many countries, products, and years are available in nominal U.S. dollars. This said, and also consistent with theory, it is possible to express bilateral trade in real terms and to explicitly introduce exchange rates in the theoretical gravity model. If this is done, one should be careful about the changes in the definition of the right-hand-side country-specific variables in the theoretical gravity model and the interpretation of their estimates. From a practical perspective however, transforming the dependent variable in real terms or introducing exchange rates explicitly should make no difference to our estimating recommendations because, as we argue in the estimation subsection, gravity should be estimated with exporter-time and importer-time fixed effects, which will absorb and fully control for the impact of consumer prices, producer prices, and exchange rates.[13] Moreover, consistently constructed price data may not be readily available. Therefore, for convenience and as long as there is no explicit need to proceed otherwise, we recommend following the standard practice of estimating gravity with nominal trade data in U.S. dollars. Finally, we note that rescaling the nominal trade values to larger units, e.g., from thousands to millions of U.S. dollars, may help with convergence, however, one should remember not to round the rescaled (nor the original) trade data.

Third, the theory implies that the dependent gravity variable should be trade flows at delivered prices, i.e., including bilateral trade costs for delivery from the source to the destination. The implication is that the theory-consistent data for gravity estimations should be on imports, including customs, insurance, and freight (c.i.f.). In the case of goods trade, another reason to prefer the

[13]In Section 3.2, we discuss how to capture possible real exchange rate effects on trade in the presence of the full set of fixed effects. It is also possible, and consistent with theory, to estimate gravity in real terms, i.e., using quantities instead of nominal values. The challenges with such estimations include data availability and the inability to aggregate quantities for different products.

importer-reported values is that importers have a greater interest in recording trade transactions completely and accurately since tariffs are collected based on these records or to ensure compliance with non-tariff measures such as SPS, TBT, etc. Services trade, on the other hand, is not subject to such measures. Therefore, the question turns into one about the most reliable or accurate firm-level data collection instruments, and many countries are more interested in the export side of services, e.g. as part of export promotion initiatives. As a result, the values for services transactions are usually more accurately recorded by exporters. Therefore, we recommend using services trade flows data on exports.

It is often the case that international trade data is reported by two countries: the exporter and the importer, e.g., Germany reports exports to Bulgaria and, for the same product and year, Bulgaria reports 'mirror' imports from Germany. Consistent with the previous discussion, we recommend using goods imports data and services exports data for gravity estimations.[14] However, we also think that the additional information from the mirror reporting is useful and should be taken into account in one of two ways. First, it may be possible to gain a significant number of observations by replacing some of the missing trade flows with the available mirror data.[15] Second, if one is reluctant to replace missing values with the available mirror flows, at a minimum, the mirror data should be used to ensure that the missing trade values are treated as missing indeed, instead of being replaced with zeros. Replacing missing values with zeros is standard in many gravity papers, however, this practice may prove dangerous given the large number of available export data that can be used to replace missing imports.

Recommendation 3: Use disaggregated data.

The level of aggregation in gravity estimations should be determined by the purpose of the analysis and by data availability. As discussed in Section 2, the trade gravity model can be derived at any level

[14] We are aware that some authors use the average of the mirrored import and export data. Based on the underlying gravity theory and our previous discussion, we would not recommend that. In addition to customs, insurance, and freight, there are other reasons for the mirrored data on exports and imports to be different, e.g., misreporting, and different product valuation. It is also possible that the value of c.i.f. imports may be smaller than the corresponding value of free-on-board (f.o.b.) exports. From that perspective, we believe that there is significant scope for improvements in the existing trade datasets. We also know that constructing consistent international trade data for many products, countries, and years is a difficult task. Thus, conditional on selecting a reliable raw dataset, we maintain our recommendation for using gravity data on goods imports and services exports.

[15] This may be particularly useful in the case of developing countries, which sometimes do not report at all or if they report, the data may be of lower quality and less reliable than the corresponding mirror flows data that is provided by their developed trade partners. We also note that the use of exporter-time and importer-time fixed effects, which we will recommend in the next section, will further mitigate such concerns.

of aggregation. An important implication of this 'separability' property of structural gravity is that gravity estimations can also be performed at any desired level of aggregation, e.g., from tariff-line data at 8- or 10-digit product code to aggregate country-level data and also for goods or services. While estimating gravity with aggregated data is often convenient, e.g., for computational or presentation reasons, our recommendation is to consider using disaggregated trade data whenever available.

The motivation for this recommendation is as follows. First, many trade policies, e.g., tariffs, are imposed and implemented at a very disaggregated level. Second, even if policies are implemented at the aggregate level, e.g., complete trade sanctions, their effects could substantially differ across products, industries, and sectors. Therefore, the estimates of aggregate policy effects may mask significant heterogeneity. This motivation is consistent with recent analysis from Breinlich et al. (2024), who also argue in favor of the use of disaggregated data and demonstrate that, when estimated with PPML, the gravity estimates that are obtained with aggregate data can be interpreted as trade-weighted averages of the corresponding disaggregated-level parameters. Third, even if the objective is to obtain aggregate estimates of the effects of certain policies, our recommendation is to pool, rather than aggregate, the disaggregated data because (i) this will improve estimating efficiency, (ii) it may allow estimating the effects of certain policies that are not possible to obtain with aggregated data, and (iii) it is preferable to obtain a common/aggregate estimate with pooled disaggregated data than to aggregate ex-post disaggregated estimates. Finally, for trade policies that are measured at the disaggregated level (e.g., tariffs), we recommend obtaining disaggregated estimates and then aggregating them to the desired level using weights based on the structural gravity model (Yotov et al., 2016) or ex-post trade weights (Breinlich et al., 2024). We elaborate on these points in Subsection 3.3.

Recommendation 4: Use panel data for consecutive years.

There are at least six good reasons to use panel data for gravity estimations.[16] First, from a pure econometric perspective, adding more years would improve estimation efficiency since the number of observations will grow faster than the number of gravity parameters (even in the presence of all fixed effects that we will recommend using in the next subsection). Second, as demonstrated in the previous section, the use of panel data is motivated by and consistent with gravity theory. Third, the use of panel data allows for flexible and comprehensive modeling of bilateral trade costs, e.g., with country-pair fixed effects that fully control for all possible time-invariant bilateral determinants of

[16]We refer the reader to Hsiao (2003) and Baltagi (2021) for a general discussion of the benefits of using panel data.

trade (e.g., Egger and Nigai, 2015; Agnosteva et al., 2019). Fourth, on a related note, the pair fixed effects will mitigate possible endogeneity concerns with the time-varying gravity variables (e.g., Baier and Bergstrand, 2007). Fifth, using panel data allows for modeling and capturing the evolution and adjustment of bilateral trade costs and the effects of various policies over time.[17] Finally, relatively long panel data (e.g., for more than 20 years) is now readily available at various levels of aggregation and for both goods and services.

According to Cheng and Wall (2005), *"[f]ixed-effects estimation is sometimes criticized when applied to data pooled over consecutive years on the grounds that the dependent and independent variables cannot fully adjust in a single year's time."* (p. 52). Similarly, in the context of RTAs, Trefler (2004) also criticizes the use of data pooled over consecutive years. As a result, using time-interval data, instead of data for consecutive years, has become a standard practice for gravity estimations (Yotov et al., 2016).[18] Most recently, Egger et al. (2022) challenged the practice of using time-interval or averaged data for estimating the effects of RTAs with the gravity model because much of the data are unnecessarily discarded and, as a result, some parameters are estimated less precisely, while others cannot be estimated at all. Egger et al. (2022) also argue that using time-interval data (i) may lead to biased estimates of the short-run as well as the long-run effects of RTAs and (ii) may not be appropriate to capture certain anticipation as well as phasing-in RTA effects.

Against this backdrop, we recommend using panel data for consecutive years. However, we also recognize that obtaining yearly estimates of the effects of various policies within a gravity setting with a rich fixed effects structure may be challenging in terms of precision and that the annual policy estimates may be subject to the influence of outlier observations. Therefore, we also recommend that, within the consecutive-year data, researchers should experiment by constraining the estimates of the effects of the policies of interest to be common across 2- or 3-years (Egger et al., 2022). Finally, if the estimates are obtained with time-interval data, researchers should be mindful of the pitfalls of discarding informative data and lost estimation efficiency.

We conclude by highlighting two related caveats when panel data is not available. Both challenges are due to the fact that without panel data one cannot include country-pair fixed effects in the econometric gravity model. As a result, it is not possible to comprehensively account for all possible

[17] We offer a more thorough discussion on the heterogeneity and adjustment of trade costs over time in Section 3.3.

[18] Trefler (2004) uses 3-year intervals, Anderson and Yotov (2016) use 4-year intervals, Cheng and Wall (2005) and Baier and Bergstrand (2007) use 5-year intervals, while Olivero and Yotov (2012) experiment with 3- and 5-year intervals.

time-invariant trade costs, and the latter should be proxied with observable variables. Several proxies (e.g., distance, contiguous borders, common official language, and colonial ties) have established themselves as 'standard' time-invariant gravity variables. However, as demonstrated by Egger and Nigai (2015) and Agnosteva et al. (2019), while these variables *correlate* very well with bilateral trade costs, they do not match *the level* of trade costs.[19] The implication is that one should be careful how to model the time-invariant trade costs in cross-section gravity settings. Our recommendation is to allow for heterogeneous effects of the standard gravity variables across various dimensions, and we elaborate on this recommendation in Subsection 3.3. On a related note, we caution the reader to the fact that it is much harder to address trade policy endogeneity in cross-section settings, and we discuss possible approaches to address this challenge in *'Recommendation 8'* from the next section.

Recommendation 5: Include domestic trade data.

To motivate this recommendation, we follow Yotov (2022) who argues that gravity equations should be estimated with domestic trade flows because: this is consistent with trade theory of the intensive margin of trade (Eaton and Kortum, 2002; Anderson and van Wincoop, 2003) and the extensive margin of trade (Anderson and Yotov, 2020a);[20] it allows for estimation of the effects of international borders and home biases (Anderson et al., 2018b; Ramondo et al., 2016) and heterogeneous domestic and regional trade costs (Anderson et al., 2018a; Donaldson, 2018); it allows for a systematic analysis of the determinants of domestic trade costs (Gurevich et al., 2021; Beverelli et al., 2024); it allows for country-specific asymmetries in the vector of international trade costs (Felbermayr and Yotov, 2021); it allows for identification of the trade-diversion effects of bilateral trade policies (Dai et al., 2014), the extraterritorial effects of sanctions (Kwon et al., 2022), the effects of non-discriminatory trade policies (Heid et al., 2021), the effects of country-specific characteristics (Beverelli et al., 2024), and the country-specific effects of trade policies (Felbermayr et al., 2024b); it leads to a solutions to "The Distance Puzzle of International Trade" (Yotov, 2012), "The Missing Globalization Puzzle" (Bergstrand et al., 2015), the puzzle that "Larger Countries Should Be Richer than Smaller Countries"

[19] Specifically, the standard gravity variables usually over-predict the trade costs for rich countries and under-predict the trade costs for poor countries. Adam et al. (2024b) perform an analysis of variance to investigate different trade cost components. They show that asymmetries play a role in bilateral trade costs and that the remaining residual variation is larger when less developed countries trade.

[20] While most of our recommendations apply for gravity estimations on the extensive margin of trade, the focus of this paper is exclusively on the intensive margin of trade. We refer the reader to Helpman et al. (2008b), Chaney (2008), Egger and Larch (2011), Egger et al. (2011), and Anderson and Yotov (2020a) for gravity analysis that incorporates the extensive margins of trade.

(Ramondo et al., 2016), and the puzzle of "The Missing WTO Effects" (Rose, 2004; Larch et al., 2024a).

Despite the theoretical motivation and multiple empirical benefits of estimating gravity with domestic (in addition to international) trade flows, most gravity estimations are still performed with international data only. At least in part, this is due to the fact that many researchers are not aware of the availability of data on domestic trade. The construction of domestic trade flows has evolved over time. Some studies have calculated domestic trade as the difference between GDP and total exports, e.g., Yotov (2012) and Larch et al. (2024a), while others have relied on input-output data, e.g., Felbermayr et al. (2018), Larch et al. (2018a), and Felbermayr and Steininger (2019). While these are viable options, one has to bear in mind that: (i) GDP is measuring value added, whereas trade flows are measured on a gross basis. Thus, comparing the two variables is not consistent and may lead to many negative values for domestic trade. Moreover, GDP cannot be used for constructing sectoral domestic sales;[21] and (ii) The input-output tables often also adjust raw data to balance the tables, which makes these tables/data less suitable for estimation. Most recently, Borchert et al. (2021, 2022) rely on gross production data to construct the *International Trade and Production Database for Estimation* (ITPD-E), which includes consistent domestic and international trade flows for many countries (more than 250), many industries (170, including agriculture, mining and energy, manufacturing, and services), and many years (1986-2019). Thus, the domestic trade data limitations have been relaxed, and we see more and more trade gravity applications that rely on data that includes international and domestic trade flows.[22]

While we are convinced of the benefits of including domestic trade flows in gravity estimations, we emphasize that estimating gravity on international trade flows only is not wrong. One also has to keep in mind that including both flows at the same time relies on the assumption that domestic trade flows and international trade flows follow the same data-generating process.[23] Nevertheless, it is still useful to be aware of the benefits and implications of including domestic trade flows in the gravity model and to acknowledge potential biases with some of the gravity estimates, e.g., downward biases in the

[21] Despite these inconsistencies, Campos et al. (2021) compare the effects of trade agreements estimated using data based on different constructions of domestic sales and find little difference in the resulting estimates.

[22] Moreover, motivated by the trade literature, we also see some data and papers that employ domestic flows beyond trade, e.g., for migration (Beverelli, 2021), investment (Carril-Caccia et al., 2023), cross-border patents (LaBelle et al., 2023), and Bradley et al. (2023) for gravity for mergers and acquisitions.

[23] If one believes that the data-generating process for domestic trade flows is different than the one for international trade flows, then one probably should estimate only with international trade flows data (or have country-specific, time-varying domestic sales fixed effects, which leads to the same results as estimating without domestic trade flows). Note, however, that the inclusion of domestic trade flows is guided by the theories underlying the gravity equation and opens up the possibility of estimating differential effects of trade costs on international versus domestic trade flows.

estimates of RTAs (Dai et al., 2014) and WTO membership (Larch et al., 2024a), that may be obtained with international trade data only. Finally, we note that introducing domestic trade flows requires proper modeling of the domestic trade costs, and we discuss this challenge in Recommendations 8 and 9 of the next subsection.

3.2 Gravity estimation

This section includes six recommendations related to the estimator, the covariates, and, more broadly, the econometric gravity specification. Some of our recommendations are guided by theory, while others are motivated by developments in the econometrics literature or by practical considerations.

Recommendation 6: Estimate gravity in its multiplicative form using PPML.

Two natural estimating options arise based on the theoretical gravity model from Section 2: (i) to log-linearize it, or (ii) to estimate equation (2) in its multiplicative form (rather than log-transforming it).[24] Perhaps not surprisingly, starting with Tinbergen (1962), the more popular option has been to log-linearize the gravity equation and end up with a specification linear in parameters that can be estimated with the standard ordinary least squares (OLS) estimator. Even though already early on it has been recognized that the gravity model can be estimated multiplicatively (Pöyhönen, 1963, 1964) and the Poisson estimator has been employed for this purpose in the literature on spatial interactions (Flowerdew, 1982; Flowerdew and Aitkin, 1982; Fotheringham and Williams, 1983), it was only relatively recently when the Poisson Pseudo Maximum Likelihood (PPML) estimator established itself as the leading gravity estimator owing to the work of Santos Silva and Tenreyro (2006).[25]

We have identified seven main reasons for using PPML for gravity estimations. First, and most important, PPML leads to consistent estimates of the multiplicative gravity model in the presence of heteroskedasticity, while OLS will likely lead to inconsistent estimates in this case, as the log-transformation of the trade flows changes the properties of the error terms (see Flowerdew, 1982; Santos Silva and Tenreyro, 2006).[26] The error terms in the log-linear gravity equation are typically

[24]Mechanically, the multiplicative gravity model is estimated after transforming its right-hand side into an exponential form, where the continuous trade cost proxies (e.g., bilateral distance) enter the exponent in logs, while the indicator gravity covariates (e.g., trade agreements and sanctions) enter it directly. We refer the reader to Yotov et al. (2016) for interpretation of the estimates of the standard gravity covariates.

[25]We refer the reader to Santos Silva and Tenreyro (2022) and the 'Log of Gravity' website for useful information about the properties and misconceptions of the PPML estimator for gravity estimations.

[26]Similar arguments have been made for the estimation of wage equations (Blackburn, 2007) and of production functions (Sun et al., 2010), and we refer the reader to Manning (1998), Manning and Mullahy (2001), and Ai and Norton (2000) for a general discussion of the problem of a logged dependent variable in the presence of heteroskedasticity.

heteroskedastic, as the conditional variance, i.e., the variation of trade flows, is different between bilateral trading pairs, even after conditioning on size, geography, and differences in trade costs. Importantly, due to the non-linear transformation by logs, heteroskedasdicity likely leads to inconsistent OLS estimates of the log-linearized model—different to the situation where the model is linear in parameters. For this reason, using robust standard errors, as often employed in the presence of heteroskedasticity, will not lead to consistent estimates of the log-linearized model. There have been alternative attempts, in different fields, to tackle the problem of heteroskedasticity.[27] However, most of these methods are either computationally intensive or impose functional-form restrictions and have not established themselves as standard tools in the literature.

The second attractive property of PPML is that, due to its multiplicative form, it takes into account information contained in the zero trade flows. Moreover, it performs well, even when the proportion of zeros is large (Santos Silva and Tenreyro, 2011), which is particularly important for gravity estimations for services and with disaggregated data.[28] Motivated by the analysis of Felbermayr and Kohler (2006) and Helpman et al. (2008a), who document the large share of zero trade flows for earlier periods, we use the second release of the International Trade and Production Database for Estimation (ITPD-E-R02) (Borchert et al., 2021, 2022) to visualize the evolution of the share of country pairs with trade in both directions, trade in one direction only, and no trade at the industry-level, 1988-2019, and to highlight several important patterns.

Figure 1 reports our findings. The top panel of the figure is based on the full dataset, where the left figure shows the zeros of the disaggregated data, i.e., based on 170 industries, and the right figure the zeros when aggregating over all 170 industries. The main messages from the top panel are that (i) the number of zeros has decreased over time, however, (ii) the zeros have remained relatively stable since 2007, (iii) the fraction of zeros is very large, i.e., more than 80%. As can be expected, the number

[27] For example, Stronge and Schultz (1978) and Fotheringham and Williams (1983) use iterative weighted methods in the context of the spatial interaction literature. In the trade literature, Anderson and van Wincoop (2003) transform the dependent variable into size-adjusted trade and use non-linear least squares, while Eaton et al. (2013) develop a finite-sample version of the gravity framework and, using bilateral trade shares as the dependent variable, estimate gravity with multinomial pseudo-maximum likelihood (MPML) following Gourieroux et al. (1984). Based on Eaton et al. (2013), some quantitative economic history papers, e.g., Barjamovic et al. (2019) and Flückiger et al. (2021), rely on the count nature of the data as an important argument to employ shares rather than levels. Sotelo (2019) shows that the difference between MPML and PPML, i.e., using trade shares or levels, can be seen as different weighting of observations, which also applies to other estimators, e.g., Gamma PML, suggested to estimate the gravity equation in its multiplicative form (Manning and Mullahy, 2001; Santos Silva and Tenreyro, 2006; Egger and Staub, 2016).

[28] Note that, while PPML can incorporate zero trade flows in the estimation, it assumes that the data-generating process for zero trade flows and positive trade flows is, up to the explicitly modeled differences in trade costs, the same. See for frameworks explicitly distinguishing between the extensive and intensive margin Helpman et al. (2008a) and Egger et al. (2011).

of zeros will be significantly larger with more disaggregated data, especially for services trade, and (iv) the share of zeros decreases to about 20% when aggregating over all 170 industries, highlighting the fact that the zeros are more prevalent in disaggregated data.[29] The figures of the bottom panel of Figure 1 use the same data, however, we have kept only the zeros that were not dropped when we estimated a PPML gravity model with three-way fixed effects. The main message from this panel is that the number of zeros that are 'relevant' for gravity estimations is significantly smaller, i.e., about 40% in the disaggregated data and about 20% in the aggregate data.

There have been various attempts to account for the zero trade flows in gravity estimations (Martin and Pham, 2020). For example, some authors have added a small, arbitrary value to replace the zero trade flows. As noted in Head and Mayer (2014), however, this approach should be avoided because the results depend on the units of measurement and the interpretation of the gravity coefficients as elasticities are lost. Eaton and Tamura (1994), Martin and Pham (2008) and Felbermayr and Kohler (2006) rely on the Tobit estimator, which treats zeros as a corner solution, This approach has been criticized because, (i) given the still large share of zero bilateral trade flows, more structure than just assuming zeros to be corner solutions is probably warranted, and (ii) the arbitrary selection of the Tobit threshold. Helpman et al. (2008a) propose a theory where zero trade flows arise from the selection of firms into exporting, estimating the resulting specification in two steps employing an Heckman (1979) sample-selection procedure.[30] While nicely motivated by theory, the resulting estimation procedure depends on strong functional form assumptions, needs a good exclusion restriction for the first stage, and is not readily extended to panel data (Santos Silva and Tenreyro, 2015).[31] Egger et al. (2011) propose a two-part framework, which models the extensive and intensive margin as two separate processes (see also Felbermayr et al., 2015, for an application of this procedure to evaluate the proposed Transatlantic Trade and Investment Partnership). Finally, it is possible to use the inverse hyperbolic sine function. However, this procedure has to be applied with caution because it is a non-linear transformation (as is the log-transformation), which means that with heteroskedastic trade data, one may end up with inconsistent estimates (Santos Silva and Tenreyro, 2006; Aihounton and Henningsen,

[29] Helpman et al. (2008a) report about 40% trade in both directions in 1997 for aggregate trade, while we have about 35%. We likely obtain a larger share of zeros as our figures are based on 258 countries, whereas Helpman et al. (2008a) cover 158 countries.

[30] The first step is a Probit model determining the probability of whether two countries trade or not, i.e., the extensive margin of trade. In the second step, the intensive margin, i.e., the volume of trade for country pairs that have positive trade, is estimated with OLS, taking into account the selection estimated in the first stage.

[31] Eaton et al. (2012) amend the standard heterogeneous-firm trade model by replacing the assumption of firms as points on a continuum with only an integer number of firms, and their theoretical model also allows for zero trade flows.

2020; Bellemare and Wichman, 2020; Chen and Roth, 2023).

We conclude our discussion of the zeros in gravity regressions with two practical observations. First, as noted earlier, while the fraction of zeros in the data is large, more than half of the zeros are 'irrelevant' in gravity regressions that are estimated with exporter-time, importer-time, and directional country-pair fixed effects. A practical implication based on this observation is that researchers may consider dropping the 'irrelevant' zeros in advance to avoid dealing with unnecessarily large datasets. Second, we have noticed that including the zeros or not in PPML gravity estimations has little or no impact on the estimated parameters. The reason is that when estimating with PPML, trade flows enter in levels (in contrast to logs, as in the log-linearized model estimated with OLS), and larger flows tend to have larger residuals leading to a larger influence in the estimation. Based on this, we recommend estimating the model with PPML on the full sample that includes the zeros but also experimenting with PPML on positive trade values only.

The third attractive property of PPML is that, despite being a non-linear estimator, its convergence is fast and reliable. Standard estimation commands, e.g., `poisson` and `glm` by Stata, have been available for a long time, and they were improved by the `ppml` command of Santos Silva and Tenreyro (2006). Despite this progress, however, until recently, the PPML estimator was criticized for being unable to converge for large datasets with many years and countries (Glick and Rose, 2016). The primary reason for convergence difficulties with gravity models is the large set of fixed effects, which are not trivial to handle even in linear settings and is an even bigger challenge in non-linear models such as PPML.[32] To address this challenge, Larch et al. (2018b) introduce the `ppml_panel_sg` command and, most recently, Correia et al. (2020) develop the `ppmlhdfe` command, which can handle fast and effectively a huge number of fixed effects and is our choice for trade gravity estimations.[33]

We conclude this discussion with a couple of practical tips to further speed up convergence in gravity estimations. First, we recommend including all dummy variables used as controls in the `absorb` option of the `ppmlhdfe` command. It is also possible to retain and interpret the estimates

[32]Following the recommendations of Yotov et al. (2016), which we reinforce later in this section, most gravity models are estimated with exporter-time, importer-time, and country-pair fixed effects. Using a balanced dataset with 60 countries and 30 years, implies $60 \times 30 + 60 \times 30 + 60 \times 60 = 7,200$ fixed effects (ignoring the dropping of fixed effects due to perfect collinearity). Going a step further, if disaggregated data are pooled across industries, e.g., in the newly released ITPD-E-R02, which has 256 countries, 170 industries, and 34 years of data, this may lead up to $256 \times 170 \times 34 + 256 \times 170 \times 34 + 256 \times 256 \times 170 = 14,100,480$ fixed effects.

[33]Besides the PPML developments in the trade literature, there are also contributions dealing with high-dimensional fixed effects in generalized linear models, e.g., the Exponential or the Gamma model (Stammann, 2017; Hinz et al., 2019). We also refer the reader to the `reghdfe` command of Correia (2016) for OLS with high-dimensional fixed effects.

of these dummy variables by saving the estimates and then transforming them relative to the proper reference group. However, this may not be necessary if the dummy variables are used purely as control variables. If the dummy variables are the covariates of interest, then they should probably be included explicitly in the model. Another practical tip is to consider rescaling (but not rounding) the dependent variable. Third, we recommend adjusting some of the default options of the `ppmlhdfe` command if speed is an issue, such as using "separation(fe)". Finally, researchers should be mindful of observations that may be 'almost' perfectly collinear with the rich set of fixed effects in the gravity model, e.g., pairs with only a few trade flows.

The fourth attractive property of PPML is that the actual distribution of the error term does not have to follow a Poisson distribution. PPML is a pseudo- (or quasi-) maximum likelihood estimator and, since it belongs to the family of generalized linear models, it will lead to consistent estimates as long as the conditional mean is correctly specified (McCullagh and Nelder, 1989; Cameron and Trivedi, 2005).[34] This also implies that the dependent variable does not have to be a count variable, even though some early (Flowerdew and Aitkin, 1982) and some recent (Barjamovic et al., 2019; Flückiger et al., 2021) contributions emphasize the count nature of the dependent variable. Other estimators of the generalized linear model family, e.g., Gamma or negative-binomial, share this property, however, they are not invariant to changes in the units of measurement of the dependent variable.

The fifth attractive property of PPML is related to the incidental parameter problem (IPP), which may arise in non-linear gravity settings with many fixed effects (Neyman and Scott, 1948). In their study of fixed effects estimators for nonlinear panel data models with two-way fixed effects, Fernández-Val and Weidner (2016) find that PPML has the unique property that it is asymptotically unbiased as both dimensions of the panel become large, whereas other estimators are consistent but asymptotically biased. Weidner and Zylkin (2021) extend their analysis to a panel data gravity setting with three-way fixed effects, and show that, as in the two-way case, the three-way PPML estimator is consistent. However, with fixed T, there is an asymptotic bias which can be corrected by a Jackknife bias correction, a bootstrap bias correction (Zylkin, 2024), and via an analytical bias correction, which can be implemented with the `ppml_fe_bias` command of Weidner and Zylkin (2021).[35] As an

[34] Santos Silva and Winkelmann (2024) show that PML estimators provide an optimal approximation to the conditional expectation using the weighted mean squared error as a criterion function. PPML also correctly identifies the average marginal effect if the regressor is normally distributed.

[35] When both N and T are large, there is also no asymptotic bias in the three-way fixed effects setting. While the two-way fixed effects have no asymptotic bias for the point estimates, the standard errors are biased. Here, the approaches discussed concerning clustering of standard errors, such as Jackknife and bootstrap confidence interval methods discussed

alternative approach to overcome the incidental parameter problem for the estimation of gravity models with two-way fixed effects, Jochmans (2017) proposed using a GMM estimator based on the quasi-differencing moment conditions of Charbonneau (2013). Recently, Yang and Zhang (2023) extend the approach used in Jochmans (2017) to allow the estimation of three-way gravity models.

The sixth attractive property of PPML is that its constant variance-to-mean ratio (CVMR) assumption is mostly satisfied in the data.[36] This is shown by Kwon et al. (2022), who rely on recent iterated GMM methods (Hansen and Lee, 2021) to test the PPML's CVMR assumption and propose the Generalized PPML (G-PPML) as a complementary estimator. The main practical implication from the empirical analysis of Kwon et al. (2022) is that, even though the PPML's CVMR assumption is satisfied most of the time, there are also instances when it does not hold. In this case, G-PPML will be more efficient and the point estimate differences between PPML and G-PPML can be as high as 40%. Based on this, our recommendation is to use the routine from Kwon et al. (2022) to test the validity of the CVMR assumption and consider using the G-PPML estimator in case the assumption is not satisfied.[37]

The seventh attractive property of PPML is that it is perfectly consistent with the structural gravity model from Section 2, i.e., the exporter(-time) and importer(-time) fixed effects in the gravity model capture exactly and only the corresponding theoretical gravity terms that they are supposed to account for. The history of the so-called 'additive property' of PPML dates back to Fotheringham and Williams (1983) Davies and Guy (1987), but, more recently, Arvis and Shepherd (2013) and, especially, Fally (2015) highlight it in the trade gravity context by showing that the origin-(year) and destination-(year) fixed effects will ensure that total predicted exports and imports for each country from the model will be equal to the total observed exports and imports for each country. From a theoretical perspective, this means that the estimation procedure takes into account production and consumption constraints. From a practical perspective, the 'additive' PPML property implies that the PPML estimator may serve as a non-linear solver to recover the structural multilateral resistances directly from the estimates of the exporter(-time) and importer(-time) fixed effects in gravity regressions. Anderson et al. (2018b) capitalize on this property to propose a simple Stata procedure to *estimate* benchmark general equilibrium effects (GE) with the gravity model.

by Pfaffermayr (2021) as well as the `ppml_fe_bias` command, could also be helpful.

[36] In contrast, the corresponding OLS and Gamma-PML assumptions are always rejected.

[37] The two Stata commands are (i) `cvmrtest`, which tests the validity of the CVMR assumption, and (ii) `gppmlhdfe`, which implements the G-PPML estimator.

Recommendation 7: Use exporter(-sector-time) and importer(-sector-time) fixed effects.

Guided by theory (see Section 2), our recommendation is to use (i) exporter and importer fixed effects in single-sector (or aggregate) cross-section settings (Hummels, 1999; Egger, 2000; Rose and van Wincoop, 2001; Redding and Venables, 2004), (ii) exporter-time and importer-time fixed effects in single-sector (or aggregate) panel settings (Baldwin and Taglioni, 2006; Olivero and Yotov, 2012), and (iii) exporter-sector-time and importer-sector-time fixed effects in panel settings with pooled sectoral data (Anderson and van Wincoop, 2004; Costinot et al., 2012b; French, 2016, 2019; Beverelli et al., 2024).[38]

Not following this recommendation may lead to biased and inconsistent gravity estimates, and we offer four reasons for using exporter(-sector-time) and importer(-sector-time) fixed effects. First, from a pure econometric perspective, the motivation for using these fixed effects is that they will fully control for and absorb all possible observable and unobservable determinants of bilateral trade flows on the exporter and the importer side. Second, from a computational point of view and as we already mentioned earlier, there are now fast and flexible estimation commands that can handle many fixed effects in linear gravity models, e.g., the `reghdfe` command of Correia (2016) for OLS gravity, and non-linear gravity settings, e.g., the `ppmlhdfe` command of Correia et al. (2020) for PPML gravity estimations. Third, from a theoretical perspective, the country fixed effects will account for the unobservable multilateral resistances of Anderson and van Wincoop (2003). Motivated by the importance of the MRs to obtain better estimates of the bilateral trade costs in gravity models (e.g., Anderson and van Wincoop, 2003), Baldwin and Taglioni (2006) dub the omission to control for the MRTs the *'Gold Medal Mistake'* in gravity estimations. Finally, as discussed earlier and owing to the 'additive' PPML property, the estimates of the exporter(-sector-time) and importer(-sector-time) fixed effects correspond exactly to the structural gravity terms that they are designed to capture.[39]

Over the years, a few alternative methods have been proposed to control the MRs. For example, Anderson and van Wincoop (2003) use an iterative non-linear least squares approach with constraints. In most cases, the use of fixed effects dominates the iterative procedure of Anderson and van Wincoop

[38] We refer to sectors loosely as the fixed effects could be exporter-product-time and importer-product-time when the data are pooled across products.

[39] The implication of this is that if the full set of exporter(-sector-time) and importer(-sector-time) fixed effects are replaced with the structural variables, the fit of the gravity model, as well as the estimates of all gravity variables, will remain unchanged. This is an important property that distinguishes gravity estimations with fixed effects from other prominent fixed effects, e.g., the Mincerian labor regression, where the fixed effects do not have a structural interpretation.

(2003), as the correlation between trade friction variables in the model and unobserved country/region-specific determinants leads to inconsistent parameter estimates in the structural approach.[40] However, there are some instances when the structural approach to construct the MRs may be useful. One such example is estimating gravity for a single country, where using proper fixed effects is impossible (Lampe et al., 2023). Others have used the so-called 'remoteness indexes', constructed as GDP-weighted averages of bilateral distance (or trade costs) (e.g., Wei, 1996; Davis and Weinstein, 2002; Baier and Bergstrand, 2009; MacDonald et al., 2020). However, Head and Mayer (2014) criticize such reduced-form approaches as they "bear little resemblance to [their] theoretical counterpart." (p. 150) and, more recently, Felbermayr and Yotov (2021) demonstrate that the a-theoretical remoteness indexes do not perform well in predicting bilateral trade flows. Finally, some papers have proposed transformations of the dependent variable, i.e., via the so-called 'ratio methods', to eliminate the MRs from the estimating model (Head and Ries, 2001; Head et al., 2010; Novy, 2013a). Such transformations may lead to challenges with the interpretation of the gravity estimates.

We conclude the analysis of the country fixed effects with a discussion of two practical considerations. First, we underscore the importance of theory to guide the selection of the fixed effects for gravity estimations. An intuitive approach to this challenge is to interpret the MRs as price indexes on the consumer and on the producer side, which vary over time and could be very heterogeneous across products, industries, and sectors. Based on this, for example, using exporter, importer, and time fixed effects in a single-sector panel setting will not control for the MRs properly. Instead, the fixed effects should be exporter-time and importer-time. Similar logic applies to the use of exporter-sector-time and importer-sector-time fixed effects in panel settings with pooled sectoral data.[41]

Second, a downside of using fixed effects is that they will not allow the identification of the effects of any country-specific variables. We are aware of three possible methods to approach this challenge. Some studies, e.g., Redding and Venables (2004), Head and Mayer (2014), Baker and Fortin (2001), have used a two-step procedure, where the fixed effects from the first-stage gravity estimation are regressed on various country-specific variables. The challenges with this approach are that (i) it

[40] As discussed by Egger and Larch (2012a), this can be tested using a Hausman test. Egger and Larch (2012a) also suggest a variant of the Anderson and van Wincoop (2003)-procedure that combines the advantages of fixed effects estimation and the structural AvW model.

[41] We also note that the size variables in the theoretical gravity model vary at the same dimensions as the MRs. Therefore, for example, using GDP variables in disaggregated gravity regressions is inconsistent with theory. Using GDP to proxy for size is not appropriate even in regressions with aggregate data. The reason is that GDP is measured as value-added, while the theory-consistent size variables (e.g., output and expenditure) are on a gross basis.

does not fully control for the structural MRs, (ii) the second-stage dependent variable is generated from a regression, and (iii) when the first-stage model is estimated with PPML, then the structural gravity variables would explain all the variation in the fixed effects. Second, consistent with theory (Yotov, 2022), using domestic trade flows allows for identifying various country-specific and non-discriminatory determinants of trade flows (see 'Recommendation 5'). We elaborate on the possibility of including country-specific variables in the gravity model in 'Recommendation 10'. Finally, Freeman et al. (2021) capitalize on gravity theory to propose a structural approach, where the gravity fixed effects are replaced by country-specific variables, including the MRs.[42] In combination with the use of domestic trade flows, the methods of Freeman et al. (2021) can deliver estimates of the uniform and the discriminatory (i.e., on international relative to domestic trade) effects of country-specific variables.[43]

Recommendation 8: Employ asymmetric country pair fixed effects.

The recommendation to estimate the gravity model with pair fixed effects, in addition to the country-specific fixed effects that we discussed in the previous recommendation, is not motivated by theory, but it is not new (e.g., see Egger and Pfaffermayr, 2003). There are at least two good reasons to do so. First, the country-pair fixed effects will fully control for and absorb all time-invariant bilateral trade costs, including any time-invariant domestic trade costs when gravity is estimated with domestic sales. While it is well-established that the standard gravity variables (e.g., bilateral distance, contiguity, common language, etc.) correlate very well with the bilateral trade costs, it is also documented now (e.g., Egger and Nigai, 2015; Agnosteva et al., 2019), that the standard gravity variables do not predict the level of trade costs very well. Specifically, and as noted earlier, they usually over-predict the trade costs for rich countries and under-predict the trade costs for poor countries. This issue is mitigated significantly by using country-pair fixed effects. Based on our experience and in combination with a set of widely used time-varying bilateral trade cost proxies (which we discuss in our next recommendation), the bilateral trade costs from gravity regressions with three-way fixed effects

[42] Specifically, Freeman et al. (2021) use the market-clearing condition in the structural gravity model to express the value of gross output as a function of a series of country-specific characteristics, including factors of production and the outward multilateral resistances. Then, they use the resulting expression to replace the exporter-(time) fixed effect with the corresponding theoretical term. The 'additive' property of the PPML estimator, which we discussed earlier, ensures consistency between gravity theory and the proposed estimation procedure.

[43] A byproduct of the methods from Freeman et al. (2021) is that they deliver estimates of the trade elasticity, including for services, without the need to use price or tariff data. See also Bergstrand et al. (2013) for an alternative approach to estimate the trade elasticity without tariffs within a structural gravity framework and based on asymmetric bilateral trade costs.

(e.g., importer-time, exporter-time, and country-pair) match the corresponding calibrated bilateral trade costs very well. Thus, in addition to controlling for many potentially omitted variables, the pair fixed effects will deliver reliable estimates of bilateral trade costs, which are important for proper bilateral trade predictions (Davis and Weinstein, 2002; MacDonald et al., 2020), when the vector of bilateral trade costs is used in the calibration of models used for counterfactual analysis (Anderson et al., 2019; Dingel and Tintelnot, 2023), or to validate the gravity model by comparing estimated vs. calibrated bilateral trade costs.[44]

Second, by absorbing all possible time-invariant bilateral trade costs, the country-pair fixed effects will mitigate potential endogeneity concerns with the bilateral (trade) policy variables in the gravity model. The issue of trade policy endogeneity has been recognized in the trade literature for a long time (e.g., Leamer, 1988), and there are various reasons to believe that most if not all, trade policy estimates are subject to endogeneity criticisms and concerns. Simultaneity, e.g., the fact that countries that trade a lot are more likely to make efforts to liberalize bilateral trade further, is an obvious candidate. Alternatively, Trefler (1993) argues that higher levels of imports will lead to intensified lobbying activity for protection and, unless this is taken into account econometrically, the estimates of the effects of trade liberalization will be downward biased. Viewed again as an omitted variable bias, Baier and Bergstrand (2007) argue that the incentive to sign an RTA for governments is potentially larger if unmeasurable domestic regulations are high, leading to a positive correlation, while the unmeasurable domestic regulations are expected to decrease trade flows, leading to a downward bias.[45]

Even though the endogeneity issue has been recognized for a long time, the trade literature has struggled to find a convincing solution. The main difficulty in applying the standard IV approach to tackle endogeneity has been the lack of good instrumental variables, which determine the trade policy but are not correlated with trade flows (Baier and Bergstrand, 2007, 2004).[46] As an alternative, Baier and Bergstrand (2009) use the matching estimator of Abadie and Imbens (2006) for the three nearest neighbors, but note that the stable-unit-treatment-value assumption (SUTVA) is hard to defend in trade settings, where the GE effects of trade agreements are non-negligible. Egger and

[44] Anderson et al. (2019) propose a hybrid approach, dubbed 'ESTIBRATED' trade costs, which combines the 'ESTImates' from the identifiable trade costs and the error term to deliver 'caliBRATED' trade costs.

[45] See for a discussion of the omitted variable bias, for example, Cameron and Trivedi (2005), section 4.7.4.

[46] Alternative instruments have been proposed with mixed success, e.g., sectoral shares of value added (Lee and Swagel, 1997), the absolute value of the difference of the two countries gross trade flows (Baier and Bergstrand, 2002), log difference in GDP, the amount of intra-industry trade, the size of the bilateral trade deficit, similarity in capital-labor ratios, and an indicator that both countries are democracies (Magee, 2003), regulation cost variables and common religion (Helpman et al., 2008a), colonial ties (Egger et al., 2011), or the number of common RTA partners (Jochmans and Verardi, 2022).

Tarlea (2021) emphasize the lack of balancing in the data, i.e., the similarity of the country pairs in the control and treatment group in terms of the propensity scores, conditional on observables, and employ entropy balancing from Hainmueller (2012) in an attempt to overcome the problem. Finally, Baier and Bergstrand (2007) propose differencing the data or using bilateral fixed effects, since they will control for many of the potential causes of endogeneity (e.g., common history, the (average) level of trade flows, institutions, political similarity, etc.). The identification of the policy variables of interest is then solely based on variation over time between pairs of countries.[47]

We conclude with six practical recommendations. First, the fast and flexible estimation commands we discussed in the previous recommendation for handling exporter-time and importer-time fixed effects should handle the large set of country-pair fixed effects effectively. Second, suppose it is not possible to use country-pair fixed effects. In that case, one should carefully model the time-invariant bilateral trade costs, e.g., by allowing the effects of distance to vary by interval (e.g., as in Eaton and Kortum, 2002) or by the level of development to capture some systematic heterogeneity across the trading partners in the sample. Third, the country-pair fixed effects will not allow identification of the effects of any time-invariant variables, however, subject to the caveats that we discussed in the previous recommendation, it may be possible to implement a two-stage procedure, where the country-pair fixed effects from the first-stage are regressed on the time-invariant variables of interest (Anderson and Yotov, 2016; Anderson et al., 2020). Fourth, the country-pair fixed effects should be asymmetric, because (i) some trade costs are asymmetric by nature, e.g., the time of crossing Eastern European borders can be very different depending on the direction of trade, and (ii) because if the underlying time-invariant trade costs are indeed asymmetric but symmetric fixed effects are used, then the time-invariant asymmetries may be misattributed to the effects of bilateral policies (Baier et al., 2019). Fifth, we recommend that the asymmetric country-pair fixed effects are allowed to vary further by product (industry, or sector) when disaggregated data are pooled to perform the analysis (Beverelli et al., 2023). The motivation for this recommendation is that trade costs vary significantly across products, industries, and sectors, e.g., due to differences in transportation costs (agricultural products vs. services). Finally, it may be beneficial to consider allowing for time-variation in the pair fixed effects, especially for panels with long time dimensions. The motivation for this is twofold. First, Anderson and Yotov (2022) propose a model that implies the use interval-pair fixed

[47] For a general discussion of the average treatment effect (ATE) methods of Baier and Bergstrand (2007), we refer the reader to Wooldridge (2010).

effects with different interval lengths to capture "gravity in transition" from the short to the long run. Second, Baier and Standaert (2024) provide evidence that the assumption that the pair fixed effects are constant over time may not hold in the long run.

Recommendation 9: Carefully model bilateral trade costs.

The inability to properly model bilateral trade costs risks omitted variable bias and may lead to inconsistent and biased estimates. While theory takes a solid stand on the composition and functional form of the country-specific variables in the gravity model,[48] there is little theoretical guidance regarding the modeling and estimation of bilateral trade costs in the empirical gravity equation. Consistent with this observation, and even though distance is considered the most widely-used and most robust proxy for trade costs in the gravity model (Disdier and Head, 2008), Chaney (2018) notes that *"[w]hile the role of size is well understood, that of distance remains mysterious."* Anderson and van Wincoop (2004) offer an informative survey of trade costs, however, we believe there is a pressing need for an updated systematic discussion of the measurement and decomposition of bilateral trade costs in the context of the gravity model. Such an important task is beyond the scope of our paper and, instead, in what follows we offer several practical recommendations, based on empirical observations and motivated by the emergence of new gravity databases.

The bilateral trade costs can be decomposed into two broad types – time-invariant vs. time-varying. Traditionally, the time-invariant trade costs have been proxied by a set of few variables, including bilateral distance, contiguous borders, colonial ties, and common language. However, as discussed in our previous recommendation, the country-pair fixed effects used in most current gravity models will fully absorb and account for these variables. Importantly, based on our experience, the asymmetric country-pair fixed effects would allow for capturing the distribution of bilateral trade costs across pairs very well. Thus, we are in strong favor of using country-pair fixed effects whenever possible. In case pair fixed effects cannot be used, the reader should be aware that the standard gravity variables cannot capture the level of trade costs well and, therefore, one should consider allowing for heterogeneous trade costs across various dimensions.

Perhaps not surprisingly, most current gravity estimations focus on the effects of various time-varying determinants of trade flows. While the objective of the specific project should guide the inclusion of these variables and there have been many time-varying determinants of trade flows that

[48] In fact, we believe that this is the main reason to often refer to the gravity model of trade as 'structural gravity'.

have been included in the gravity model, we offer several broad practical recommendations. The first one is to include a set of time-varying control variables, which will not significantly decrease the sample size (see 'Recommendation 1'). Several time-varying variables, for which reliable data are now available for most/all countries, have proven to be important determinants of trade flows. These include RTAs (with data from Egger and Larch, 2008; Baier and Bergstrand, 2021)[49], WTO membership (with data from Mayer and Zignago, 2011; Gurevich and Herman, 2018), currency unions (with data from de Sousa, 2012) and economic sanctions (with data from Felbermayr et al., 2020; Syropoulos et al., 2024). We consider these time-varying variables to be standard because they have been studied by many authors and, accordingly, we recommend that they should be controlled for routinely in gravity regressions.[50]

We also recommend the inclusion of two other covariates in the gravity model – tariffs and a set of indicators that will capture globalization effects – however, these variables do not enter our list of 'standard' time-varying covariates due to data considerations. In addition to being an important impediment to trade, the role of tariffs for gravity estimations is special because, since tariffs are direct price-shifters (see Section 2), their estimate can be used to recover the trade elasticity (Heid and Larch, 2016; Fontagné et al., 2020; Anderson and Yotov, 2022). There are several challenges with the inclusion of tariffs in the gravity model: (i) they are endogenous; (ii) the data on tariffs may not be reliable (Teti, 2020); (iii) tariff evasion may lead to biased trade elasticity estimates (Egger and Larch, 2012b); and (iii) tariff aggregation is required in most cases. Against this backdrop, it is unsurprising that most gravity regressions do not include tariffs.[51]

Despite the common belief that the world has become 'smaller' (Leamer and Levinsohn, 1995) or 'flatter' (Friedman, 2006), and that the impact of distance is dead (Cairncross, 1997), the gravity estimates of the distance effects have remained remarkably stable over time (Disdier and Head, 2008). More generally, Coe et al. (2007) argue that the gravity model of bilateral trade cannot capture declining trade costs and conclude that *"globalization is everywhere but in estimated gravity models"*

[49] The continuously updated RTA data from Mario Larch's Regional Trade Agreements Database from Egger and Larch (2008) are available at https://www.ewf.uni-bayreuth.de/en/research/RTA-data/index.html.

[50] Without an attempt to provide an exhaustive list, some other time-varying determinants of trade flows that have been studied with the gravity model include non-tariff measures (Fontagné et al., 2005), transportation costs (Egger, 2005), trade missions (Head and Ries, 2010), piracy (Bensassi and Martinez-Zarzoso, 2012), commercial trade law (Gil-Pareja et al., 2020), COVID (Baldwin and Dingel, 2021), time as a trade barrier (Egger and Larch, 2013; Oberhofer et al., 2021), business travel costs (Soderlund, 2023), and international air travel (Wang et al., 2021).

[51] The inclusion of trade agreements in gravity regressions mitigates the challenge of omitting tariffs, however, our recommendation is to attempt to include both RTAs and tariffs in the gravity model, whenever possible. The reason for this suggestion is that RTAs should promote trade beyond just eliminating tariffs.

(p. 3). Capitalizing on two theoretical properties of the gravity model, i.e., that gravity can only identify relative trade costs and that the gravity model includes domestic trade flows, Yotov (2012) and Bergstrand et al. (2015) show that the gravity model can capture the declining effects of distance and the effects of globalization more broadly. The implication for gravity estimations is that when domestic trade data are available, the gravity model should include a set of time-varying indicator variables, which take a value of one for international trade and zero for domestic trade. These time-varying border indicators will capture common globalization trends and, as demonstrated by Bergstrand et al. (2015), their omission may lead to significant biases in the estimates of the effects of various determinants of bilateral trade flows (e.g., free trade agreements). Still, they can also be allowed to vary across different dimensions (e.g., Larch et al., 2023; Adam et al., 2024a), and we discuss possibilities to obtain heterogeneous trade cost estimates in Section 3.3.

The inclusion and identification of home biases, border effects, and globalization in the presence of country-time fixed effects are only possible (and recommended) when the gravity model is estimated with domestic trade flows. While, as discussed earlier, the inclusion of domestic trade flows may be very beneficial for various purposes, this also poses the challenge of modeling domestic trade costs. Similar to the comprehensive treatment of the time-invariant international trade costs with pair fixed effects, the pair fixed effects will also absorb any time-invariant domestic trade costs. The fixed effects will control for time-invariant home biases. In addition, as we just discussed, we also recommend allowing for globalization trends with time-varying border variables. Finally, depending on the data, the econometric specification, and the specific objectives of the analysis, we recommend the inclusion of explicit proxies for domestic trade costs, e.g., internal distance (Borchert and Yotov, 2017), domestic language (Gurevich et al., 2021), institutional quality and development (Beverelli et al., 2024), etc.

We conclude with four potential caveats. First, some trade policies have become excessively complex, e.g., most modern RTAs include many provisions. This may pose challenges with estimating their effects (e.g., loss of data, overfitting, multicollinearity, etc.), which should be addressed on a case-by-case basis, e.g., by using different methods for variable selection. For example, Breinlich et al. (2021) and Bergstrand and Paniagua (2024) propose data-driven methods for selecting the most important provisions from the World Bank's Deep Trade Agreements Database (Mattoo et al., eds, 2020). Second, even if a policy is non-discriminatory by design, its effects can be heterogeneous across various dimensions. We discuss trade cost heterogeneity in more detail in Section 3.3. Third, despite having theoretical gravity foundations that imply that the vector of bilateral trade costs should distinguish between fixed vs. variable trade costs, we are not aware of an established approach (and data) to tackle this issue, and we view this as an important gap in the gravity literature. Finally, we are aware of papers that construct bilateral gravity covariates by combining country-specific variables. This practice can lead to wrong conclusions, and we offer more details in our next recommendation, where we discuss the effects of country-specific variables.

Recommendation 10: Allow for non-discriminatory and country-specific trade costs.

Many policies that affect trade flows are non-discriminatory against trade partners (e.g., export subsidies) or country-specific (e.g., sanitary and phytosanitary standards (SPS)). However, according to Head and Mayer (2014), the effects of such policies cannot be identified within the gravity model: *"In the presence of importer and exporter fixed effects a variety of potentially interesting trade determinants can no longer be identified in a gravity equation. Notably, (1) anything that affects exporters' propensity to export to all destinations (such as having hosted the Olympics or being an island), (2) variables that affect imports without regard to origin, such as country-level average applied tariff, and (3) sums, averages, and differences of country-specific variables."* (pp. 157-158). Theory offers an easy solution to this challenge by implying that the gravity model can be estimated with domestic and international trade flows (see 'Recommendation 5').

Capitalizing on the theory-consistent use of domestic trade flows, Heid et al. (2021) demonstrate that, even in the presence of the full set of exporter-time and importer-time fixed effects, the gravity model can deliver estimates of the effects of non-discriminatory trade policies on the exporter and on the importer side. Identification comes from the fact that such policies do not apply to domestic trade. Extending the methods of Heid et al. (2021) and focusing on the impact of institutional quality, Beverelli et al. (2024) show that the gravity model can also identify the effects of any country-specific variable.[52] Unlike the effects of non-discriminatory trade policies that can be estimated separately for imports vs. exports, the effects of the country-specific variables in the gravity model are identified for international trade (regardless of its direction) relative to domestic trade, e.g., stronger institutions promote international relative to domestic trade by 10%. However, following the recommendations of Beverelli et al. (2024), directional estimates (i.e., for exports vs. imports) of the effects of the country-

[52]Mechanically, identification of the effects of a country-specific variable on international trade is obtained by interacting it with a border dummy variable, which takes a value of one for international trade and it is zero for domestic trade. As a result, the interaction is bilateral and, therefore, can be identified even with exporter and importer fixed effects.

specific variables (e.g., SPS, corporate income taxes, etc.) can also be obtained for trade between groups of countries, e.g., rich to rich vs. rich to poor vs. poor to rich vs. poor to poor countries.

Using domestic trade flows can also help correct two mistakes we have seen in some gravity applications. First, some authors have attempted to estimate the effects of exchange rates in gravity specifications with exporter-time and importer-time fixed effects that only include international trade flows. This is impossible because, even though the exchange rate seems bilateral by definition, e.g., U.S. Dollar vs. Japanese Yen, it is actually *a country-specific variable*. To see this, note that, *ceteris paribus*, when the Dollar appreciates against the Yen, it also appreciates against the Euro. As a result, the exchange rates are fully absorbed by the country-time fixed effects in the gravity model and their effects cannot be identified when only international trade data are used. However, as demonstrated by Anderson et al. (2016), this is not the case when the model includes domestic trade flows. In this case, one can obtain estimates of the effects of exchange rates on international (relative to domestic) trade even in the presence of the full set of exporter-time and importer-time fixed effects. The source of identification is the fact that the exchange rate does not change domestically.

Another mistake we have seen often in papers that try to identify the effects of country-specific variables in the gravity models on international trade only is when the country-specific variables are combined into a bilateral covariate. The danger is that, even though the new variable is bilateral by construction, it may still be perfectly collinear with the country-time fixed effects in the gravity model. For example, one cannot identify the effects of a bilateral variable constructed as the sum of two country-specific variables. We offer several solutions to this problem. First, rely on more complicated (e.g., non-linear) combinations of the country-specific variables. This makes identification possible, however, the interpretation of the resulting estimate is difficult. Second, do not include exporter-time and importer-time fixed effects. However, such specifications would be subject to the 'gold medal mistake' for gravity estimations (Baldwin and Taglioni, 2006). Third, employ the two-stage procedure that we discussed in 'Recommendation 7'. Fourth, implement the methods of Freeman et al. (2021).[53] Fifth, if domestic trade flows are included, identification of the new bilateral covariate (or, more precisely, of its interaction with the border dummy variable) is possible even when it is constructed as

[53] Freeman et al. (2021) implement a two-step estimation procedure, where the first step gravity model delivers estimates of the multilateral resistances (MRs) from the country-time fixed effects. The second step replaces the exporter-time and/or importer-time fixed effects with the constructed MRs and output and expenditures (or further, underlying determinants, such as factor endowments), allowing the inclusion of country-specific variables. We remind the reader that not including the proper set of country fixed effects in the gravity model may lead to inconsistent and biased estimates. Thus, it is essential to pay special attention to the country variables used to replace the fixed effects.

a linear combination of the country-specific variables, and its effects are identified relative to domestic sales, which should be kept in mind when interpreting the results. However, sixth, if domestic trade flows are used, then there is no need to construct new bilateral covariates as one can include the country-specific variables directly by interacting them with the international border dummy, and the interpretation is still the average effect on international trade relative to domestic trade. Finally, the advantage of the methods of Freeman et al. (2021) applied to a model with domestic trade is that one can simultaneously identify not only the differential effect of the country-specific variable on international relative to domestic trade but also its uniform impact on sales regardless of whether they are international or domestic.

The methods that we discussed here unveil opportunities to study the effects of (m)any country-specific variables on international trade within the gravity model. However, these opportunities may come with some costs and caveats. First, one has to be very clear and careful with the interpretation of the estimates on the country-specific variables in the gravity model (or any transformations of them). Second, one has to think carefully about possible omitted variables. Third, as most bilateral policies, many country-specific gravity covariates may be endogenous. Similar to the benefits of using country-pair fixed effects for bilateral policies, the country-time fixed effects in the gravity model will mitigate endogeneity concerns for the country-specific covariates. Moreover, if the identification is based on the use of domestic sales, Beverelli et al. (2024) argue, based on insights from Nizalova and Murtazashvili (2016), that the interaction of the country-specific variable and the international border dummy leads to consistent estimates if one believes that international borders are exogenous.[54]

Recommendation 11: Cluster the standard errors.

Since PPML is a quasi-maximum likelihood estimator and the proportionality of the conditional expectation and variance will in general not hold, the PPML standard errors and various corresponding test statistics may be flawed. Therefore, one should consider adjusting the PPML standard errors accordingly. Several methods have been proposed to address this challenge. For example, Santos Silva and Tenreyro (2006) suggest using Huber-White robust standard errors, while most gravity papers cluster the standard errors by country pair. Adapting the recommendations of Cameron et al. (2011) to the three-way panel structure in a typical gravity model, Egger and Tarlea (2015) propose a multi-way clustering by importer, exporter, and time. In a series of papers, Pfaffermayr (2019, 2021, 2022) shows

[54] We refer the reader to Frankel (1997) for an IV approach to control for potential endogeneity of income/output.

that in a cross-section gravity setting two-way cluster-robust standard errors could be substantially downward biased when importer and exporter fixed effects are included. The Monte-Carlo evidence suggests that clustering may not be necessary in the presence of fixed effects. Jacknife procedures, bootstrapping, and the correction procedure proposed by Pustejovsky and Tipton (2017) could help to improve inference, whereas the last approach shows even better performance under the assumption of independent heteroskedastic disturbances. Based on the existing evidence, and since there is no analytical guidance or a uniformly agreed best approach to handle the standard errors in the gravity model, our recommendation is to experiment with alternative clusterings, e.g., (i) by country pair, (ii) by exporter, importer, and year, or (iii) by exporter-year, importer-year, and country pair.

3.3 Heterogeneous trade costs and policy effects

The effects of bilateral trade costs and trade policies in the gravity model may vary across many dimensions, e.g., by product, industry, or sector, by country, by country pair, over time, and depending on the direction of trade flows. As discussed earlier, the directional country-pair fixed effects in our preferred gravity specification (see Recommendation 8) will allow for and perfectly control for all possible observable and unobservable symmetric and asymmetric time-invariant trade costs. Moreover, we also discussed possibilities and the need to allow for heterogeneous time-invariant trade costs in cross-sectional settings, i.e., when panel data are unavailable (see Recommendations 4 and 8). Therefore, the primary focus of this subsection will be on the heterogeneity of the effects of bilateral (and time-varying) policies, and we summarize our recommendations on this front in four categories.

Recommendation 12: Obtain disaggregated policy estimates.

Per 'Recommendation 3', we believe that estimating gravity with disaggregated data has benefits. Importantly, this will allow for obtaining disaggregated trade costs and policy estimates. One reason for this recommendation is that many policies, e.g., tariffs and non-tariff measures, are imposed at a very disaggregated level. However, even if policies are implemented at the aggregate level, e.g., some RTAs or complete trade sanctions, their effects are usually very different across products, industries, and sectors.[55] Therefore, estimates of aggregate policy effects may mask significant heterogeneity. Ob-

[55] Similarly, the effects of distance can be quite heterogeneous across sectors too, e.g., due to different transportation costs, modes of transportation, etc. Using the directional country-pair fixed effects in our preferred specification will allow for and fully absorb such heterogeneous effects of distance or any other time-invariant bilateral variables. However, if it is impossible to include pair fixed effects, then allowing for a heterogeneous impact of the time-invariant gravity

taining disaggregated policy estimates would be much more informative and helpful for policy analysis and recommendations. It could also help detect possible issues with data and point to potential outliers that may be driving the aggregate estimates. Against this backdrop, and due to the separability property of the theoretical gravity model, i.e., that structural gravity holds at any level of aggregation, our practical recommendation is to obtain disaggregated gravity estimates for each product, industry, or sector rather than pooling the disaggregated data together into a single estimating sample and using interactions (for example, Larch et al. (2020) estimated the EU-Turkey Customs Union effects at the two-digit and three-digit level).

There are also certain settings when pooling the disaggregated data may be desirable. For example, obtaining sectoral instead of product-level estimates may be better for presentation purposes. Rather than aggregating the product-level data to the sectoral level, we recommend that the sectoral estimates be obtained by pooling the data for the individual products to the desired sectoral level. Pooling disaggregated data may also be beneficial for gaining estimation efficiency. For example, panel data are not available when policy changes are recent. Therefore, it may be beneficial to pool disaggregated data to gauge the impact and statistical significance of such policies. Beverelli et al. (2023) is a recent example, where disaggregated data from the USITC's ITPD-E database were pooled across industries to study the impact of the recently implemented WTO Trade Facilitation Agreement.

Importantly, gravity theory offers clear guidance on some of the fixed effects that should be employed when pooling disaggregated data together. Specifically, the dimensions of the multilateral resistances imply that they should be controlled for in pooled regressions by exporter-industry-time and the importer-industry-time fixed effects. In addition, we recommend that the pair fixed effects are allowed to vary across the disaggregated dimension as well, e.g., country-pair-industry fixed effects should be used with industry-level data. The motivation is that the bilateral trade costs may vary across the industries that have been pooled together for the estimation. Similarly, it may be desirable to also use industry-specific time-varying border indicators to allow for heterogeneous globalization trends at the industry level.

Recommendation 13: Allow for dynamic adjustments in the trade costs.

covariates may be desirable. If there is significant underlying heterogeneity at the disaggregate level, not considering this may lead to inconsistent and biased estimates and wrong policy conclusions. We refer the reader to French (2019) and Breinlich et al. (2024) for recent analysis of the interpretation, bias, and links between aggregate estimates and the corresponding disaggregated parameters.

There are good reasons to expect that the estimates of the standard gravity variables, such as distance and language, and the effects of various policy variables, such as trade agreements, sanctions, and tariffs, should evolve over time.[56] In fact, not properly accounting for dynamic effects may lead to inconsistent estimates and significant biases in the estimates of the effects of these variables and other gravity covariates (Bergstrand et al., 2015). The dynamic effects of some of the gravity covariates have been studied extensively in the literature, e.g., the impact of distance and the phasing-in effects of trade agreements, while others have attracted attention only recently, e.g., the globalization effects in gravity estimates and the evolution of the effects of tariffs from the short to the long run. Our objective is to discuss several prominent examples and some new developments focusing on the gravity model. Regardless of the application in question, to properly account for the dynamic evolution of trade costs, especially of the effects of trade policy, we reiterate our earlier Recommendation 4 to use consecutive year data.

Perhaps not surprisingly, i.e., due to its prominent role as the most widely used and most robust proxy for bilateral trade costs, one of the most studied applications related to the evolution of trade costs is on the impact of *distance* on trade over time. Disdier and Head (2008) survey the literature on the evolution of the effects of distance on trade and based on a meta-analysis of 1,467 estimates from 103 papers, conclude that *"the estimated negative impact of distance on trade rose around the middle of the century and has remained persistently high since then."*(p. 37), i.e., despite significant improvements in transportation and communication, the estimated effects of physical distance on international trade flows have remained constant over time. This mystery is known as *"The Distance Puzzle of Trade"*.

Capitalizing on gravity theory and recognizing that the structural gravity model can only identify *relative* trade costs, Yotov (2012) argues that the impact of distance on international trade should be measured relative to the change in domestic trade costs and demonstrates that, once this is done, the distance puzzle disappears and there is strong empirical evidence that the impact of distance on trade has fallen over time.[57] Several other papers have provided alternative solutions and explanations for the distance puzzle (e.g., Buch et al. (2004), Carrère and Schiff (2005), Brun et al. (2005), Boulhol

[56]Theoretical foundations for the dynamic evolution of bilateral links between trading partners are described in Arkolakis (2010), Head et al. (2010), Chaney (2014), Crucini and Davis (2016), Anderson and Yotov (2020b), Anderson and Yotov (2022), Egger et al. (2023), and Larch et al. (2024b).

[57]Mechanically, the simple (but theory-consistent) adjustment that Yotov (2012) proposes is to estimate the gravity equation with domestic, in addition to international, trade flows (see Recommendation 5).

and de Serres (2010), Lin and Sim (2012), Carrère et al. (2013), Larch et al. (2016), and Borchert and Yotov (2017)). Finally, we note that the time-varying effects of distance could be identified even in the presence of country-pair fixed effects in the gravity model.

Another application of the evolution of trade costs over time that has attracted significant attention is the phasing-in effects of RTAs. Larch and Yotov (2024b) offer a recent survey of the evolution of the methods and the estimates of the effects of RTAs. Prominently, Baier and Bergstrand (2007) use 5-year interval data to find that the effects of RTAs are exhausted about 10 to 15 years after their implementation. As discussed in Recommendation 4, a potential limitation of analyses with interval data is that they may omit some adjustments and anticipation effects, thus leading to downward biases in the RTA estimates. Using consecutive-year data Egger et al. (2022) find evidence for positive RTA effects in the years that immediately precede the RTA entry into force. This is consistent with the arguments from Moser and Rose (2014) and, as discussed in Larch and Yotov (2024b), we would interpret positive 1- or 2-year lead estimates of the effects of RTAs as an indicator of economic adjustment in response to the RTA, rather than a sign of unaccounted endogeneity. Concerning the phasing-in effects of RTAs, the existing estimates suggest strong initial effects in the few years after entry into force, which gradually decrease over time and are fully exhausted about 10 to 15 years after the agreements have entered into force. Finally, in practical terms and following Egger et al. (2022), our recommendation is to use consecutive-year data but to consider constraining the policy estimates to be common biannually for every three years.

Bhavnani et al. (2002) define the *"Missing Globalization Puzzle"* as *"the failure of declining trade-related costs to be reflected in estimates of the standard gravity model of bilateral trade"* (p.1), and conclude that *"globalization is everywhere but in estimated gravity models"* (p.3). Bergstrand et al. (2015) demonstrate that when the estimations are performed with domestic trade flows, the gravity model is well-suited to capture globalization effects with a series of time-varying border indicators, which take a value of one for international trade and zero for domestic trade.[58] The main finding of Bergstrand et al. (2015) is that the impact of international borders has fallen steadily over time, i.e., the *"Missing Globalization Puzzle"* is solved, and two additional implications of their analysis are (i) that if the common globalization trends and effects are not properly accounted for in gravity

[58] Anderson and Yotov (2020b) build a short-run gravity model that offers a theoretical foundation for the solution of the *"The Missing Globalization Puzzle"*, and conclude that globalization was always present in the gravity model, it was just hiding in the country-specific fixed effects that are typically used in gravity estimations to account for the structural multilateral resistances.

regressions, then the effects of RTAs and other policy variables may be over-predicted, and (ii) that the time-varying border variables are enough to absorb and control for any additional evolution in the effects of distance or other time-invariant gravity covariates over time. Against this backdrop, we recommend that all gravity regressions that use domestic trade flows should also include time-varying border indicators, which may vary by country or by groups of countries (e.g., developed vs. developing nations).

Finally, and most recently, there has been a renewed interest in estimating the evolution of the effects of trade costs, tariffs in particular, and the corresponding trade elasticities from the short to the long run. In part, this literature has been motivated by the observation that "[i]nternational real business cycle (IRBC) models need low elasticities, in the range of 1 to 2, to match the quarterly fluctuations in trade balances and the terms of trade, but static applied general equilibrium models need high elasticities, between 10 and 15, to account for the growth in trade following trade liberalization." (Ruhl, 2008), which he dubs *'the international elasticity puzzle'*.[59] Using French firm-level data, Fontagné et al. (2018) revisited the puzzle to conclude that it was even more severe than originally thought. Yilmazkuday (2019b) and Anderson and Yotov (2020b) provide alternative solutions to the puzzle, while Yilmazkuday (2019a), Anderson and Yotov (2022), and Boehm et al. (2023) offer estimates of the evolution of the trade elasticity from the short to the long run.

Within the gravity framework, Anderson and Yotov (2022) propose a model, that nests the standard, long-run gravity equation (see Section 2) and traces the evolution of the trade elasticity from the short to the long-run.[60] From an econometric perspective, their recommendation for gravity estimations from the short to the long run is to use interval-pair fixed effects with different interval lengths.[61] The trade elasticities that they obtain vary between 0.5 in the short run to about 5 in the long run, thus offering a solution to the *'the international elasticity puzzle'*.[62] Another practical implication for

[59]See Backus et al. (1994), Zimmermann (1997), Heathcote and Fabrizio (2002) and Feenstra et al. (2018) for examples of low trade elasticities that vary between 1 and 2 from the IRBC literature, and also Eaton and Kortum (2002), Anderson and van Wincoop (2003), Broda et al. (2006), Broda and Weinstein (2006), Egger et al. (2012), Hillberry and Hummels (2013), Simonovska and Waugh (2014), Head and Mayer (2014), Soderbery (2015), Caliendo and Parro (2015), Feenstra et al. (2018), and Fontagné et al. (2020) for examples of high trade trade elasticities that vary between 4 and 12 from the trade literature.

[60]Specifically, to model the evolution of bilateral trade costs over time, Anderson and Yotov (2022) extend the traditional Lucas and Prescott (1971) transition formulation, which combines the roles of adjustment cost and depreciation in a single adjustment parameter, to allow this parameter to vary over time-lag intervals. A broader implication of Anderson and Yotov (2022) is that models with actual dynamic adjustments may lead to different estimating equations, which have to be carefully considered and taken into account, see Egger et al. (2023) or Larch et al. (2024b) for examples.

[61]This is consistent with recent findings from Baier and Standaert (2024), who provide evidence that the assumption that the pair fixed effects are constant over time may not hold in the long run due to changes in the economic landscape.

[62]Yilmazkuday (2019a) does not rely on a gravity framework. Instead, he uses a panel structural vector autoregressive

gravity estimations from Anderson and Yotov (2022) is that the estimates of tariffs in gravity regressions can be smaller than one. Boehm et al. (2023) also estimate a gravity-type model and recover the evolution of the trade elasticities from the estimates on tariffs with local projections. A puzzling result from Boehm et al. (2023) is that both their short- and long-run trade elasticity estimates (0.8 and 2, respectively) are significantly smaller than the elasticities from the trade literature and fall within the bounds of what Ruhl (2008) and the IRBC literature would classify as short-term elasticities (see footnote 59). Since Boehm et al. (2023) do not implement some of the established recommendations for gravity estimations, e.g., the PPML estimator, it would be interesting to determine what drives the difference between their estimates and those from the trade literature.

Recommendation 14: Allow for other types and sources of heterogeneity.

As evident from the discussion in the previous two recommendations, obtaining heterogeneous gravity estimates across products/industries/sectors and over time are two natural applications where gravity theory offers useful guidance and which have been popular in the related literature. However, there are many other possible dimensions of the heterogeneity of policy effects that have been explored in the gravity literature, e.g., (i) across groups of countries depending on their economic development, e.g., poor vs. rich (see Beverelli et al. (2024) for institutional quality, Subramanian and Wei (2007) for the effects of WTO membership, Santos Silva and Tenreyro (2010) for currency unions), (ii) across individual countries (see Felbermayr et al. (2024b) for the effects of the WTO), (iii) across types of policies (see Larch and Yotov (2024b) across types of RTAs, and Nitsch (2003) and Glick and Rose (2016) across currency unions), (iv) across country pairs within trade agreements (see Baier et al. (2019) for RTAs and Larch et al. (2020) for the EU-Turkey Customs Union), (v) depending on the direction of trade flows (see Baier et al. (2019) for RTAs, Larch et al. (2020) for the EU-Turkey Customs Unions, and Felbermayr et al. (2024a) for sanctions), (vi) depending on country-size (see Micco et al. (2003) for currency unions, and Bergstrand and Clance (2023) for an application of quantile gravity), (vii) depending on the strength of the bilateral trade relationships (see Chen and Novy, 2022), and many others.

Against this backdrop, and based on significant evidence that we have accumulated from the existing literature, we are convinced that the effects of trade costs and various trade policies are

model for the imports of a single country (the United States) to obtain estimates of the trade elasticities that vary between 1 in the short, to 5 in the medium run and 7 in the long run.

indeed very heterogeneous across multiple dimensions. Moreover, exploring such heterogeneity may generate valuable policy insights. Finally, we also see an additional potential benefit from allowing for heterogeneous policy effects from an identification perspective. To describe this possibility and its potential caveats, we follow Yalcin et al. (2024), who estimate the effects of the recent sanctions on Russia due to the war in Ukraine in a gravity model with three-way fixed effects and international trade flows only. In such a setting, it is only possible to identify the average impact of the effects of the sanctions on trade between Russia and the sanctioning states relative to Russia's trade with third/non-sanctioning countries. However, if one allows for heterogeneous, country-specific sanction effects, it is also possible to identify the effects of the sanctions on bilateral trade costs between Russia and each sender (e.g., US, the EU, etc.) and also between Russia and some (but not all) third countries (e.g., China, India, and Turkey). The caveat with this procedure is that the country-specific sanction effects can only be identified relative to at least one reference country's trade with Russia. Thus, all the country-specific estimates are relative, and it is only meaningful to interpret their ranking but not their magnitudes and statistical significance. From a practical perspective, the implication is to attempt to select the reference country or group of countries very carefully based on economic reasoning and intuition, so that the relative estimates are as close as possible to being absolute.

In sum, We recommend obtaining heterogeneous policy estimates and exploring various dimensions of heterogeneity. However, we caution the reader that allowing for heterogeneity across multiple dimensions would inevitably decrease estimation efficiency and may result in data loss, overfitting, and multicollinearity. The problem is exacerbated by using high dimensional fixed effects, which we motivated earlier. Ideally, researchers should explore heterogeneity motivated by economic theory or policy considerations. Moreover, we recommend that heterogeneity-related experiments are implemented consecutively by isolating the sources of heterogeneity one at a time. Finally, following (Pfaffermayr and Yotov, 2024, Drexel University), our general recommendation is to obtain heterogeneity estimates within the same sample, e.g., if the objective is to obtain estimates for rich vs. poor countries, this should be done by introducing interactions within the full estimating rather than by splitting the estimating sample. As discussed in Recommendation 3, an important exception is estimating gravity with disaggregated data.

Recommendation 15: Consider using heterogeneity-robust DiD methods.

As noted by Nagengast and Yotov (2024), commonly used estimation methods, including gravity settings, have come under scrutiny from econometric papers that analyze settings with staggered DiD designs, where the estimates are possibly biased in the presence of treatment effect heterogeneity, e.g., due to negative weights or the so-called 'forbidden comparisons' that (mis)use already-treated units in the control group (e.g., Borusyak and Jaravel, 2017; de Chaisemartin and D'Haultfœuille, 2020). In light of the discussion in this section and the ample evidence that the effects of various policies on trade are widely heterogeneous both across groups or over time, the issues of negative weights and 'forbidden comparisons' may be particularly pronounced, as demonstrated by Nagengast and Yotov (2024) in the case of RTAs and by Nagengast et al. (2024) in the case of EU membership.

To address these challenges, Nagengast and Yotov (2024) combine the established practices for gravity estimations with recent developments from the difference-in-differences econometrics literature,[63] and they revisit the impact of RTAs to find that the 'extended two way fixed effects' (ETWFE) are significantly larger. In terms of practical implementation, we offer three tips. First, the ETWFE estimator requires the parallel trends assumptions and the no-anticipation assumption to hold. Thus, it is important to test these assumptions before obtaining the main estimates. Second, Nagengast et al. (2024) propose a fast and flexible estimation command – **jwdid** – that can be used to obtain heterogeneity-robust policy estimates in gravity settings and aggregate them to various levels of aggregation, e.g., by cohort, by cohort and time, by country, etc. Third, following Nagengast et al. (2024), we note that while the **jwdid** command allows the inclusion of additional time-varying bilateral control variables, our recommendation is only to include such covariates if they are exogenous, and we refer the reader to de Chaisemartin and D'Haultfœuille (2023) and Callaway et al. (2024a,b) for possible approaches to address this issue in linear settings. Thus, in practical terms, we recommend that the effects of alternative policies be estimated from separate regressions, and we look forward to the development of ETWFE methods to estimate gravity equations with multiple treatments. We also expect ETWFE methods to handle continuous treatments in gravity settings would be very welcome.[64]

[63] See, for example, Borusyak and Jaravel (2017); Hull (2018); de Chaisemartin and D'Haultfœuille (2020); Callaway and Sant'Anna (2021); Goodman-Bacon (2021); Sun and Abraham (2021); Wooldridge (2021); Goldsmith-Pinkham et al. (2022); de Chaisemartin and D'Haultfœuille (2023); Borusyak et al. (forthcoming); Wooldridge (2023).

[64] An alternative approach to testing robustness to the presence of other treatments that may potentially confound the main effect of interest is to omit conspicuous episodes from the analysis (Nagengast et al., 2024).

4 A benchmark estimating gravity model

In this section, we synthesize our recommendations to specify a comprehensive estimating gravity model, and we motivate each of its elements by referencing specific recommendations. We present our estimating equation without explicit sector, industry, or product notation to simplify the exposition. However, due to the separability property of the structural gravity equation, the proposed econometric model can be applied at any desired level of aggregation, i.e., it can be estimated for any specific product, industry, or sector. Whenever appropriate, we also discuss implications for estimations with pooled disaggregated data. Against this backdrop, we specify the following econometric model:

$$\begin{aligned} X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \overrightarrow{\gamma}_{ij} + BLTRL_{ij,t}\boldsymbol{\beta}_1 + \beta_2\tau_{ij,t} + BRDR_{ij,t}\boldsymbol{\beta}_3 + EXS_{i,t}{\times}BRDR_{ij}\boldsymbol{\beta}_4\right] \times \\ \exp\left[IMP_{j,t}{\times}BRDR_{ij}\boldsymbol{\beta}_5 + CNTRY_{j,t}{\times}BRDR_{ij}\boldsymbol{\beta}_6 + \beta_7 EXRT_{ij,t}{\times}BRDR_{ij}\right] \times \varepsilon_{ij,t}, \qquad \forall\, \mathbf{i}, \mathbf{j}. \end{aligned}$$

Here, $\mathbf{X_{ij,t}}$ denotes nominal trade flows in common currency at delivered prices, which have not been constructed using statistical inference (Recommendation 2) from all possible exporters i to all possible importers j (Recommendation 1) for consecutive years t (Recommendation 4), including domestic trade (Recommendation 5). The data can be at any level of aggregation and, if available, more disaggregated data should be used (Recommendations 3 and 12).

The preferred gravity estimator is PPML, and the estimation can be performed with the family of `hdfe` commands (Recommendation 6).

$\pi_{\mathbf{i,t}}$ and $\chi_{j,t}$ are exporter-time and importer-time fixed effects (Recommendation 7). If disaggregated data are pooled for the estimations (e.g., across industries), then each of the two fixed effects should also vary across the level of disaggregation, e.g., exporter-industry-time and importer-industry-time.

$\overrightarrow{\gamma}_{\mathbf{ij}}$ are directional country-pair fixed effects (Recommendation 8). If it is impossible to include country-pair fixed effects, one should add time-invariant bilateral trade costs and model them carefully, e.g., to allow for possible heterogeneous effects (Recommendations 8, 12, and 14). To estimate gravity from the short to the long run, the country-pair fixed effects may also be allowed to vary by time interval (Recommendation 13). Finally, if gravity is estimated with pooled disaggregated data, e.g., by industry or product, then the country-pair fixed effects should also vary by industry or product, e.g., directional country-industry-pair fixed effects (Recommendation 8).

$\mathbf{BLTRL_{ij,t}}$ is a vector of time-varying bilateral covariates (Recommendation 9), whose effects may be allowed to vary at the level of aggregation of the dependent variable or higher (Recommendation 12), over time (Recommendation 13), and across countries, pairs, and depending on the direction of trade flows (Recommendation 14). The effects of the trade policies that are captured by indicator variables may also be estimated using heterogeneity-robust DiD methods when the necessary assumptions are fulfilled (Recommendation 15).

$\tau_{ij,t}$ denotes $\log(1 + tariff_{ij,t})$, where $tariff_{ij,t}$ is the ad-valorem tariff. Ideally, the tariffs should vary at the level of aggregation of the dependent variable, e.g., if the model is estimated with pooled industry data, then the tariffs should also vary by industry, however, their estimates can be common across industries or they can be by industry-specific (Recommendation 9).

$\mathbf{BRDR_{ij,t}}$ is a vector of time-varying border indicators, which can be included only when the estimating sample also includes domestic trade flows (Recommendations 5 and 9). The time-varying border dummies can be allowed to vary by country or by country group, e.g., for trade from developed to developing countries (Recommendation 14). If the model is estimated with pooled disaggregated data then the effects of the border indicators should also be allowed to vary at the disaggregated level.

$\mathbf{EXS_{i,t}}$, $\mathbf{IMP_{j,t}}$, $\mathbf{CNTRY_{j,t}}$, and $\mathbf{EXRT_{ij,t}}$ are vectors of non-discriminatory export support policies, non-discriminatory import protection policies, country-specific characteristics, and exchange rates, whose effects can be identified in the presence of domestic trade flows (Recommendation 10).

Standard errors should be clustered by (i) country-pair, (ii) by exporter, importer, and time, or (iii) by exporter-time, importer-time, and country-pair (Recommendation 11).

Finally, $\varepsilon_{ij,t}$ is a multiplicative error term in our econometric model.[65] Given the rich structure of three-way fixed effects, our preferred specification would control fully for all time-invariant bilateral determinants of trade flows (through the country-pair fixed effects) and any time-varying and/or time-invariant country-specific characteristics on the exporter and the importer side (through the exporter-time and the importer-time fixed effects, respectively). Thus, the remaining unexplained variation of the dependent variable in our model can only be due to potentially omitted time-varying bilateral variables, and our recommendations on modeling bilateral trade costs and allowing for heterogeneity

[65] As demonstrated by Santos Silva and Tenreyro (2006), the error term in the PPML gravity model can be additive or multiplicative.

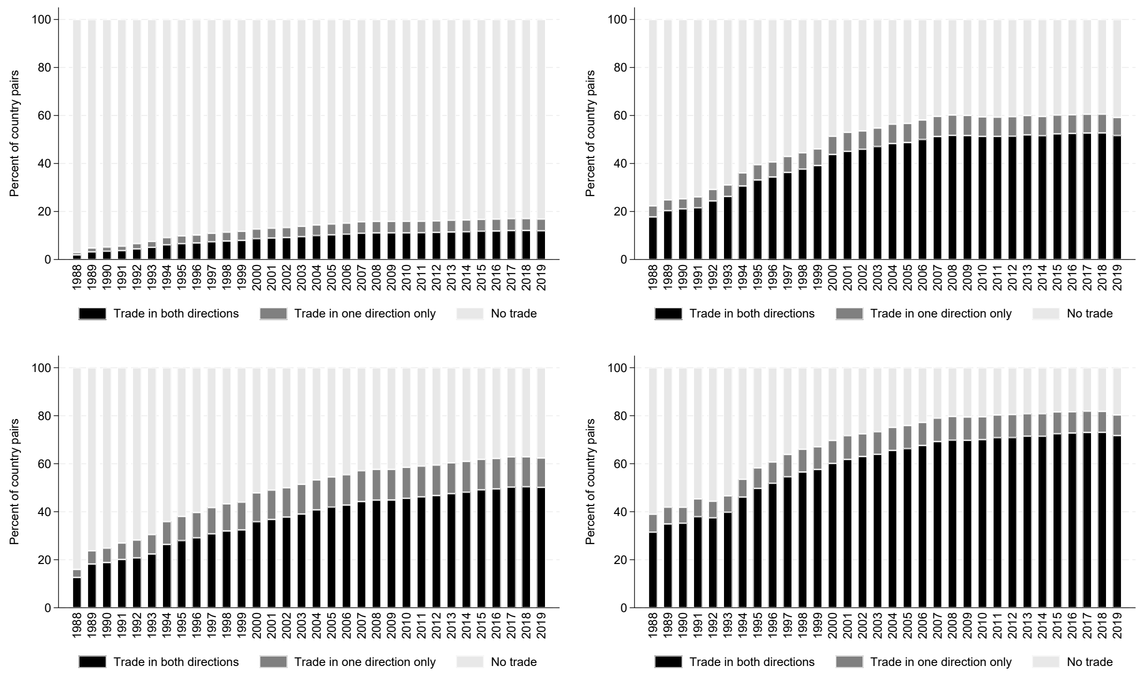
in the effects of these variables should be useful to mitigate such concerns.

Before we conclude, we note that we are keenly aware that it may not be possible to implement and apply all of our recommendations at all times, e.g., due to data limitations. However, we also believe that it is useful to have a benchmark estimating model and be aware of the benefits from being able to apply the recommendations and stay close to the proposed specification as well as being aware of the potential caveats of not being able to do so.

5 Conclusion

We traced the evolution of the trade gravity literature to synthesize the great progress that has been made into fifteen recommendations for gravity estimations, which reflect the current state of the literature and can be applied to gravity applications in trade and beyond. We also identified several areas where further developments are needed, and we look forward to seeing this new work soon.

Figure 1: The prevalence and relevance of zeros in bilateral trade flows



Note: This figure uses the *International Trade and Production Database for Estimation* (ITPD-E) to summarize the fractions of zero trade flows, trade flows in one direction, and positive bilateral trade flows in the world, 1988-2019. The figures in the top panel are based on all the data with filled zeros. The figures in the bottom panels use the same data, however, we have kept only the zeros that were not dropped when we estimated a PPML gravity model with three-way fixed effects. The figures on the left show the zeros of the disaggregated, 170 industries. The figures on the right show the zeros in a dataset where trade flows are aggregated over all industries. See the main text for further details.

References

Abadie, Alberto and Guido W. Imbens, "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, 2006, *74* (1), 235–267.

Adam, Hanna L. and Mario Larch, "A Structural Regional Trade and Migration Framework for Policy Evaluation: Quantifying the Regional Eectsof the 2004 EU Enlargement," *unpublished mnauscript*, 2024.

_ , _ , and Jordi Paniagua, "Spain, Split and Talk: Quantifying Regional Independence," *Regional Studies,* forthcoming, 2024.

_ , _ , and Michael Nower, "Closing the Gap: Explaining Cross-Country Variation in Asymmetric Trade Costs," *unpublished manuscript*, 2024.

Agnosteva, Delina E., James E. Anderson, and Yoto V. Yotov, "Intra-National Trade Costs: Assaying Regional Frictions," *European Economic Review*, 2019, *112*, 32–50.

Ai, Chunrong and Edward C. Norton, "Standard Errors for the Retransformation Problem with Heteroscedasticity," *Journal of Health Economics*, 2000, *19* (5), 697–718.

Aihounton, Ghislain B.D. and Arne Henningsen, "Units of Measurement and the Inverse Hyperbolic Sine Transformation," *The Econometrics Journal*, 2020, *24* (2), 334–351.

Allen, Treb, Costas Arkolakis, and Yuta Takahashi, "Universal Gravity," *Journal of Political Economy*, 2020, *128* (2), 393–433.

Anderson, James E., "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 1979, *69* (1), 106–116.

_ and Eric van Wincoop, "Borders, Trade, and Welfare," *Brookings Trade Forum*, 2001, pp. 207–230.

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

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

_ and Yoto V. Yotov, "The Changing Incidence of Geography," *American Economic Review*, 2010, *100* (5), 2157–2186.

_ and _ , "Terms of Trade and Global Efficiency Effects of Free Trade Agreements, 1990-2002," *Journal of International Economics*, 2016, *99* (C), 279–298.

_ and _ , "Pound for Pound Export Diversification," School of Economics Working Paper Series 2020-14, LeBow College of Business, Drexel University December 2020.

_ and _ , "Short Run Gravity," *Journal of International Economics*, 2020, *126*, September.

_ and _ , "Gravity in Transition," School of Economics Working Paper Series 2022-14, LeBow College of Business, Drexel University December 2022.

_ , Ingo Borchert, Aaditya Mattoo, and Yoto V. Yotov, "Dark Costs, Missing Data: Shedding Some Light on Services Trade," *European Economic Review*, 2018, *105* (C), 193–214.

_ , Mario Larch, and Yoto V. Yotov, "GEPPML: General Equilibrium Analysis with PPML," *The World Economy*, October 2018, *41* (10), 2750–2782.

_ , _ , and _ , "Trade and Investment in the Global Economy: A Multi-country Dynamic Analysis," *European Economic Review*, 2019, *120* (103311).

_ , _ , and _ , "Transitional Growth and Trade with Frictions: A Structural Estimation Framework," *Economic Journal*, 2020, *130* (630), 1583–1607.

_ , Mykyta Vesselovsky, and Yoto V. Yotov, "Gravity with Scale Economies," *Journal of International Economics*, 2016, *100*, 174–193.

Arkolakis, Costas, "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 2010, *118* (6), 1151–1199.

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

Arvis, Jean-Francois and Ben Shepherd, "The Poisson Quasi-Maximum Likelihood Estimator: A Solution to the "Adding up" Problem in Gravity Models," *Applied Economics Letters*, April 2013, *20* (6), 515–519.

Aytun, Ugur, Julian Hinz, Cem Ozguzel, and Joschka Wanner, "Firm-level Gravity with Fixed Effects," Technical Report, Manuscript 2024.

Backus, David K., Patrick J. Kehoe, and Finn E. Kydland, "Dynamics of the Trade Balance and the Terms of Trade: the J-curve," *American Economic Review*, 1994, *84* (1), 84–103.

Baier, Scott L. and Jeffrey H. Bergstrand, "On the Endogeneity of International Trade Flows and Free Trade Agreements," unpublished manuscript, 2002.

— and — , "The Economic Determinants of Free Trade Agreements," *Journal of International Economics*, 2004, *64* (1), 29–63.

— and — , "Do Free Trade Agreements Actually Increase Members' International Trade?," *Journal of International Economics*, 2007, *71* (1), 72–95.

— and — , "Estimating the Effects of Free Trade Agreements on International Trade Flows using Matching Econometrics," *Journal of International Economics*, 2009, *77* (1), 63–76.

— and — , "Approximating General Equilibrium Impacts of Trade Liberalizations using the Gravity Equation: Applications to NAFTA and the European Economic Area," In: Brakman, S., van Bergeijk, P. (eds.), The Gravity Equation in International Trade: Advances and Applications. Cambridge University Press, Cambridge, UK., 2010.

— and — , "NSF-Kellogg Institute Data Base on Economic Integration Agreements," *Kellogg Institute*, 2021, *https://kellogg.nd.edu/nsf-kellogg-institute-data-base-economic-integration-agreements.*

— and Samuel Standaert, "Gravity, Globalization and Time-varying Heterogeneity," *European Economic Review*, 2024, *163*, 104679.

—, Yoto V. Yotov, and Thomas Zylkin, "On the Widely Differing Effects of Free Trade Agreements: Lessons from Twenty Years of Trade Integration," *Journal of International Economics*, 2019, *116*, 206–226.

Baker, Michael and Nicole Fortin, "Occupational Gender Composition and Wages in Canada, 1987-1988," *Canadian Journal of Economics*, 2001, *34* (2), 345–376.

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

— and Jonathan I. Dingel, "Telemigration and Development: On the Offshorability of Teleworkable Jobs," NBER Working Papers 29387, National Bureau of Economic Research, Inc October 2021.

Baltagi, Badi H., *Econometric Analysis of Panel Data*, Springer International Publishing, 2021.

Barjamovic, Gojko, Thomas Chaney, Kerem Coşar, and Ali Hortaçsu, "Trade, Merchants, and the Lost Cities of the Bronze Age," *The Quarterly Journal of Economics*, 2019, *134* (3), 1455–1503.

Bellemare, Marc F. and Casey J. Wichman, "Elasticities and the Inverse Hyperbolic Sine Transformation," *Oxford Bulletin of Economics and Statistics*, 2020, *82* (1), 50–61.

Benassy-Quere, Agnes, Lionel Fontagne, and Amina Lahreche-Revil, "How Does FDI React to Corporate Taxation?," *International Tax and Public Finance*, September 2005, *12* (5), 583–603.

Bensassi, Sami and Inmaculada Martinez-Zarzoso, "How Costly is Modern Maritime Piracy to the International Community?," *Review of International Economics*, November 2012, *20* (5), 869–883.

Bergstrand, Jeffrey H., "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence," *Review of Economics and Statistics*, 1985, *67* (3), 474–481.

— , "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory of Trade," *Review of Economics and Statistics*, 1989, *71* (1), 143–53.

— **and Jordi Paniagua**, "Do Deep Trade Agreements' Provisions Actually Increase – or Decrease – Trade and/or FDI?," CESifo Working Paper Series 11526, CESifo 2024.

— **and Matthew Clance**, "Quantile Gravity: Economic Integration Agreements, Least Traded Goods, and Least Developed Economies," *Unpublished Manuscript*, 2023.

— **and Peter H. Egger**, "A Knowledge-and-Physical-Capital Model of International Trade Flows, Foreign Direct Investment, and Multinational Enterprises," *Journal of International Economics*, 2007, *73* (2), 278–308.

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

— **, Peter H. Egger, and Mario Larch**, "Gravity Redux: Structural Estimation of Elasticities and General Equilibrium Comparative Statics under Asymmetric Bilateral Trade Costs," *Journal of International Economics*, 2013, *89* (1), 110–121.

Beverelli, Cosimo, "Pull Factors for Migration: The Impact of Migrant Integration Policies," *Economics and Politics*, September 2021.

— **, Alexander Keck, Mario Larch, and Yoto Yotov**, "Institutions, Trade, and Development: Identifying the Impact of Country-specific Characteristics on International Trade," *Oxford Economic Papers*, 2024, *76* (2), 469–494.

— **, Isabella Gourevich, Inga Heiland, Alexander Keck, Mario Larch, and Yoto Yotov**, "Trade and Welfare Effects of the WTO Trade Facilitation Agreement," WTO Staff Working Papers ERSD-2023-04, World Trade Organization (WTO), Economic Research and Statistics Division 2023.

Bhavnani, Rikhil, Natalia T. Tamirisa, Arvind Subramanian, and David T. Coe, "The Missing Globalization Puzzle," IMF Working Papers 2002/171, International Monetary Fund October 2002.

Blackburn, McKinley L., "Estimating Wage Differentials Without Logarithms," *Labour Economics*, 2007, *14* (1), 73–98.

Boehm, Christoph E., Andrei A. Levchenko, and Nitya Pandalai-Nayar, "The Long and Short (Run) of Trade Elasticities," *American Economic Review*, April 2023, *113* (4), 861–905.

Borchert, Ingo and Yoto V. Yotov, "Distance, globalization, and international trade," *Economics Letters*, 2017, *153* (C), 32–38.

— **, Mario Larch, Serge Shikher, and Yoto V. Yotov**, "The International Trade and Production Database for Estimation (ITPD-E)," *International Economics*, 2021, *166*, 140–166.

— **,** — **,** — **, and** — , "The International Trade and Production Database for Estimation - Release 2 (ITPD-E-R02)," *USITC Working Paper 2022–07–A.*, 2022.

Borusyak, Kirill and Xavier Jaravel, "Revisiting Event Study Design," *Harvard University Working Paper*, 2017.

— **,** — **, and Jann Spiess**, "Revisiting Event Study Designs: Robust and Efficient Estimation," *The Review of Economic Studies*, forthcoming.

Boulhol, Hervé and Alain de Serres, "Have Developed Countries Escaped the Curse of Distance?," *Journal of Economic Geography*, January 2010, *10* (1), 113–139.

Bradley, Sebastien, Federico Carril-Caccia, and Yoto Yotov, "Reassessing the Effects of Corporate Income Taxes on Mergers and Acquisitions Using Empirical Advances in the Gravity Literature," School of Economics Working Paper Series 2023-8, LeBow College of Business, Drexel University December 2023.

Breinlich, Holger, Dennis Novy, and J. M. C. Santos Silva, "Trade, Gravity, and Aggregation," *The Review of Economics and Statistics*, September 2024, *106* (5), 1418–1426.

— **, Valentina Corradi, Nadia Rocha, Michele Ruta, João M.C. Santos Silva, and Tom Zylkin**, "Machine Learning in International Trade Research - Evaluating the Impact of Trade Agreements," CEP Discussion Papers dp1776, Centre for Economic Performance, LSE June 2021.

Broda, Christian and David E. Weinstein, "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 2006, *121* (2), 541–585.

— **, Joshua Greenfield, and David Weinstein**, "From Groundnuts to Globalization: A Structural Estimate of Trade and Growth," *NBER Working Paper No. 12512*, 2006.

Brun, Jean-Francois, Céline Carrère, Patrick Guillaumont, and Jaime de Melo, "Has Distance Died? Evidence from a Panel Gravity Model," *World Bank Economic Review*, 2005, *19* (1), 99–120.

Buch, Claudia M., Jorn Kleinert, and Farid Toubal, "The Distance Puzzle: On the Interpretation of the Distance Coefficient in Gravity Equations," *Economics Letters*, June 2004, *83* (3), 293–298.

Cairncross, Frances, *The Death of Distance*, Boston, Mass.: Harvard Business School Press, 1997. Includes bibliographical references and index.

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

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230. Themed Issue: Treatment Effect 1.

— **, Andrew Goodman-Bacon, and Pedro H. C Sant'Anna**, "Difference-in-differences with a Continuous Treatment," Working Paper 32117, National Bureau of Economic Research February 2024.

— **,** — **, and** — , "Event-Studies with a Continuous Treatment," Working Paper 32118, National Bureau of Economic Research February 2024.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics - Methods and Applications*, Cambridge, United Kingdom: Cambridge University Press, 2005.

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

Campos, Rodolfo G., Jacopo Timini, and Elena Vidal, "Structural Gravity and Trade Agreements: Does the Measurement of Domestic Trade Matter?," *Economics Letters*, 2021, *208*, 110080.

Carrère, Céline and Maurice Schiff, "On the Geography of Trade. Distance is Alive and Well," *Revue économique*, 2005, *56* (6), 1249–1274.

— **, Jaime de Melo, and John Wilson**, "The Distance Puzzle and Low-Income Countries: An Update," *Journal of Economic Surveys*, 2013, *27* (4), 717–742.

Carril-Caccia, Federico, Juliette Milgram-Baleix, and Jordi Paniagua, "Does Terrorism Affect Foreign Greenfield Investments?," *Defence and Peace Economics*, 2023, *34* (6), 827–844.

Chaney, Thomas, "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 2008, *98* (4), 1707–1721.

— , "The Network Structure of International Trade," *American Economic Review*, 2014, *104* (11), 3600–34.

— , "The Gravity Equation in International Trade: An Explanation," *Journal of Political Economy*, February 2018, *126* (1), 150–177.

Charbonneau, Karyne B., "Multiple Fixed Effects in Nonlinear Panel Data Models Theory and Evidence," *Ph.D. dissertation, Princeton University, availabel for download at https://dataspace.princeton.edu/handle/88435/dsp015q47rn86j*, 2013.

Chen, Jiafeng and Jonathan Roth, "Logs with Zeros? Some Problems and Solutions," *The Quarterly Journal of Economics*, 2023, *139* (2), 891–936.

Chen, Natalie and Dennis Novy, "Gravity and Heterogeneous Trade Cost Elasticities," *The Economic Journal*, 2022, *132* (644), 1349–1377.

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

Coe, David T., Arvind Subramanian, and Natalia T. Tamirisa, "The Missing Globalization Puzzle: Evidence of the Declining Importance of Distance," *IMF Staff Papers*, 2007, *54* (1), 34–58.

Correia, Sergio, "A Feasible Estimator for Linear Models with Multi-Way Fixed Effects," *unpublished manuscript available at http://scorreia.com/research/hdfe.pdf*, 2016.

—, Paulo Guimarães, and Thomas Zylkin, "Fast Poisson Estimation with High-Dimensional Fixed Effects," *The Stata Journal*, 2020, *20* (1), 95–115.

Costinot, Arnaud and Andrés Rodriguez-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. Rogoff, Elsevier Ltd., Oxford, 2014.

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

—, —, and —, "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 2012, *79* (2), 581–608.

Crucini, Mario J. and J. Scott Davis, "Distribution Capital and the Short- and Long-run Import Demand Elasticity," *Journal of International Economics*, 2016, *100* (C), 203–219.

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

Davies, Richard B. and Clifford M. Guy, "The Statistical Modeling of Flow Data When the Poisson Assumption Is Violated," *Geographical Analysis*, 1987, *19* (4), 300–314.

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

De Benedictis, Luca and Daria Taglioni, "The Gravity Model in International Trade," *chapter 4 in Luca De Benedictis and Luca Salvatici (Ed.), The Trade Impact of European Union Preferencial Policies*, 2011, *Springer.*

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, September 2020, *110* (9), 2964–2996.

— and —, "Two-way Fixed Effects and Differences-in-Differences Estimators with Several Treatments," *Journal of Econometrics*, 2023, *236* (2), 105480.

de Sousa, José, "The Currency Union Effect on Trade is Decreasing Over Time," *Economics Letters*, 2012, *117* (3), 917–920.

Dingel, Jonathan I. and Felix Tintelnot, "Spatial Economics for Granular Settings," *unpublished manuscript*, September 2023.

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

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, apr 2018, *108* (4-5), 899–934.

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

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

—, —, and Sebastian Sotelo, "International Trade: Linking Micro and Macro," *in: Acemoglu, D., M. Arellano an E. Dekel (eds.), "Advances in Economics and Econometrics - Tenth World Congress", Volume II, Applied Economics, Cambridge: Cambridge University Press*, 2013, pp. 329–370.

—, —, B. Neiman, and J. Romalis, "Trade and the Global Recession," *American Economic Review*, 2016, *106* (11), 3401–38.

— , **Samuel S. Kortum, and Sebastian Sotelo**, "International Trade: Linking Micro and Macro," NBER Working Papers 17864, National Bureau of Economic Research, Inc February 2012.

Egger, Peter H., "A Note on the Proper Econometric Specification of the Gravity Equation," *Economics Letters*, January 2000, *66* (1), 25–31.

— , "On the Impact of Transportation Costs on Trade in a Multilateral World," *Southern Economic Journal*, 2005, *71* (3), 592–606.

— **and Filip Tarlea**, "Multi-way Clustering Estimation of Standard Errors in Gravity Models," *Economics Letters*, 2015, *134* (C), 144–147.

— **and** — , "Comparing Apples to Apples: Estimating Consistent Partial Effects of Preferential Economic Integration Agreements," *Economica*, April 2021, *88* (350), 456–473.

— **and Kevin Staub**, "GLM Estimation of Trade Gravity Models with Fixed Effects," *Empirical Economics*, 2016, *50* (1), 137–175.

— **and Mario Larch**, "Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis," *Journal of International Economics*, 2008, *76* (2), 384–399.

— **and** — , "An Assessment of the Europe Agreements' Effects on Bilateral Trade, GDP, and Welfare," *European Economic Review*, 2011, *55* (2), 263–279.

— **and** — , "Estimating Consistent Border Effects in Gravity Models with Multilateral Resistance," *The World Economy*, 2012, *35* (9), 1121–1125.

— **and** — , "Tariff Evasion Effects in Quantitative General Equilibrium," *Economics Letters*, 2012, *116* (2), 262–264.

— **and** — , "Time Zone Differences as Trade Barriers," *Economics Letters*, 2013, *119* (2), 172–175.

— **and Michael Pfaffermayr**, "The Proper Panel Econometric Specification of the Gravity Equation: A Three-way Model with Bilateral Interaction Effects," *Empirical Economics*, July 2003, *28* (3), 571–580.

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

— , **Mario Larch, and Kevin E. Staub**, "Trade Preferences and Bilateral Trade in Goods and Services: A Structural Approach," *CEPR Working Paper No. 9051*, 2012.

— , — , **and Yoto V. Yotov**, "Gravity Estimations with Interval Data: Revisiting the Impact of Free Trade Agreements," *Economica*, January 2022, *89* (353), 44–61.

— , — , **Kevin E. Staub, and Rainer Winkelmann**, "The Trade Effects of Endogenous Preferential Trade Agreements," *American Economic Journal: Economic Policy*, 2011, *3* (3), 113–143.

— , — , **Sergey Nigai, and Yoto V. Yotov**, "Trade Costs in the Global Economy: Measurement, Aggregation and Decomposition," WTO Staff Working Papers ERSD-2021-2, World Trade Organization (WTO), Economic Research and Statistics Division 2021.

— , **Reto Foellmi, Ulrich Schetter, and David Torun**, "Gravity with History: On Incumbency Effects in International Trade," *CEPR Discussion Paper No. 18421*, 2023.

Fajgelbaum, Pablo D. and Amit K. Khandelwal, "Measuring the Unequal Gains from Trade," *The Quarterly Journal of Economics*, 2016, *131* (3), 1113–1180.

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

Feenstra, Robert C, Philip Luck, Maurice Obstfeld, and Katheryn N Russ, "In Search of the Armington Elasticity," *Review of Economics and Statistics*, 2018, *100* (1), 135–150.

Felbermayr, Gabriel, Aleksandra Kirilakha, Constantinos Syropoulos, Erdal Yalcin, and Yoto V. Yotov, "The Global Sanctions Data Base," *European Economic Review*, 2020, *129* (103561).

— **and Marina Steininger**, "Revisiting the Euro's Trade Cost and Welfare Effects," *Jahrbücher für Nationalökonomie und Statistik*, 2019, *239* (5–6), 917–956.

— **and Wilhelm Kohler**, "Exploring the Intensive and Extensive Margins of World Trade," *Review of World Economics*, 2006, *142* (4), 642–674.

— **and Yoto V. Yotov**, "From Theory to Policy with Gravitas: A Solution to the Mystery of the Excess Trade Balances," *European Economic Review*, 2021, *139* (C).

— **, Benedikt Heid, Mario Larch, and Erdal Yalcin**, "Macroeconomic Potentials of Transatlantic Free Trade: A High Resolution Perspective for Europe and the World," *Economic Policy*, 2015, *83* (3), 491–537.

— **, Constantinos Syropoulos, Erdal Yalcin, and Yoto V. Yotov**, "On the Heterogeneous Effects of Sanctions on Trade," *Canadian Journal of Economics*, 2024, *Forthcoming.*

— **, Jasmin Gröschl, and Thomas Steinwachs**, "The Trade Effects of Border Controls: Evidence from the European Schengen Agreement," *Journal of Common Market Studies*, March 2018, *56* (2), 335–351.

— **, Mario Larch, Erdal Yalcin, and Yoto Yotov**, "On the Heterogeneous Trade and Welfare Effects of GATT/WTO Membership," *Review of World Economics*, 2024, *Forthcoming, https://doi.org/10.1007/s10290-023-00520-6.*

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

Flowerdew, Robin, "Fitting the Lognormal Gravity Model to Heteroscedastic Data," *Geographical Analysis*, 1982, *14* (3), 263–267.

— **and Murray Aitkin**, "A Method of Fitting the Gravity Model Based on the Poisson Distribution," *Journal of Regional Science*, 1982, *22* (2), 191–202.

Flückiger, Matthias, Erik Hornung, Mario Larch, Markus Ludwig, and Allard Mees, "Roman Transport Network Connectivity and Economic Integration," *The Review of Economic Studies*, 2021, *89* (2), 774–810.

Fontagné, Lionel Gérard, Houssein Guimbard, and Gianluca Orefice, "Product-Level Trade Elasticities: Worth Weighting For," CESifo Working Paper Series 8491, CESifo 2020.

Fontagné, Lionel, Friedrich von Kirchbach, and Mondher Mimouni, "An Assessment of Environmentally-related Non-tariff Measures," *The World Economy*, 2005, *28* (10), 1417–1439.

— **, Philippe Martin, and Gianluca Orefice**, "The International Elasticity Puzzle is Worse Than You Think," *Journal of International Economics*, 2018, *115* (C), 115–129.

Fotheringham, A. Stewart and Peter A. Williams, "Further Discussion on the Poisson Interaction Model," *Geographical Analysis*, 1983, *15* (4), 343–347.

Frankel, J.A., *Regional Trading Blocs*, Washington, DC: Institute for International Economics, 1997.

Freeman, Rebecca, Mario Larch, Angelos Theodorakopoulos, and Yoto Yotov, "Unlocking New Methods to Estimate Country-specific Trade Costs and Trade Elasticities," School of Economics Working Paper Series 2021-17, LeBow College of Business, Drexel University November 2021.

French, Scott, "The composition of trade flows and the aggregate effects of trade barriers," *Journal of International Economics*, 2016, *98* (C), 114–137.

— , "omparative Advantage and Biased Gravity," Technical Report, UNSW Working Paper Version 2019.

Friedman, Thomas L., *The World is Flat*, updated and expanded ed., London [u.a.]: Penguin Books [u.a.], 2006.

Gil-Pareja, Salvador, Rafael Llorca-Vivero, and Jordi Paniagua, "Trade Law and Trade Flows," *The World Economy*, March 2020, *43* (3), 681–704.

Glick, Reuven and Andrew K. Rose, "Currency Unions and Trade: A Post-EMU Reassessment," *European Economic Review*, 2016, *87* (C), 78–91.

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár, "Contamination Bias in Linear Regressions," Working Paper 30108, National Bureau of Economic Research June 2022.

Goodman-Bacon, Andrew, "Difference-in-differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277. Themed Issue: Treatment Effect 1.

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

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

_ , _ , Farid Toubal, and Yoto Yotov, "One Nation, One Language? Domestic Language Diversity, Trade and Welfare," School of Economics Working Paper Series 2021-8, LeBow College of Business, Drexel University January 2021.

Hainmueller, Jens, "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis*, 2012, *20* (1), 25–46.

Hansen, Bruce E. and Seojeong Lee, "Inference for Iterated GMM Under Misspecification," *Econometrica*, May 2021, *89* (3), 1419–1447.

Head, Keith and John Ries, "Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade," *American Economic Review*, 2001, *91* (4), 858–876.

_ and _ , "Do Trade Missions Increase Trade?," *Canadian Journal of Economics*, August 2010, *43* (3), 754–775.

_ 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. Rogoff, Elsevier Ltd., Oxford, 2014.

_ , _ , and John Ries, "The Erosion of Colonial Trade Linkages after Independence," *Journal of International Economics*, 2010, *81* (1), 1–14.

_ , _ , and Mathias Thoenig, "Welfare and Trade Without Pareto," *American Economic Review Papers and Proceedings,* forthcoming, 2014.

Heathcote, Jonathan H. and Perri. Fabrizio, "Financial Autarky and International Business Cycles," *Journal of Monetary Economics*, 2002, *49* (3), 601–627.

Heckman, James, "Sample Selection Bias as a Specification Error," *Econometrica*, 1979, *47*, 153–161.

Heid, Benedikt and Mario Larch, "Gravity with Unemployment," *Journal of International Economics*, 2016, *101*, 70–85.

_ , _ , and Yoto V. Yotov, "A Simple Method to Estimate the Effects of Non-discriminatory Trade Policy within Structural Gravity Models," *Canadian Journal of Economics*, 2021, *March.*

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

_ , Oleg Itskhoki, and Stephen Redding, "Wages, Unemployment and Inequality with Heterogeneous Firms and Workers," *NBER Working Paper No. 14122*, 2008.

Hillberry, Russell and David Hummels, "Trade Elasticity Parameters for a Computable General Equilibrium Model," in Peter B. Dixon and Dale Jorgenson, eds., *Handbook of Computable General Equilibrium Modeling*, Vol. 1 of *Handbook of Computable General Equilibrium Modeling*, Elsevier, 2013, pp. 1213–1269.

Hinz, Julian, Alexander Hudlet, and Joschka Wanner, "Separating the Wheat from the Chaff: Fast Estimation of GLMs with High-Dimensional Fixed Effects," *unpublished manuscript available at https://julianhinz.com/research/glmhdfe_technical_note/*, 2019.

Hsiao, Cheng, *Analysis of Panel Data* number 34. In 'Econometric Society monographs.', 2. ed., 8. print. ed., Cambridge [u.a.]: Cambridge University Press, 2003.

Hull, Peter, "Estimating Treatment Effects in Mover Designs," 2018.

Hummels, David, "Toward a Geography of Trade Costs," *GTAP Working Papers. Paper 17, http://docs.lib.purdue.edu/gtapwp/17*, 1999.

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

— **and Vincenzo Verardi**, "Instrumental-variable Estimation of Exponential-regression Models with Two-way Fixed Effects with an Application to Gravity Equations," *Journal of Applied Econometrics*, sep 2022, *37* (6), 1121–1137.

Kox, Henk L.M. and Hugo Rojas-Romagosa, "How Trade and Investment Agreements Affect Bilateral Foreign Direct Investment: Results from a Structural Gravity Model," *The World Economy*, 2020, *43* (12), 3203–3242.

Kwon, Ohyun, Constantinos Syropoulos, and Yoto Yotov, "The Extraterritorial Effects of Sanctions," School of Economics Working Paper Series 2022-3, LeBow College of Business, Drexel University January 2022.

LaBelle, Jesse, Inmaculada Martinez-Zarzoso, Ana Maria Santacreu, and Yoto Yotov, "Cross-border Patenting, Globalization, and Development," School of Economics Working Paper Series 2023-7, LeBow College of Business, Drexel University December 2023.

Lampe, Markus, Kevin Hjortshøj O'Rourke, Lorenz Reiter, and Yoto V. Yotov, "The Empire Project: Trade Policy in Interwar Canada," NBER Working Papers 31238, National Bureau of Economic Research, Inc May 2023.

Larch, Mario, Aiko F. Schmeißer, and Joschka Wanner, "A Tale of (Almost) 1001 Coefficients: The Deep and Heterogeneous Effects of the EU-Turkey Customs Union," *Journal of Common Market Studies,* forthcoming, 2020.

— **and Yoto V. Yotov**, "Deep Trade Agreements and FDI in Partial and General Equilibrium: A Structural Estimation Framework," Accepted 2022-7, The World Bank Economic Review 2024.

— **and —** , "Estimating the Effects of Trade Agreements: Lessons from 60 Years of Methods and Data," *The World Economy*, 2024, *47* (5), 1771–1799.

— **, Joschka Wanner, and Yoto V. Yotov**, "Bi- and Unilateral Trade Effects of Joining the Euro," *Economics Letters*, 2018, *171*, 230–234.

— **, — , — , and Thomas Zylkin**, "Currency Unions and Trade: A PPML Re-assessment with High-dimensional Fixed Effects," *Oxford Bulletin of Economics and Statistics*, 2018, *81* (3), 487–510.

— **, — , — , and —** , "Currency Unions and Trade: A PPML Re-assessment with High-dimensional Fixed Effects," *Oxford Bulletin of Economics and Statistics*, June 2019, *81* (3), 487–510.

— **, José-Antonio Monteiro, Roberta Piermartini, and Yoto Yotov**, "On the Effects of GATT/WTO Membership on Trade: They are Positive and Large After All," *Canadian Journal of Economics, forthccoming*, 2024.

— **, Leandro Navarro, and Dennis Novy**, "High-Frequency Gravity," *unpublished manuscript*, 2024.

— **, Pehr-Johan Norbäck, Steffen Sirries, and Dieter Urban**, "Heterogeneous Firms, Globalization and the Distance Puzzle," *World Economy*, 2016, *39* (9), 1307–1338.

— **, Shawn W. Tan, and Yoto V. Yotov**, "A Simple Method to Ex-ante Quantify the Unobservable Effects of Trade Liberalization and Trade Protection," *Journal of Comparative Economics*, 2023, *51* (4), 1200–1213.

Leamer, Edward E., "Cross-Section Estimation of the Effects of Trade Barriers," *In Empirical Methods for International Trade, edited by Robert C. Feenstra. Cambridge, Mass.: MIT Press*, 1988.

— **and James Levinsohn**, "International Trade Theory: The Evidence," Chapter 26 in the Handbook of International Economics Vol. 3, eds. Gene M. Grossman and Kenneth S. Rogoff, Elsevier Ltd., Oxford, 1995, *3*, 1339–1394.

Lee, Jong-Wha and Phillip Swagel, "Trade Barriers and Trade Flows Across Countries and Industries," *The Review of Economics and Statistics*, 1997, *79* (3), 372–382.

Lin, Faqin and Nicholas C.S. Sim, "Death of Distance and the Distance Puzzle," *Economics Letters*, 2012, *116* (2), 225–228.

Lucas, Robert E. and Edward C. Prescott, "Investment Under Uncertainty," *Econometrica*, 1971, *39* (5), 659–681.

MacDonald, Ms. Margaux, Mr. Roberto Piazza, Johannes Eugster, and Ms. Florence Jaumotte, "Are Bilateral Trade Balances Irrelevant?," IMF Working Papers 2020/210, International Monetary Fund September 2020.

Magee, C.S., "Endogenous Preferential Trade Agreements: An Empirical Analysis," *Contributions to Economic Analysis & Policy*, 2003, *2*, Articel 15.

Manning, Willard G., "The Logged Dependent Variable, Heteroscedasticity, and the Retransformation Problem," *Journal of Health Economics*, 1998, *17* (3), 283–295.

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

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

— **and** — , "Estimating the gravity model when zero trade flows are frequent and economically determined," *Applied Economics*, 2020, *52* (26), 2766–2779.

Mattoo, Aaditya, Nadia Rocha, and Michele Ruta, eds, *Handbook of Deep Trade Agreements*, https://openknowledge.worldbank.org/handle/10986/34055, Washington, DC: World Bank, 2020.

Mayer, T. and S. Zignago, "Notes on CEPII's Distances Measures: the GeoDist Database," *CEPII Working Paper No. 2011-25*, 2011.

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

McCullagh, P. and J.A. Nelder, *Generalized Linear Models*, 2 ed., London/New York: Chapman and Hall, 1989.

Melitz, Marc J. and Stephen J. Redding, "New Trade Models, New Welfare Implications," *American Economic Review*, March 2015, *105* (3), 1105–46.

Mercado, Rogelio, "Bilateral Capital Flows: Gravity, Push and Pull," *International Finance*, September 2022.

Micco, Alejandro, Ernesto H. Stein, and Guillermo Luis Ordonez, "The Currency Union Effect on Trade: Early Evidence from EMU," Research Department Publications 4339, Inter-American Development Bank, Research Department July 2003.

Moser, Christoph and Andrew K. Rose, "Who Benefits From Regional Trade Agreements? The View From the Stock Market," *European Economic Review*, 2014, *68*, 31–47.

Nagengast, Arne and Yoto V. Yotov, "Staggered Difference-in-Differences in Gravity Settings: Revisiting the Effects of Trade Agreements," Technical Report, American Economic Journal: Applied Economics, Forthcoming 2024.

— **, Fernando Rios-Avila, and Yoto Yotov**, "The European Single Market and Intra-EU Trade: An Assessment with Heterogeneity-Robust Difference-in-Differences Methods," School of Economics Working Paper Series 2024-5, LeBow College of Business, Drexel University April 2024.

Neyman, J. and E.L. Scott, "Consistent Estimation from Partially Consistent Observations," *Econometrica*, 1948, *16* (1), 1–32.

Nitsch, Volker, "Honey, I Shrunk the Currency Union Effect on Trade," *The World Economy*, 2003, *Forthcoming, https://doi.org/10.1007/s10290-023-00520-6.*

Nizalova, Olena Y. and Irina Murtazashvili, "Exogenous Treatment and Endogenous Factors: Vanishing of Omitted Variable Bias on the Interaction Term," *Journal of Econometric Methods*, 2016, *5* (1).

Novy, Dennis, "Gravity Redux: Measuring International Trade Costs With Panel Data," *Economic Inquiry*, 2013, *51* (1), 101–121.

— , "International Trade without CES: Estimating Translog Gravity," *Journal of International Economics*, 2013, *89* (2), 271–282.

Oberhofer, Harald, Michael Pfaffermayr, and Richard Sellner, "Revisiting Time as a Trade Barrier: Evidence from a Panel Structural Gravity Model," *Review of International Economics*, November 2021, *29* (5), 1382–1417.

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

Paniagua, Jordi, "A Gravity Model for Foreign Re-investment," *Economics Bulletin*, 2015, *35* (1), 627–632.

Persyn, Damiaan and Wouter Torfs, "A Gravity Equation for Commuting with an Application to Estimating Regional Border Effects in Belgium," *Journal of Economic Geography*, 03 2015, *16* (1), 155–175.

Pfaffermayr, Michael, "Gravity Models, PPML Estimation and the Bias of the Robust Standard Errors," *Applied Economics Letters*, feb 2019, *26* (18), 1467–1471.

— , "Confidence Intervals for the Trade Cost Parameters of Cross-section Gravity Models," *Economics Letters*, apr 2021, *201*, 109787.

— , "Cross-section Gravity Models, PPML Estimation and the Bias Correction of the Two-way Cluster-robust Standard Error," *unpublished manuscript*, 2022.

— **and Yoto V. Yotov**, "Gravity Estimations on Sub-samples," *Unpublished Manuscript*, 2024, Drexel University.

Poot, Jacques, Omoniyi Alimi, Michael P. Cameron, and David C. Mará, "The Gravity Model of Migration: The Successful Comeback of an Ageing Superstar in Regional Science," *INVESTIGACIONES REGIONALES - Journal of REGIONAL RESEARCH*, 2016, pp. 63–86.

Pustejovsky, James E. and Elizabeth Tipton, "Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models," *Journal of Business & Economic Statistics*, 2017, *36* (4), 672–683.

Pöyhönen, Pentti, "A Tentative Model for the Volume of Trade Between Countries," *Weltwirtschaftliches Archiv*, 1963, *90* (1).

— , *Toward a General Theory of International Trade*, Ekonomiska Samfundets tidskrift, Institute of economics, University of Helsinki, 1964.

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

Ravenstein, Ernest George, "The Laws of Migration: Part 1," *Journal of the Statistical Society of London*, 1885, *48* (2), 167–235.

— , "The Laws of Migration: Part 2," *Journal of the Royal Statistical Society*, 1889, *52* (2), 241–305.

Redding, S. and A. Venables, "Economic Geography and International Inequality," *Journal of International Economic*, 2004, *62* (1), 53–82.

Rose, Andrew K., "Do We Really Know That the WTO Increases Trade?," *American Economic Review*, 2004, *94* (1), 98–114.

— **and Eric van Wincoop**, "National Money as a Barrier to International Trade: The Real Case for Currency Union," *American Economic Review*, 2001, *91* (2), 386–390.

Ruhl, Kim J., "The International Elasticity Puzzle," unpublished manuscript, available for download at `http://www.kimjruhl.com/`, 2008.

Santos Silva, João M.C. and Rainer Winkelmann, "Misspecified Exponential Regressions: Estimation, Interpretation, and Average Marginal Effects," *Review of Economics and Statistics*, 2024, pp. 1–25.

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

— **and** — , "On the Existence of the Maximum Likelihood Estimates in Poisson Regression," *Economics Letters*, 2010, *107* (2), 310–312.

— **and** — , "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator," *Economics Letters*, 2011, *112* (2), 220–222.

— and — , "Trading Partners and Trading Volumes: Implementing the Helpman-Melitz-Rubinstein Model Empirically," *Oxford Bulletin of Economics and Statistics*, 2015, *77* (1), 93–105.

— and — , "The Log of Gravity at 15," *Portuguese Economic Journal*, 2022, *21* (3), 423–437.

Shikher, Serge, "Predicting the Effects of NAFTA: Now We Can Do It Better!," *Journal of International and Global Economic Studies*, 2012, *5* (2), 32–59.

— , "Putting Industries Into the Eaton–Kortum Model," *The Journal of International Trade & Economic Development*, 2012, *21* (6), 807–837.

Simonovska, Ina and Michael E. Waugh, "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

Soderbery, Anson, "Estimating Import Supply and Demand Elasticities: Analysis and Implications," *Journal of International Economics*, 2015, *96* (1), 1–17.

Soderlund, Bengt, "The Importance of Business Travel for Trade: Evidence from the Liberalization of the Soviet Airspace," *Journal of International Economics*, 2023, *145* (C).

Sotelo, Sebastian, "Practical Aspects of Implementing the Multinomial PML Estimator," *unpublished manuscript, available for download at http://www-personal.umich.edu/∼ssotelo/*, 2019.

Stammann, Amrei, "Fast and Feasible Estimation of Generalized Linear Models with High-Dimensional k-way Fixed Effects," *unpublished manuscript available at https://arxiv.org/abs/1707.01815*, 2017.

Stronge, William B. and Ronald R. Schultz, "Heteroscedasticity and the Gravity Model," *Geographical Analysis*, 1978, *10*, 279–86.

Subramanian, Arvind and Shang-Jin Wei, "The WTO Promotes Trade, Strongly But Unevenly," *Journal of International Economics*, 2007, *72* (1), 151–175.

Sun, Kai, Daniel J. Henderson, and Subal C. Kumbhakar, "Biases in Approximating Log Production," *Journal of Applied Econometrics*, 2010, *26* (4), 708–714.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199. Themed Issue: Treatment Effect 1.

Syropoulos, Constantinos, Gabriel Felbermayr, Aleksandra Kirilakha, Erdal Yalcin, and Yoto V. Yotov, "The Global Sanctions Data Base–Release 3: COVID-19, Russia, and Multilateral Sanctions," *Review of International Economics*, February 2024, *32* (1), 12–48.

Teti, Feodora, "30 Years of Trade Policy: Evidence from 5.7 Billion Tariffs," ifo Working Paper Series 334, ifo Institute - Leibniz Institute for Economic Research at the University of Munich 2020.

Tinbergen, Jan, *Shaping the World Economy: Suggestions for an International Economic Policy*, New York: The Twentieth Century Fund, 1962.

Trefler, Daniel, "Trade Liberalization and the Theory of Endogeneous Protection: An Econometric Study of U.S. Import Policy," *Journal of Political Economy*, 1993, *101* (1), 138–160.

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

Wang, Feicheng, Zheng Wang, and Zhou Zhou, "Business Flies: The Trade Promoting Effect of Air Connectivity," University of Göttingen Working Papers in Economics 425, University of Goettingen, Department of Economics 2021.

Wei, Shang-Jin, "Intra-National versus International Trade: How Stubborn are Nations in Global Integration?," *NBER Working Papers No 5531*, 1996.

Weidner, Martin and Thomas Zylkin, "Bias and Consistency in Three-way Gravity Models," *Journal of International Economics*, 2021, *132* (103513).

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, 2nd ed., Cambridge, Massachusetts: The MIT Press, 2010.

—, "Two-way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators," *Available at SSRN 3906345*, 2021.

—, "Simple Approaches to Nonlinear Difference-in-Differences with Panel Data," *The Econometrics Journal*, 2023.

Yalcin, Erdal, Gabriel Felbermayr, Heider Kariem, Aleksandra Kirilakha, Ohyun Kwon, Constantinos Syropoulos, and Yoto V. Yotov, "The Global Sanctions Data Base - Release 4: The Heterogeneous Effects of the Sanctions on Russia," *Unpublished Manuscript*, 2024.

Yang, Yimin and Huili Zhang, "Three-way gravity models with multiplicative unobserved effects," *The Econometrics Journal*, 2023, *26* (3), 422–443.

Yilmazkuday, Hakan, "Estimating the Trade Elasticity Over Time," *Economics Letters*, 2019, *183* (C), 1–1.

—, "Understanding the International Elasticity Puzzle," *Journal of Macroeconomics*, 2019, *59* (C), 140–153.

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

—, "On the Role of Domestic Trade Flows for Estimating the Gravity Model of Trade," *Contemporary Economic Policy*, 2022, *40* (3), 526–540.

—, Roberta Piermartini, Jose-Antonio Monteiro, and Mario Larch, *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, Geneva: UNCTAD and WTO, 2016.

Zimmermann, C., "International Real Business Cycles among Heterogeneous Countries," *European Economic Review*, 1997, *41* (2), 319–356.

Zylkin, Tom, "Bootstrap for Gravity Models," Unpublished Manuscript, University of Richmond 2024.