KIEL
WORKING
PAPER

From Theory to Policy with Gravitas: A Solution to the Mystery of the Excess Trade Balances

Gabriel J. Felbermayr and Yoto V. Yotov

ABSTRACT

FROM THEORY TO POLICY WITH GRAVITAS: A SOLUTION TO THE MYSTERY OF THE EXCESS TRADE BALANCES *

Gabriel J. Felbermayr and Yoto V. Yotov

Bilateral trade balances often play an important role in the international trade policy debate. Disturbingly, several studies argue that the gravity model of trade fails when confronted with bilateral trade balances data, dubbing this "The Mystery of the Excess Trade Balances". Capitalizing on developments in the related literature, we solve the mystery and show that the gravity model explains bilateral trade balances well. We also uncover a new property of the structural multilateral resistances as asymmetric trade costs. This is exactly what solves the mystery. Our analysis suggests that, on average, there is little room for trade cost asymmetries due to 'unfair' policies to lead to bilateral trade imbalances. However, we also identify sectors (e.g., Mining and Services), where the modeling of the direct bilateral trade costs in the gravity model can be improved, including the possibility for 'unfair'/asymmetric policies to play a more significant role.

Gabriel J. Felbermayr
Kiel Institute for the World Economy;
Kiel University

Email: felbermayr@ifw-kiel.de

Yoto V. Yotov
School of Economics, Drexel University;
Center for International Economics, ifo
Institute; CESifo; and ERI-BAS

Email: yotov@drexel.edu

"When net exports are negative, that is, when a country runs a trade deficit by importing more than it exports, this subtracts from growth."

<div align="right">(Navarro and Ross, 2016)</div>

"We use the canonical 'gravity model' of bilateral trade to form predictions about bilateral trade balances ... Our results paint a dismal picture. The central explanations that economists provide to explain bilateral imbalances fail miserably ... These failures require that we move beyond the simple gravity framework."

<div align="right">(Davis and Weinstein, 2002)</div>

1 Introduction

Trade balances play a very important role in current trade policy negotiations and popular discussions, and always have.[1] Bilateral trade balances, in particular, seem to be widely taken as an indicator of "fairness" in trade relations. Frequent tweets of former U.S. President Donald Trump very clearly point into this direction. Most economists understand that bilateral trade balances are of no relevance for whether and to what extent a country benefits from international trade. Therefore, they should not matter for the design of a welfare-maximizing trade policy. However, academic economists also recognize the important role that trade balances play in actual trade policy negotiations, and they call for more rigorous scholarly work in this area; *"Because the bilateral trade deficit has real consequences on trade policies, we should definitely solve the technical problem of measuring it accurately."* (Feenstra et al., 1999).[2] Yet, as noted by Paul Krugman in his New York Times column, *"[s]omewhat surprisingly, there's not a lot of economic literature on the causes of bilateral*

---

[1] The opening quote of our paper is from an op-ed on Trump's economic plan, written by Commerce Secretary Wilbur Ross and trade adviser Peter Navarro in September 2016. Similarly, Davis and Weinstein (2002) note that "[b]ilateral trade deficits are a perennial policy issue" (p. 170), and motivate their work on trade balances with a quote from the former Deputy Assistant U.S. Trade Representative for Japan and China, Merit Janow, that during the first George Bush administration, "High deficits coupled with the continuing allegations from U.S. business interests about the closed nature of the Japanese market were resulting in serious domestic pressures for improved access to the Japanese market" (Janow, 1994, p.55).

[2] Bilateral balances are measured with substantial error. This is particularly true for services trade and primary income. Braml and Felbermayr (forthcoming) show that even the sign of the US-EU current account balance is essentially unknown. In this paper we focus on goods trade which is more accurately measured in international statistics.

trade imbalances." (May 31, 2017). In fact, the only academic paper that Krugman cites in his article is Davis and Weinstein (2002).

Quite disturbingly, and as captured by the second opening quote, Davis and Weinstein (2002) find that the most successful empirical trade model, the gravity equation, fails to predict trade (im)balances "miserably". To reach this conclusion, Davis and Weinstein (2002) use a canonical gravity model and plot fitted against actual bilateral trade balances. Their main finding is that the gravity model predicts balances that are an order of magnitude smaller than the corresponding actual balances. In addition, Davis and Weinstein (2002) run a regression that obtains a coefficient of fitted on actual trade balances that is only 0.06, with a corresponding $R^2 = 0.07$. Based on these results, and based on similar findings at the sectoral level, Davis and Weinstein (2002) conclude that the canonical gravity model fails to explain bilateral trade balances and dub this failure 'The Mystery of the Excess Trade Balances'. More recently, the IMF's World Economic Outlook (2019) and MacDonald et al. (2020) confirm the 'mystery' of Davis and Weinstein (2002) stating that *"[t]he gravity model explains the levels of trade balances less well than it explains unidirectional trade flows ... consistent with the long-established observation of Davis and Weinstein, 2002."*

If the results of Davis and Weinstein (2002), the World Economic Outlook (2019), and MacDonald et al. (2020) hold up to scrutiny, the overall validity of the structural gravity equation as one of the most successful empirical models in (international) economics, c.f., Costinot and Rodriguez-Clare (2014) and Head and Mayer (2014), would have to be questioned. In addition, it would also cast a shadow over the many quantitative trade models that are built around the gravity equation, c.f., Arkolakis et al. (2012) and Costinot and Rodriguez-Clare (2014). Finally, if the workhorse model cannot explain bilateral imbalances, then there might be an intellectual basis for the claim that those are due to some 'unfair' (asymmetric) manipulation of trade costs by trade partners.

Motivated by this, we make the following contributions. First, based on structural gravity theory, we show that the empirical gravity model actually does a good job in predicting trade

imbalances so that there is no 'mystery of the excess trade balances'. This result reinforces the tight connection between gravity estimations and the data, it validates the use of the gravity model for counterfactual analysis with trade imbalances, and it opens avenues for further research that integrates international trade and macroeconomics. Second, from a methodological perspective, we demonstrate that the structural multilateral resistance terms (MRs) are the key vehicles that translate the aggregate into bilateral trade imbalances, and we uncover a new property of the MRs as asymmetric trade costs. This is exactly what solves the mystery. A broader implication of this result for gravity estimations is that the use of a-theoretical proxies for the multilateral resistances is dangerous. Third, from a policy perspective, we find that, on average, there is little room for direct bilateral trade cost asymmetries (e.g., due to 'unfair' policies) to explain bilateral trade imbalances. Nonetheless, our analysis also points to sectors (e.g., Mining and Services), where the modeling of direct bilateral trade costs must be improved, including the possibility for 'unfair'/asymmetric policies to play a more significant role.

To solve the balances mystery, we capitalize on four innovations in the theoretical and empirical gravity literature since Davis and Weinstein (2002) that move gravity estimations closer in line with theory. First, following the recommendations of Anderson and van Wincoop (2003) we introduce properly defined multilateral resistance terms. In addition, we employ theory-consistent country-specific size variables, i.e., gross output and expenditure. Third, consistent with gravity theory, c.f., Arkolakis et al. (2012) or Ramondo et al. (2016), and following the estimation recommendations of Yotov et al. (2016), we estimate the gravity model with domestic trade flows, in addition to the standardly used international trade flows. Finally, following Santos Silva and Tenreyro (2006; 2011), we estimate the gravity model with the Poisson Pseudo Maximum Likelihood (PPML) estimator.[3]

---

[3]Better modeling of bilateral trade costs, especially in combination with the introduction of domestic trade flows, is a fourth possible improvement to the original specification of Davis and Weinstein (2002). While we are able to quantify the potential contribution of this adjustment, we do not implement it in our main analysis due to its *ad hoc* nature. Instead, we only implement the aforementioned four adjustments and we show that these are sufficient for the gravity model to predict bilateral trade balances quite well.

To perform the main analysis we employ the dataset of Baier et al. (2016), which includes consistently constructed domestic and international trade flows for goods. To demonstrate that the main finding of Davis and Weinstein (2002) was not an artifact of their specific data set, we reproduce the results from their original specification with the data of Baier et al. (2016). The mystery is replicated with the new data: the gravity model predicts balances that are an order of magnitude smaller than the corresponding actual balances. The regression coefficient of the fitted on the actual bilateral trade imbalances is equal to 0.08 with an R^2 of 0.05. We also implement the gravity specifications from the World Economic Outlook (2019) and MacDonald et al. (2020). Just like these studies, once again we confirm that the mystery of the excess trade balances is present in the dataset of Baier et al. (2016). Thus, our main results cannot be attributed to the different data.[4]

Using the same data, we implement the four aforementioned adjustments. The mystery of the excess trade balances disappears. A graphical illustration reveals that the points capturing actual vs. predicted balances are aligned close to a 45-degree line. In addition, the regression coefficient of fitted imbalances on actual trade imbalances is equal to 0.79 (as compared to the original 0.06) and the R^2 equal to 0.86 (as compared to the original 0.07). A series of sensitivity experiments including (i) sectoral estimations, (ii) yearly estimations, (iii) the use of an alternative data set, i.e., the World Input-Output Database (WIOD), (iv) using very disaggregated sectoral data, (v) controlling for outliers, and (vi) better modeling of bilateral trade costs all confirm the robustness of our main finding: The modern incarnation of the gravity model is well suited to predict bilateral trade balances.

Two of our sensitivity experiments have interesting policy implications. First, when we model the partial bilateral trade costs as fully symmetric based on a panel specification with pair fixed effects, we obtain a regression coefficient of fitted on actual bilateral trade imbalances that is equal to 0.87 with a corresponding R^2 of 0.88. The implication is that there is relatively little room for improvement in the fit between the actual and the predicted

---

[4]We also demonstrate the robustness of our results with regard to alternative data.

balances that can be achieved by allowing for asymmetric direct bilateral trade costs. Second, the disaggregated sectoral analysis delivers fit statistics that are an order of magnitude larger as compared to the original mystery numbers, however, we do observe worse fit in some sectors, e.g., Services, Mining, and some manufacturing sectors. The poorer fit can be used to identify sectors where the modeling of the direct bilateral trade costs could be improved. This could also be an indicator for some 'unfair' (asymmetric) manipulation of trade costs by trade partners.

So what exactly solves the mystery and explains the bilateral (im)balances? While we cannot offer an analytical answer to this question, our approach is to simultaneously introduce all but one of the major improvements that we propose at a time. The analysis reveals that when all other improvements are implemented at the same time, the gravity model performs quite well when it is estimated with the OLS estimator. Thus, we cannot attribute the solution of the mystery to PPML. We also confirm that the structural gravity model predicts bilateral trade balances very well based on data for international trade flows only, i.e., without data on domestic trade flows. Thus, the introduction of domestic trade flows/costs is not necessary to solve the balances mystery either. Third, we show that, while necessary, properly accounting for the aggregate trade balances is definitely not sufficient to solve the mystery of the bilateral balances. Finally, the analysis suggests that the key improvement that resolves the balances mystery is theory-consistent treatment of multilateral resistances.

To gauge and demonstrate the importance of the MRs for solving the mystery, we implement the iterative procedure of Anderson and van Wincoop (2003). We find that the fit between the predicted and the actual trade balances improves significantly even only after one iteration with the structural MRs, and it continues to improve with every additional iteration that takes us closer to the true MR terms. Based on this, we conclude that the structural multilateral resistances are the key channel for solving the mystery. We uncover that the MRs operate as components of asymmetric trade costs. Thus, our paper complements the seminal work of Anderson and van Wincoop (2003). Furthermore, our findings

reinforce the message from Baldwin and Taglioni (2006) that not controlling for the multilateral resistances is a 'gold medal mistake', and the argument from Head and Mayer (2014) that proxying for the structural multilateral resistances with a-theoretical remoteness indexes is not enough.

Our paper is related to various strands of research. First, we already mentioned the literature on bilateral trade balances, including the empirical study of Davis and Weinstein (2002), the World Economic Outlook, IMF (2019), and MacDonald et al. (2020). Our contribution to these papers is that we demonstrate that the structural gravity model performs quite well to explain bilateral trade (im)balances. Second, our work is also related to Reyes-Heroles (2016), who studies the relationship between trade costs and trade imbalances in a general equilibrium setting, as well as Dekle et al. (2007) who show how bilateral balances and welfare would be affected across the globe by setting all aggregate trade balances to zero. The relationship to these papers is that our analysis complements and validates the calibrated counterfactual approaches of Reyes-Heroles (2016) and Dekle et al. (2007), which implicitly assume that the gravity model works well for predicting bilateral trade balances, while we demonstrate that this is indeed the case with estimation analysis.

Third, Cunat and Zymek (2018) develop and calibrate a structural gravity model to find that large asymmetric trade costs are needed to explain the empirical patterns. We complement their analysis by decomposing the asymmetries in total bilateral trade costs into their direct/partial and general equilibrium (GE) components and by demonstrating that even when the partial/direct bilateral trade costs are perfectly symmetric, the corresponding GE bilateral trade costs may still be quite asymmetric (due to the structural MRs). Thus, we believe that our findings are consistent with and complement the results from Cunat and Zymek (2018). Fourth, we also mentioned some of the most influential papers from the recent structural gravity literature, i.e., Arkolakis et al. (2012) Costinot and Rodriguez-Clare (2014) and Head and Mayer (2014). To that list, we have to add the first theoretical gravity model of Anderson (1979) as well as the two seminal papers of Eaton and Kortum (2002)

and Anderson and van Wincoop (2003) that popularized the structural gravity model in the early 2000s.

Finally, in spirit and approach, our paper is related to several studies that resolve prominent puzzles in the economics literature by capitalizing on theoretical developments. Most notably, Anderson and van Wincoop (2003) resolve the "Canadian Border Puzzle" of McCallum (1995) by introducing and properly controlling for the structural multilateral resistance terms. Yotov (2012) resolves the "Distance Puzzle of International Trade", c.f., Disdier and Head (2008), and "Missing Globalization Puzzle", c.f., Coe et al. (2002) by recognizing that the theoretical gravity system can only identify relative trade costs and that the puzzles disappear once the effects of globalization are measured relative to the changes in domestic trade costs. Most recently, Ramondo et al. (2016) stress the counterfactual positive correlation between country size and welfare implied by the standard gravity model. They show that properly accounting for domestic trade frictions eliminates this empirical issue. In relation to this literature, we believe that our most important methodological contribution is to interpret multilateral resistances as components of asymmetric trade costs and to exploit this property to solve the balances mystery.

The remainder of the paper is structured as follows. Section 2 presents and reproduces the bilateral balances mystery of Davis and Weinstein (2002). Section 3 offers a brief review of the structural gravity model, shows how the puzzle disappears when modern gravity tools are applied, and demonstrates the robustness of our findings. Section 4 analyzes the contribution of alternative factors to the success of gravity in predicting bilateral trade imbalances. Section 5 offers a discussion on how exactly the mystery was solved. Section 6 concludes.

2 The Mystery of the Excess Trade Balances

This section describes the 'The Mystery of the Excess Trade Balances' and reproduces the analysis of Davis and Weinstein (2002) to demonstrate that the mystery is present in an alternative and more recent dataset of international trade. We also confirm the mystery using the gravity specifications from the World Economic Outlook (2019) and from MacDonald et al. (2020). First, we present the econometric gravity model as it appears in Davis and Weinstein (2002):

$$
\begin{aligned}
\ln E_{cc'} \quad = \quad & \beta_0 + \beta_1 \ln(s_{c'} X_c) + \beta_2 \ln(DIST_{cc'}) + \beta_3 \ln(REMOTE_c) \\
& + \beta_4 ADJ_{cc'} + \beta_5 FTAEC_{cc'} + \epsilon_{cc'},
\end{aligned}
\tag{1}
$$

where $E_{cc'}$ denotes exports from country c to country c'. $s_{c'}$ is the world share of spending of importer c', which is constructed as the GDP of country c' plus its current account as a share of world GDP. X_c is the GDP of exporter c. $DIST_{cc'}$ is the bilateral distance between countries c and c'. $REMOTE_c$ is a remoteness index for exporter c, which is constructed as an inverse distance-weighted average of rest-of-world GDP's. $ADJ_{cc'}$ is an indicator variable for a common border between countries c and c'. Finally, dummy variable $FTAEC_{cc'}$ takes a value of one if both countries in a pair were part of NAFTA or the EC, and it is equal to zero otherwise.[5]

Davis and Weinstein (2002) estimate Equation (1) on a cross-section of data for the year 1996. Then, they take the exponential of the fitted values to calculate estimated bilateral trade balances, $\hat{E}_{cc'} - \hat{E}_{c'c}$, and they plot them against the actual bilateral trade balances $E_{cc'} - E_{c'c}$. The results appear in Figure 1 of Davis and Weinstein (2002), which is included as Panel A of Figure 1 of this paper. As noted by Davis and Weinstein (2002), Panel A of Figure 1 reveals that the gravity model predicts balances that are an order of magnitude

---

[5]The only difference between Equation (1) and the original specification from Davis and Weinstein (2002) is that in their specification (see their Equation (7) on page 172) the ADJ and the FTAEC variables appear in logs. We believe that this is a typo since ADJ and FTAEC are indicator variables and, therefore, they should enter Equation (1) in levels.

smaller than the actual imbalances.

In addition to the visual presentation of the mystery of the excess trade balances, the authors offer a series of other statistics, which we report in Panel A of Table 1. Specifically, they construct the ratio of the variance of predicted balances to actual balances to find that it is just 0.05. Another interesting result is that the gravity model performs very poorly even to predict the sign of the bilateral trade balances. It is successful only 54 percent of the time, i.e., as good as a coin flip. Finally, the authors run a regression that obtains a coefficient of fitted imbalances on actual trade imbalances that is equal to 0.06 and an R^2 value of 0.07. Based on these results and based on similar findings that are obtained at the sectoral level, Davis and Weinstein (2002) conclude that the canonical gravity model fails to explain actual bilateral trade balances. They call this failure 'The Mystery of the Excess Trade Balances' and argue that standard explanations of bilateral imbalances based on triangular trade or the distribution of aggregate balances over trade partners have little to offer, neither on the aggregate nor on the sectoral level.

Next, we reproduce the results from Equation (1) and from Figure 1 of Davis and Weinstein (2002) with an alternative dataset. This dataset is constructed by Baier et al. (2016). It covers manufacturing trade and has several features that are needed for the estimation of the gravity equation that we develop in the next section. Specifically, in addition to international trade flows, the dataset includes domestic trade flows. Moreover, it can be used to construct total output and total expenditures. For the purpose of replicating the analysis of Davis and Weinstein (2002), we only employ the international trade flows from the Baier et al. (2016) database.[6] The other variables in Equation (1) come from several sources. Data on bilateral distances and common borders stem from the CEPII distances database, data on trade agreements from Mario Larch's RTA database. Finally, data on GDP and the current account are from the World Bank's World Development Indicators.

We use a cross-section of data from Baier et al. (2016) pertaining to the year 2000 and

---

[6]Further description of the dataset, including a list of the countries covered, appears in a Supplementary Data Appendix.

follow the steps from Davis and Weinstein (2002) to obtain the results in Panel B of Figure 1, which confirms the mystery of the excess balances.[7] Also similar to the original analysis of Davis and Weinstein (2002), in Panel B of Table 1, we find that the ratio of the variance of predicted balances to actual balances is small, just 0.19. In addition, we find that the gravity model predicts the sign of the bilateral trade balances correctly only in 53 percent of the cases. We also regress fitted imbalances on actual trade imbalances and find a coefficient of 0.09 and an R^2 of 0.04. We also note that the underlying gravity regression on trade flows in levels delivers the standard strong gravity fit, $R^2 = 0.74$. In combination with the very low R^2 from the second-stage (i.e., trade balances) regression, this result suggests that while the 'canonical' (a-theoretical) gravity model predicts the correlation between trade flow levels very well, it does not necessarily perform well in explaining trade imbalances.

Next, we follow the approach of IMF (2019) and MacDonald et al. (2020) to construct two remoteness indexes; one on the importer side and one on the exporter side. The exporter remoteness index is constructed as the weighed-average of the bilateral trade costs based on the estimates that we obtain from specification (1), with expenditures used as weights. Similarly, the importer remoteness index is also calculated as a weighted-average of the bilateral trade costs based on the estimates from specification (1), but with outputs used as weights.[8] This approach improves on the main specification of Davis and Weinstein (2002) on two fronts. First, instead of just using distance, we take into account the impact of contiguity and trade agreements. Thus, the vector of bilateral trade costs used for the construction of the remoteness indices should be more accurate. Second, and potentially more important, the use of output and expenditure as the weights allows for the construction of two remoteness indexes, one on the exporter and one on the importer side.

The results from this experiment appear in Panel C of Figure 1 and in Panel C of Table

---

[7]We pick the year 2000 because results are most striking for that particular cross-section. However, using 1996 (as in Davis and Weinstein (2002)) yields very similar results. In the sensitivity analysis we obtain estimates that confirm the robustness of our results for every year between 1991 and 2006.

[8]Since we did not have access to the data and codes in IMF (2019) and MacDonald et al. (2020), we produced the results for this analysis with our data and based on the description of the corresponding specifications from these documents.

1. The figure and the corresponding statistics are very much in line with the original findings of Davis and Weinstein (2002) and with our estimates from Panel B of Figure 1 and in Panel B of Table 1. Thus, we also confirm the mystery and the conclusions from IMF (2019) and MacDonald et al. (2020). The only three studies, which we are aware of, that test the performance of the gravity model on bilateral trade imbalances are Davis and Weinstein (2002), the World Economic Outlook, IMF (2019), and MacDonald et al. (2020) and, as noted earlier, all three studies find that gravity fails. Based on these results, we conclude that the mystery of the excess trade balances is present in the data from Baier et al. (2016) and we proceed to solve this puzzle in the next section.

3   Solving the Mystery of the Excess Trade Balances

We capitalize on the developments in the theoretical and in the empirical gravity literatures to propose four adjustments to the original specification of Davis and Weinstein (2002), which resolve the mystery of the excess balances. In Section 3.1, we briefly review the structural gravity model. This is important because, as we demonstrate later, it is exactly the close adherence to gravity theory that enables us to solve the mystery and to describe the underlying mechanisms that led to the solution. Then, we introduce and motivate each of the proposed adjustments, and we combine them to obtain a new version of the empirical gravity equation, which we estimate with the same dataset that we used in the previous section to solve the mystery of the trade balances. In Section 3.2, we discuss the results from a series of sensitivity experiments that demonstrate the robustness of our main findings and offer additional policy insights.

3.1 From Theory to Policy with Gravitas

As famously demonstrated by Arkolakis et al. (2012), the following structural gravity equation is representative of a very wide class of underlying general equilibrium trade models:[9]

$$E_{cc'} = \frac{Y_c E_{c'}}{Y} \left(\frac{t_{cc'}}{\Pi_c P_{c'}} \right)^{1-\sigma} \qquad \forall c, c', \tag{2}$$

where $E_{cc'}$ is defined earlier as the exports from c to c'. Y_c is the value of output in origin c, $E_{c'}$ is the value of expenditure at destination c', and Y denotes the value of world output. $t_{cc'}$ denotes the bilateral frictions that act directly on trade flows between c and c', e.g., bilateral distance, tariffs, etc. In addition to the direct bilateral frictions, $t_{cc'}$, the total bilateral trade cost term includes the multilateral resistances of Anderson and van Wincoop (2003):

$$(P_{c'})^{1-\sigma} = \sum_c \left(\frac{t_{cc'}}{\Pi_c} \right)^{1-\sigma} \frac{Y_c}{Y}, \tag{3}$$

$$(\Pi_c)^{1-\sigma} = \sum_{c'} \left(\frac{t_{cc'}}{P_{c'}} \right)^{1-\sigma} \frac{E_{c'}}{Y}. \tag{4}$$

The multilateral resistances (MRs) are general equilibrium trade cost terms that consistently aggregate bilateral trade costs on the consumers and on the producers in each country as if they were, respectively, buying from and shipping to a single/unified world market, c.f., Anderson and Yotov (2010).[10] If trade were frictionless, i.e., if $t_{cc'} = 1$ for all c, c', then theory implies that the right-hand side of Equation (2) collapses to $(Y_c E_{c'})/Y$. Thus, one can interpret the term $\left(\frac{t_{cc'}}{\Pi_c P_{c'}} \right)^{1-\sigma}$ as a measure of *total* bilateral trade frictions that drive a wedge between realized trade flows, $E_{cc'}$, and frictionless trade, $(Y_c E_{c'})/Y$. As demonstrated in the empirical analysis, the definition and decomposition of total bilateral trade costs in the structural gravity model into their partial/direct components ($t_{cc'}$) and

---

[9]We refer the reader to Anderson (2011), Costinot and Rodriguez-Clare (2014), Head and Mayer (2014), and Yotov et al. (2016) for recent surveys of the theoretical structural gravity literature.

[10]When trade costs are symmetric, and in absence of aggregate trade imbalances, the model implies that all bilateral balances are zero. Below, we find that aggregate imbalances explain the lion's share of bilateral imbalances; trade cost asymmetries are not essential to match data and theory.

the general equilibrium term $(\Pi_c P_{c'})$ has very important implications for the solution of the trade balances mystery.

Guided by Equation (1), we propose our first theoretically-motivated adjustment to the original specification of Davis and Weinstein (2002). Specifically, we control properly for the country-specific size variables, i.e. gross output at the country of origin (Y_c) and total expenditure at the destination $(E_{c'})$. Theory implies that the coefficients on the country-specific size variables should be equal to one. Importantly, using the theory consistent size variables would allow us to capture aggregate trade imbalances. Below, we demonstrate that this is a necessary but not sufficient component for solving the mystery. We also remind the reader that the size variable in Davis and Weinstein (2002) also captured aggregate trade balances by construction. Thus, we do not view this adjustment as a major improvement to the original gravity specification.

Our second adjustment is to properly account for the structural multilateral resistances, $(\Pi_c)^{1-\sigma}$ and $(P_{c'})^{1-\sigma}$. Even though, as can be seen from system (3)-(4), one cannot decompose analytically the separate impacts of the size variables vs. the multilateral resistances, we introduce these two components separately as this enables us to emphasize the importance of the multilateral resistances for solving the balances mystery. Similar to the country-specific size variables, theory implies that the coefficients on the multilateral resistance terms should also be equal to one.

It is also important to note that, while it has become customary in the empirical gravity literature (e.g., Hummels (2001) and Feenstra (2004)) to use exporter and importer fixed effects to control for country-specific size variables and for the multilateral resistances, we do not use any fixed effects in our specifications. Instead, we rely on the structural gravity system (2)-(4) to construct theoretically consistent multilateral resistance terms and we include them directly in our gravity specifications along with the country-specific gross output and total expenditure variables.[11]

---

[11]There are at least three possible (and equivalent) approaches to construct theory consistent multilateral resistances (MRs). In this paper we employ each of these methods. In chronological order and in level of

There are four reasons for our choice to use the four country-specific variables instead of fixed effects. First, we do not want to inflate the predictive power of the model by using fixed effects. Second, we want to stay as close as possible to the original specification of Davis and Weinstein. Third, explicitly controlling for size vs. the multilateral resistances enables us to compare their relative contributions to resolving the mystery of the excess trade balances. Finally, having explicit expressions for the multilateral resistances makes it possible to measure the importance of their theory-consistent construction relative to the a-theoretical remoteness indices employed in Davis and Weinstein (2002) and the World Economic Outlook, IMF (2019).

The third adjustment that we propose to the original specification of Davis and Weinstein (2002) is to estimate the gravity model with the PPML estimator. Following Santos Silva and Tenreyro (2006; 2011), we favor PPML over OLS because, due to heteroskedasticity, the OLS estimator delivers not only biased but also inconsistent gravity estimates. Santos Silva and Tenreyro (2006; 2011) show that the PPML estimator addresses this deficiency. In addition, due to its multiplicative form, the PPML estimator takes into account the information that is contained in zero trade flows. The OLS estimator throws this potentially useful and important information away.

The fourth adjustment that we propose is also motivated by theory. Specifically, the dependent variable in gravity estimations should include not only international trade flows but *domestic* trade flows as well. First, the inclusion of domestic trade flows is consistent with structural gravity theory, as captured by Equation (1). Second, the use of domestic

---

difficulty to apply, the oldest and most computationally intensive possibility is to implement the original iterative procedure of Anderson and van Wincoop (2003). We rely on this iterative procedure in Section 5 to demonstrate the importance of the structural MRs for solving the puzzle. The second alternative is to obtain estimates of the bilateral trade costs and then solve the non-linear system (3)-(4) for the multilateral resistances. We use this method also in Section 5, in order to measure the importance of properly controlling for aggregate trade balances. Finally, the third (most recent and easiest to implement) procedure capitalizes on the additive property of the Poisson Pseudo Maximum Likelihood (PPML), c.f., Arvis and Shepherd (2013) and Fally (2015), which allows PPML to be used as a non-linear solver to recover the multilateral resistances directly from the fixed effects in a structural gravity regression. Note that the three approaches deliver absolutely identical MR results. Thus, our main findings are identical regardless of the chosen method. For computational simplicity, for most of the analysis we rely on the PPML estimator to construct the multilateral resistances that are employed in our regressions.

trade flows allows for identification of domestic frictions. As demonstrated by Ramondo et al. (2016), this removes the counterfactual prediction that larger countries should be much richer than smaller ones. Third, following Yotov (2012), the use of domestic trade flows ensures proper measurement of the evolving impact of distance and globalization in the structural gravity model. Finally, the inclusion of domestic trade flows allows for identification of the effects of country-specific determinants of trade flows, c.f., Beverelli et al. (2017), as well as non-discriminatory effects of trade policies, c.f., Heid et al. (2017) which, in turn, allow for the identification of asymmetric trade costs.

In combination, our four adjustments lead to the following estimating gravity model:

$$
\begin{aligned}
E_{cc'} \;=\; & \exp[\beta_0 + \beta_1 \ln(Y_c) + \beta_2 \ln(E_{c'}) + \beta_3 \ln\left(\Pi_c^{1-\sigma}\right) + \beta_4 \ln\left(P_{c'}^{1-\sigma}\right) + \beta_5 \ln(DIST_{cc'})] \times \\
& \exp[\beta_6 ADJ_{cc'} + \beta_7 FTA_{cc'} + \beta_8 BRDR_{cc'} + \beta_9 BRDR\_GDP_c] + \epsilon_{cc'}.
\end{aligned} \tag{5}
$$

Here, in addition to using the theoretically-motivated variables for country-size and multilateral remoteness along with the original proxies for bilateral trade costs from Davis and Weinstein (2002), we have added two new variables due to the introduction of domestic trade flows. Specifically, $BRDR_{cc'}$ is a dummy variable that takes a value of one for international trade and of zero for domestic trade, and $BRDR\_GDP_c$ is defined as the interaction between $BRDR_{cc'}$ and national GDP. The intrinsically different nature of domestic vs. international trade requires that, when gravity is estimated with domestic trade flows, one should at least introduce the border variable $BRDR_{cc'}$, e.g., to allow for home bias effects. We also introduce the interaction between this variable and GDP, as a key country-specific economic indicator. Note that the interaction term $BRDR\_GDP_c$ results in asymmetric bilateral trade costs. This is consistent with the findings of Waugh (2010) and Fieler (2011) that poorer countries face higher trade costs. To get even closer to these studies, in the robustness analysis we interact the border dummy with GDP per capita.[12] The results

---

[12] In addition, we could easily allow for differential effects of internal distance and development, c.f., Anderson et al. (2018), institutions, Beverelli et al. (2018), or domestic language, e.g., as in Gurevich et al.

are very similar. More important, in the empirical analysis, we show that gravity predicts the bilateral trade balances well (i) without the interaction term $BRDR\_GDP_c$, i.e., with perfectly symmetric trade costs, and (ii) even without domestic trade flows altogether.

In principle, we could model both international and domestic trade costs better, with more proxies and/or with country-specific fixed effects. In addition to allowing for asymmetric country-specific trade costs, there are many other improvements that we could introduce to the modeling of bilateral trade costs in Davis and Weinstein (2002). For example, following Eaton and Kortum (2002), we could split the distance variable into intervals. In addition, following Baier et al. (2016), we could allow for agreement-specific and directional/asymmetric effects of trade agreements. We could also introduce a series of additional proxies for trade costs, e.g., WTO membership, currency unions, etc. Finally, we could employ pair fixed effects to capture all time-invariant bilateral determinants of trade flows in a panel setting. In order to avoid inflating the predictive power of the model by adding more a-theoretical bilateral trade cost variables, we obtain our main results with exactly those trade cost proxies that Davis and Weinstein (2002) have used. In the robustness analysis, we experiment with alternative specifications of the trade costs in our econometric model, and discuss implications.

We apply the proposed adjustments to the same data that we employed in the previous section to obtain the results in Panel D of Figure 1. The puzzle disappears. The points capturing actual vs. predicted balances are close to a 45-degree line in Figure 1. In addition, Panel D of Table 1 reports a regression coefficient of fitted imbalances on actual trade imbalances equal to 0.79 (as compared to the original 0.09) and an R^2 value of 0.86 (as compared to 0.04). Moreover, the ratio of the variance of predicted balances to actual balances is significantly higher 0.72 (as compared to 0.19). Furthermore, the improved gravity specification predicts the sign of the bilateral trade balances correctly in 72 percent of the cases (as compared to 53 percent). Based on these results, we conclude that the

_____

(2021). However, we do not want to inflate the fit of the model by adding more proxies for domestic trade costs.

mystery of the excess trade balances is resolved once the gravity model is estimated in accordance with the latest developments from the literature. Before we investigate which of the proposed amendments to the gravity equation matter most for its empirical success, we report the results from a series of sensitivity checks.

3.2   Robustness Experiments

We start by capitalizing on the sectoral dimension of the data by Baier et al. (2016). The motivation for this analysis is twofold. First, since the structural gravity theory is separable, i.e., the gravity model can be derived for individual sectors (e.g., Anderson and van Wincoop (2004)), the four adjustments that we propose can be implemented at the sectoral level so that we can test the performance of gravity sector by sector. Second, a key point in Davis and Weinstein (2002) is that cross-country differences in the patterns of demand and supply may give rise to bilateral trade imbalances as a result of "triangular trade". For this reason, they recommend that the model should be estimated and tested at the sectoral level.

Our findings are presented in Table 2. To ease comparison, the top row of the table reports our main results. The next eight rows report estimates per sector. Finally, the last row of Table 2 reports estimates that are based on pooled sector-level bilateral trade data rather than on aggregate data (as in the first row). Two main findings stand out. First, the structural gravity model performs well for each of the manufacturing sectors in the data. We do note, however, that we observe some clear variation across sectors. For example, the fit is not as good for the Chemicals industry, while it is best for Textiles, Paper, and Wood manufacturing products. Combined with the standardly strong fit of the gravity model to predict sectoral trade flow in levels, the variation in the fit of the model to predict bilateral trade balances across sectors points to the need and opportunity to model sector-specific (possibly asymmetric) trade costs better. Second, we note that estimation based on pooled sectoral data (i.e., the last row) performs slightly better than relying on aggregate manufacturing data (i.e., the first row). These results are consistent with the intuition from

17

Davis and Weinstein (2002) and reinforce our methods and main findings.

In the analysis so far, we stayed very close to the original specification of trade costs from Davis and Weinstein (2002) with few standard gravity proxies. In our next experiment, we study the importance of the direct/partial bilateral trade costs, $t_{cc'}$, for successfully predicting the trade balances.[13] The difference between the three panels in Table 3 and Figure 2 is in the definition of the underlying bilateral trade cost vectors. To ease comparison, the results in Panel A of Table 3 and Panel A of Figure 2 reproduce our main findings, which are based on specification (5). The results in Panel B of Table 3 and Panel B of Figure 2 are also based on specification (5) but without the interaction $BRDR\_GDP_c$, i.e., with perfectly symmetric direct bilateral trade costs. While not perfect, the fit between the actual and the predicted bilateral trade balances from this experiment is quite good. The implication is that even with a very limited set of trade costs proxies and perfectly symmetric bilateral trade costs, the gravity model does a good job in predicting bilateral trade balances.

We reinforce this result in our next experiment, where we employ the following panel specification that would enable us to capture direct bilateral trade costs better:[14]

$$
\begin{aligned}
E_{cc',t} = \; & \exp[\beta_0 + \beta_1 \ln(Y_{c,t}) + \beta_2 \ln(E_{c',t}) + \beta_3 \ln\left(\Pi_{c,t}^{1-\sigma}\right) + \beta_4 \ln\left(P_{c,t'}^{1-\sigma}\right)] \times \\
& \exp[\beta_5 FTA_{cc',t} + \mathbf{BRDR_{cc',t}} + \mu_{cc'}] + \epsilon_{cc'}.
\end{aligned}
\tag{6}
$$

In addition to the time dimension, the main difference between (6) and (5) is that our panel specification uses symmetric pair fixed effects, $\mu_{cc'}$. The idea is that the pair fixed effects capture all observable and unobservable time-invariant trade costs.[15] In addition, following Bergstrand et al. (2015), we use a vector of time-varying border dummies that control for common globalization trends. Importantly, even though much more comprehensive, the

---

[13]We thank Peter Neary for very stimulating discussions on these issues.

[14]Equation (6) is also motivated by theory, e.g., see Olivero and Yotov (2012) for a dynamic gravity model with country-specific asset accumulation.

[15]The pair fixed effects absorb $\ln(DIST_{cc'})$, $ADJ_{cc'}$, and $BRDR_{cc'}$ from specification (5). In order to keep the direct bilateral trade costs perfectly symmetric, we also do not include the interaction $BRDR\_GDP_c$. Note that the $\mu_{cc'}$ term is non-directional, i.e., $\mu_{cc'} = \mu_{c'c}$.

underlying bilateral trade costs vector is once again fully symmetric. To estimate (6), we employ a panel version of the data from Baier et al. (2016), over the period 1991-2006.

For consistency with our main results, in Panel C of Table 3 and Panel C of Figure 2 we report estimates that are based on the trade cost vector obtained from the panel specification (5), but only for the year 2000. The main message from this exercise is that, even with perfectly symmetric trade costs only, the gravity model predicts the bilateral trade imbalances very well. Specifically, the $R^2 = 0.88$ from Panel C of Table 3 suggests that, on average across country pairs, there is relatively little room for improvement in the fit between the actual and the predicted trade balances that can be achieved by introducing asymmetries in direct bilateral trade costs.

The rest of the robustness experiments that we perform appear in the Supplementary Appendix. First, we confirm our results with data for every year of the manufacturing dataset of Baier et al. (2016). Overall, the estimates across years are similar with moderate variation, thus demonstrating that the structural gravity model performs well in each year of the sample. Second, we test the robustness of our results to the presence of outliers. Specifically, we drop the countries with the largest actual bilateral trade imbalances, e.g, USA, China, Germany, Canada, and Japan. The performance of the gravity model is weaker as compared to the model with all countries, but the test statistics are still an order of magnitude larger as compared to those from Davis and Weinstein (2002).[16] Third, we study the influence of asymmetric trade costs. Staring with the panel specification with symmetric pair fixed effects from Panel C of Table 3, we show that introduction country-specific asymmetries through interactions between the border dummy variable, $BRDR_{cc'}$, and GDP or GDP per capita do not improve the fit of the model, while using asymmetric pair fixed effects in the underlying gravity regression leads to almost perfect fit. See the Supplementary Appendix for further details.

We also obtain estimates with an alternative dataset. Specifically, we employ the WIOD

---

[16]Conversely, this result shows that the model fits the trade flows of those very important trading nations very well, which of course is good news.

dataset, c.f., Timmer et al. (2015), which has several advantages and some caveats. On the positive side, WIOD offers complete and detailed sectoral coverage for the countries in the data and this database includes consistently constructed domestic trade flows. Importantly, it also covers services trade. The downside of WIOD is that country coverage is limited (to 43 countries) and that the trade data have been adjusted to match the underlying IO linkages. With these caveats in mind, we use WIOD to obtain estimates for each of the 56 sectors in WIOD for the first and for the last year of the data (i.e., for 2000 and 2014). The results from the WIOD dataset, which appear in the Supplementary Appendix, confirm our main findings. We also document some differences across sectors. For example, the model performs worse for Services and Mining. Even though the statistics for these sectors are an order of magnitude better than the corresponding indexes from Davis and Weinstein (2002). A possible explanation for this result is that our three standard proxies for bilateral trade costs do not perform well in Mining and the services sectors. One implication of this analysis is that an inferior fit of the gravity model to predict bilateral trade imbalances could indicate the need to improve the vector of bilateral trade costs for given sectors.

We draw three conclusions based on the analysis in this section. First, and most important, we demonstrate that the structural gravity model can predict bilateral trade balances quite well. Second, consistent with the intuition from Davis and Weinstein (2002), we find that the fit between the actual and the trade imbalances that are obtained with disaggregated data is better as compared to the corresponding fit from analysis with aggregate data. Third, our analysis suggests that there is relatively little room left for asymmetries in the direct/bilateral trade costs to contribute to the bilateral trade imbalances, thus, implying that asymmetries in the direct bilateral trade costs are not necessary to resolve the balances puzzle.

4   Who Did It? What Solved the Mystery?

In this section we ask: What Solved the Mystery? While we cannot offer an analytical answer to this question, our approach is to simultaneously introduce all but one of the improvements that we implemented in the previous section. The goal is to draw reduced-form evidence about the relative importance of the alternative and competing factors that, in combination, explain the trade balances. All experiments share a common approach. In each case, we construct predicted trade flows from an alternative specification of the empirical gravity model. Then, we construct bilateral trade balances and regress them on the actual trade balances from our data. We report the same statistics as in Davis and Weinstein (2002) in Table 4. In addition, we visualize our main findings in Figure 3.

We start by investigating the importance of the estimator. To this end, (i) we estimate gravity including domestic trade flows; (ii) we add theory-motivated size variables (gross output and expenditure), and (iii) we use the structural multilateral resistance terms. However, instead of the PPML estimator, we use OLS.[17] Our estimates are presented in Panel A of Table 4, and they reveal that the gravity model performs quite well even when the first stage results are obtained with the OLS estimator. Panel A of Figure 3 visualizes the success of the OLS estimator in predicting bilateral trade balances. Thus, we conclude that the use of the PPML was not necessary to solve the balances mystery. Nevertheless, we remind the reader that PPML has the attractive properties of accounting for heteroskedasticity in the trade data and taking advantage of the information contained in the zero trade flows, c.f., Santos Silva and Tenreyro (2006; 2011), and being perfectly consistent with the structural gravity model, c.f., Arvis and Shepherd (2013) and Fally (2015).

Next, we investigate the importance of using domestic trade flows for explaining the bilateral trade balances. The results in Panel B of Table 4 and Panel B of Figure 3 are obtained with data on international trade flows only (i.e., without observations on domestic

---

[17]Once again, we emphasize that the OLS estimator does not use any fixed effects. Instead, we use the theory motivated size variables and the structural multilateral resistances.

trade flows) but (i) with the PPML estimator; (ii) after controlling for the theory moti-
vated size variables (gross output and expenditure); and (iii) with the structural multilateral
resistances. The results without domestic trade flows are a bit weaker but comparable to
our main findings, and they perform quite well relative to the original results from Davis
and Weinstein (2002). Thus, we conclude that domestic trade flows are also not necessary
to solve the balances mystery. Nevertheless, we remind the reader that the inclusion of
domestic trade flows in gravity estimations is theory-consistent and, in addition, it allows
identification of the effects of country-specific and non-discriminatory policies as well as to
capture asymmetric trade costs.

In the third experiment, we focus on the importance of theory-motivated size variables,
i.e. gross output and expenditure. The results from this specification are obtained (i) with
the PPML estimator, (ii) with domestic trade flows, and (iii) with the structural multilateral
resistances, however, instead of gross output and expenditure we employ the original size
variables from Davis and Weinstein (2002). The corresponding results, which we present in
Panel C of Table 4 and Panel C of Figure 3 reveal that the use of the theory-consistent size
variables instead of the original proxies for size from Davis and Weinstein (2002) does not
make a big difference. This result should not be surprising because, just like gross output and
expenditure, the original size proxies controlled for aggregate trade balances by construction.

The results from our last experiment are obtained (i) using the PPML estimator, (ii)
including domestic along with international trade flows, and (iii) employing theory-consistent
size variables. However, instead of using the structural multilateral resistances, we employ
the original remoteness index from Davis and Weinstein (2002).[18] Our findings, which appear
in Panel D of Table 4 and Panel D of Figure 3, reveal that proper control for the structural
multilateral resistances makes the biggest difference for resolving the mystery of the bilateral
trade imbalances. Thus, our findings complement the seminal work of Anderson and van
Wincoop (2003) about the importance of the structural MRs, they reinforce the main message

[18]We also experiment with the remoteness indexes from IMF (2019) and MacDonald et al. (2020) to obtain
very similar results. See the analysis in the next section.

from Baldwin and Taglioni (2006) that not controlling for the multilateral resistances is a 'gold medal mistake', and the argument from Head and Mayer (2014) that proxying for the structural multilateral resistances with a-theoretical remoteness indexes may not be enough. We add to these arguments about the importance of the MRs by showing that the structural multilateral resistances are key for solving the mystery of the bilateral trade imbalances.

5   The Multilateral Resistances as Asymmetric Trade Costs

Based on the analysis in the previous section, we argue that the two necessary and sufficient components to solve the trade balances mystery are the aggregate trade imbalances and the structural multilateral resistances. While we cannot decompose analytically the influence of the aggregate trade balances vs. the multilateral resistances (since the size variables are also employed in the construction of the MRs), in this section we offer insights about the relative importance of these two forces, and we uncover a new property of the multilateral resistances as asymmetric GE trade costs. In order to demonstrate and decompose the role of the aggregate trade balances and the structural MRs most clearly, we stay close to the original specification of Davis and Weinstein (2002) and we perform the analysis in this section using the OLS estimator, with data on international trade flows only (i.e., no domestic trade flows), and only with the symmetric trade cost proxies from specification (5). Thus, the only variables that we change in our experiments are the MRs and the aggregate trade balances.

To demonstrate the importance of the structural multilateral resistances, in our first experiment, we implement the iterative procedure of Anderson and van Wincoop (2003). Specifically, we (i) obtain estimates of the bilateral trade costs from a specification that does not control for the multilateral resistances; (ii) use the resulting trade cost vector in system (3)-(4) to construct the corresponding structural multilateral resistances; (iii) obtain new trade cost estimates from a specification that also includes the MRs from step (ii); (iv)

use the new trade cost vector from step (iii) in system (3)-(4) to construct new multilateral resistances; (v) iterate until convergence, i.e., until the vectors of bilateral trade costs and the corresponding MRs stop changing. Throughout the analysis, both in the estimations and to solve system (3)-(4), we use and treat the actual aggregate trade imbalances as exogenous.

Figure 4 illustrates our findings by plotting the fitted lines that correspond to each iteration from our procedure. For comparison, Figure 4 also includes the fitted lines that are obtained with the a-theoretical remoteness indexes of Davis and Weinstein (2002) (labeled DW) and of IMF (2019) and MacDonald et al. (2020) (labeled IMF). Three findings stand out. First, as argued before and as captured by fitted lines DW and IMF in Figure 4, the gravity model fails to predict bilateral trade imbalances with the a-theoretical remoteness indexes. Second, the fit improves significantly only after one iteration with the structural MRs. (See fitted line 'SG Iter.1', where SG denotes Structural Gravity.) Third, the fit continues to improve with every additional iteration that takes us closer to the true structural MR terms. We believe that this experiment clearly demonstrates the important role of the structural multilateral resistances for resolving the balances mystery.

Next, we study the role of the aggregate trade balances for solving the mystery. Once again, we implement an iterative procedure, but this time we change the aggregate trade imbalances. We start by noting that, with perfectly balanced aggregate trade, our specification (where the direct bilateral trade costs are also perfectly symmetric by construction) could only predict perfectly symmetric bilateral trade flows. This is exactly our departing point, i.e., initially we assume balanced aggregate trade. Then, we gradually allow for a uniform change in the aggregate trade balances that we use in the model toward the actual aggregate trade balances that we see in the data. Specifically, in the first step we assume that the aggregate trade balances in our model are 33 percent of the corresponding actual aggregate trade balances. Then, we increase the aggregate balances in the model to 50 percent of the actual balances, then to 67 percent, then we go to 100 percent.[19] At each step, we use system

---

[19]Mechanically, we define the trade imbalance multiplicatively as $\phi_c = E_c/Y_c$. Thus, in our experiment, we first set $\phi_c = 1$, i.e., $E_c = Y_c$, then we use $\phi_c = 0.33 \times E_c/Y_c$, then $\phi_c = 0.5 \times E_c/Y_c$, etc.

(3)-(4) to obtain and control for the corresponding structural multilateral resistances.

Figure 5 illustrates our findings. The clear, and probably not surprising, message from Figure 5 is that the better we match the aggregate trade imbalances, the better the structural gravity model predicts the bilateral trade imbalances. In combination with (i) our findings from the previous section that proper account for the aggregate trade imbalances is not sufficient to resolve the mystery, and (ii) the results that we presented earlier in this section about the importance of the structural MRs, this implies that, for given aggregate trade imbalances, the structural multilateral resistances are the main vehicle that translates the aggregate trade imbalances into correctly predicted bilateral trade imbalances. Thus, solving the mystery of the excess trade balances.

Our analysis reveals an important new property of the structural multilateral resistances, namely, that they induce asymmetric *total* bilateral trade costs even when the *direct/partial* bilateral trade costs are perfectly symmetric. To see this, we refer the reader back to the discussion in Section 3.1 about the distinction between partial/direct bilateral trade costs, $t_{cc'}{}^{1-\sigma}$, vs. total bilateral trade costs, $\left(\frac{t_{cc'}}{\Pi_c P_{c'}}\right)^{1-\sigma}$. Figure 6 illustrates by plotting the direct/partial and total bilateral trade costs from the specification with symmetric trade costs, which we used to obtain the results in Panel B of Table 3 and Panel B of Figure 2. Panel A of Figure 6 assures that the direct bilateral trade costs, $t_{cc'}{}^{1-\sigma}$, are indeed perfectly symmetric. More importantly, Panel B of Figure 6 shows that the corresponding total bilateral trade costs, $\left(\frac{t_{cc'}}{\Pi_c P_{c'}}\right)^{1-\sigma}$, are clearly asymmetric, which is due to the structural MRs. The policy implication of this result is that bilateral trade imbalances are indeed due to asymmetric bilateral trade costs, however, it is the GE components of the bilateral trade costs rather than the direct/partial trade costs that play the more important role. Davis and Weinstein (2002) and MacDonald et al. (2020) fail in predicting bilateral trade balances because the terms they use instead of the proper MR expressions do not generate asymmetric total trade costs.

6 Conclusion

Capitalizing on developments in the empirical gravity literature and adhering more closely to the structural gravity theory, we implement four improvements to the original gravity specification of Davis and Weinstein (2002) to resolve the mystery of the excess trade balances. We show that properly constructed multilateral resistance terms imply asymmetric total trade costs. This property is exactly what solves the mystery. An important implication of this result for gravity estimations is that the use of a-theoretical proxies for the multilateral resistances is dangerous. From a policy perspective, our analysis implies that, on average, there is little room for 'unfair' policy asymmetries to lead to bilateral trade imbalances. However, we also identify some sectors (e.g., Mining and Services), where the modeling of the direct bilateral trade costs in the gravity model should be improved, including the possibility for 'unfair'/asymmetric policies to play a more significant role. From a broader perspective, the success of the empirical gravity equation in predicting bilateral trade balances further validates the use of the gravity model for counterfactual analysis and points to potentially fruitful research that combines the structural gravity model of trade with macroeconomic frameworks.

References

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

＿ , "The Gravity Model," *Annual Review of Economics*, 2011, *3*, 133–160.

＿ **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.

Anderson, James E. and Yoto V. Yotov, "Terms of Trade and Global Efficiency Effects of Free Trade Agreements, 1990-2002," *Journal of International Economics*, 2016, *99* (C), 279–298.

＿ **, 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.

Arkolakis, Costas, Arnaud Costinot, and Andres 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.

Baier, Scott L., Yoto V. Yotov, and Thomas Zylkin, "On the widely differing effects of free trade agreements: Lessons from twenty years of trade integration," *School of Economics Working Paper Series, 2016-15*, 2016.

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

Bergstrand, Jeffrey H., Mario Larch, and Yoto V. Yotov, "Economic Integration Agreements, Border Effects, and Distance Elasticities in the Gravity Equation," *European Economic Review,* forthcoming, 2015.

Beverelli, Cosimo, Alexander Keck, Mario Larch, and Yoto V. Yoto, "Institutions, Trade and Development: A Quantitative Analysis," Manuscript 2017.

＿ **, ＿ , ＿ , and Yoto Yotov**, "Institutions, Trade and Development: A Quantitative Analysis," School of Economics Working Paper Series 2018-3, LeBow College of Business, Drexel University February 2018.

Braml, Martin and Gabriel Felbermayr, "What Do We Really Know About the Transatlantic Current Account?," *CESifo Economic Studies*, forthcoming.

Coe, D., A. Subramanian, N. Tamirisa, and R. Bhavnani, "The Missing Globalization Puzzle," *IMF Working Paper No. 171*, 2002.

Costinot, Arnaud and Andres 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.

Cunat, Alejandro and Robert Zymek, "Bilateral Trade Imbalances," Technical Report, University of Edinburgh February 2018.

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

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Unbalanced Trade," *American Economic Review: Papers and Proceedings*, 2007, *97*, 351–355.

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.

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

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

Feenstra, R., W. Hai, W.T. Woo, and S. Yao, "Discrepancies in international data: an application to China-Hong Kong entrepÃŽt trade," *American Economic Review*, 1999, *89*, 338–343.

Feenstra, Robert C., *Advanced International Trade: Theory and Evidence*, Princeton, New Jersey: Princeton University Press, 2004.

Fieler, Ana Cecília, "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 2011, *79* (4), 1069–1101.

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

Head, Keith 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.

Heid, Benedikt, Mario Larch, and Yoto V. Yoto, "Estimating the Effects of Non-discriminatory Trade Policies within Structural Gravity Models," CESifo Working Paper Series 6735, CESifo Group Munich 2017.

Hummels, David, "Toward a Geography of Trade Costs," unpublished manuscript, 2001.

IMF, "World Economic Outlook, Growth Slowdown, Precarious Recovery," *International Monetary Fund*, 2019.

Janow, M. E., "Trading with an Ally: Progress and Discontent in U.S.-Japan Trade Relations," New York: Norton pp. 53-95, The United States, Japan, and Asia: Challenges for U.S. policy. February 1994.

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

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

Navarro, P. and W. Ross, "Scoring the Trump Economic Plan: Trade, Regulatory, & Energy Policy Impacts," Technical Report, White Paper September 2016.

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

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

Reyes-Heroles, Ricardo M., "The Role of Trade Costs in the Surge of Trade Imbalances," *Manuscript*, 2016.

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

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

Timmer, M. P., E. Dietzenbacher, B. Los, R. R. Stehrer, and G. J. de Vries, "An Illustrated User Guide to the World Input?Output Database: the Case of Global Automotive Production," *Review of International Economics*, 2015, *23*, 575–605.

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

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

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

Figure 1: The Mystery of the Excess Trade Balances
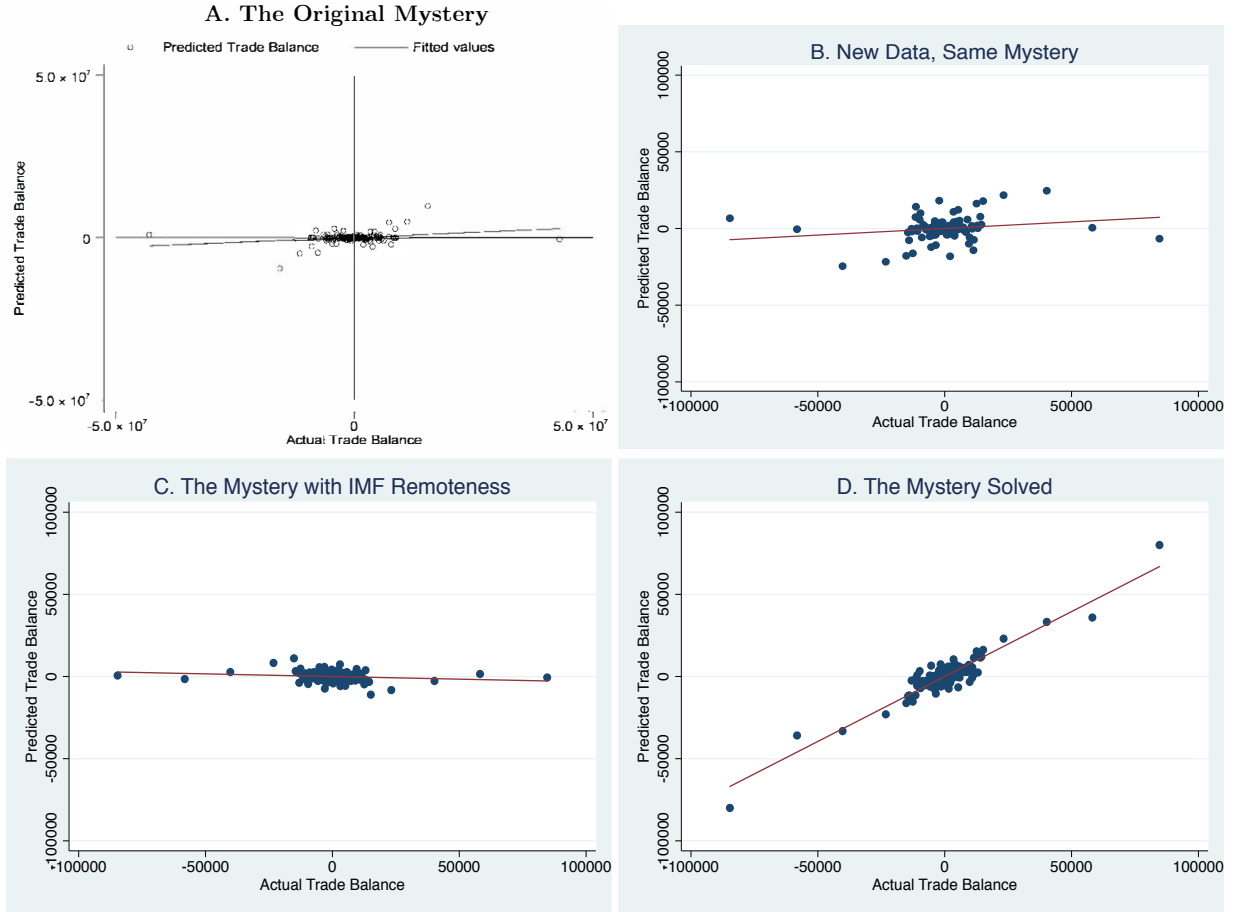
Note: Panel A of this figure is Figure 1 from Davis and Weinstein (2002), which visualizes the mystery of excess trade balances by showing that predicted balances are an order of magnitude smaller as compared to actual balances. Panel B reproduces the results from Davis and Weinstein (2002) with the data from Baier et al. (2016), and confirms the mystery of excess trade balances. To obtain the results in Panel C, we use the same data as those in Panel B, but we follow the approach of IMF (2019) and MacDonald et al. (2020) to construct the proxies for the multilateral resistances. Finally, Panel D is obtained with the same data as in the previous two panels but after introducing the adjustments to the original specification of Davis and Weinstein (2002), which we describe in the text. Alternative statistical tests for the performance of the model, which correspond to the four panels in this figure, appear in Table 1. See text for further details.

Table 1: The Mystery of Excess Trade Balances

| Regression Fit (R^2)	Regression Coefficient	Variance Ratio	Sign (%)
A. Original Specification, Original Data			
0.07	0.06	0.05	0.54
B. Original Specification, New Data			
0.04	0.09	0.19	0.53
C. New Data, IMF Remoteness Indexes			
0.02	-0.03	0.05	0.46
D. New Data, Structural Gravity Specification			
0.86	0.79	0.72	0.72

Notes: This table reports alternative statistical tests that demonstrate the performance of the gravity model to predict bilateral trade balances. The panels of the table correspond to the panels of Figure 1. Specifically, the indexes in Panel A come directly from Davis and Weinstein (2002), i.e., they are based on their original specification and are obtained with their original data. The underlying specification in Panel B is the same as in Davis and Weinstein (2002), however, the data used are from Baier et al. (2016). The results in Panel C are obtained with the same data as those in Panel B, but we follow the approach of IMF (2019) and MacDonald et al. (2020) to construct the proxies for the multilateral resistances. Finally, Panel D is obtained with the same data as in the previous two panels but after introducing the adjustments to the original specification of Davis and Weinstein (2002), which we describe in the text. See text for further details.

Table 2: Robustness: Manufacturing, Sectoral Analysis

	R^2	Reg.Coeff	Var.Ratio	Sign (%)
Total Manufacturing	0.86	0.79	0.72	0.72
Food	0.76	0.61	0.49	0.70
Textile	0.92	0.92	0.92	0.73
Wood	0.94	0.99	1.04	0.71
Paper	0.93	1.04	1.17	0.71
Chemicals	0.62	0.47	0.35	0.72
Minerals	0.78	0.79	0.80	0.71
Metals	0.77	0.75	0.74	0.62
Machinery	0.91	0.85	0.79	0.77
Aggregated Manufacturing	0.88	0.84	0.80	0.75

Notes: This table reproduces the main results from Panel D of Table 1 for each of the main manufacturing sectors covered in the dataset of Baier et al. (2016). For comparison purposes, the top row of the table reproduces the main results for total manufacturing from Panel D of Table 1. The last row aggregates the results from the individual sectors. The rest of the rows report estimates for individual sectors. See text for further details.

Figure 2: On the Importance of the Direct/Partial Bilateral Trade Costs
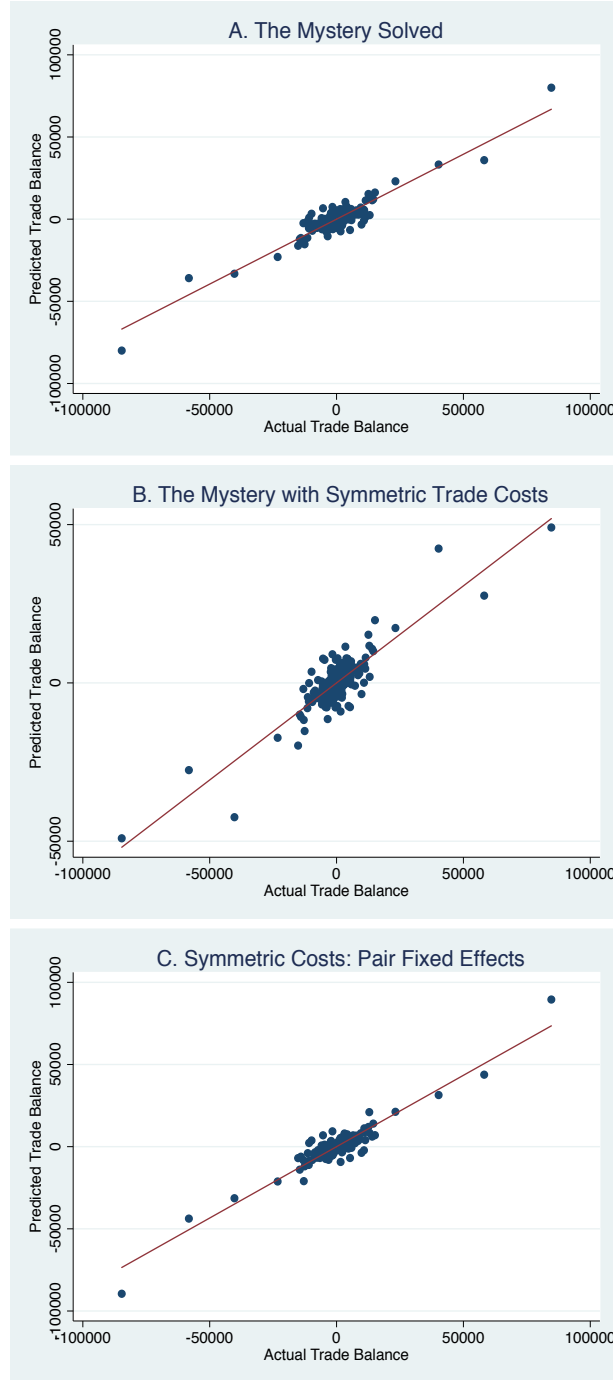
Note: This figure compares the predicted vs. actual bilateral trade balances from three underlying gravity specifications. In each case, we implement all four adjustments that we propose. The difference between the three specifications is in the definition of the bilateral trade costs that are used to predict trade flows in a structural gravity specification. For comparison purposes, the results in Panel A replicate our main findings from Panel D of Figure 1. The results in Panel B are based on symmetric trade costs that are obtained in a cross section with standard gravity variables. Finally, Panel C is based on predictions that are obtained from an underlying panel estimation, where bilateral trade costs are proxied by symmetric pair fixed effects. Alternative statistical tests for the performance of the model, which correspond to the three panels in this figure, appear in Table 3. See text for further details.

Table 3: On the Importance of the Direct/Partial Bilateral Trade Costs

| Regression Fit (R^2)	Regression Coefficient	Variance Ratio	Sign (%)
A. Main Specification			
0.86	0.79	0.72	0.72
B. Symmetric Trade Costs. Gravity Variables			
0.75	0.61	0.50	0.70
C. Symmetric Trade Costs. Symmetric Pair Fixed Effects			
0.88	0.87	0.86	0.74

Notes: This table offers statistics that compare predicted vs. actual trade balances from three underlying gravity specifications. In each case, we implement all four adjustments that we propose. The difference between the three specifications is in the definition of the bilateral trade costs that are used to predict trade flows in a structural gravity specification. For comparison purposes, the results in Panel A replicate our main findings from Panel D of Figure 1. The results in Panel B are based on symmetric trade costs that are obtained in a cross section with standard gravity variables. Finally, Panel C is based on predictions that are obtained from an underlying panel estimation, where bilateral trade costs are proxied by symmetric pair fixed effects. Figure 2 offers corresponding visualizations. See text for further details.

Table 4: Who Did It? What Solved the Mystery?

Regression Fit (R^2)	Regression Coefficient	Variance Ratio	Sign (%)
A. The Mystery and the OLS Estimator			
0.73	0.97	1.30	0.72
B. The Mystery and Domestic Trade Flows			
0.82	0.71	0.61	0.72
C. The Mystery and Theory-consistent Size			
0.78	0.59	0.45	0.67
D. The Mystery and the Multilateral Resistances			
0.29	0.28	0.26	0.67

Notes: This table reports statistics for the relative importance of the alternative improvements that we implement to resolve the mystery of the trade balances. The structure and panels in this table correspond to the panels of Figure 3. Specifically, Panel A implements all improvements except the PPML estimator. Instead, the results in this panel are obtained with the OLS estimator. Panel B implements all improvements except the addition of domestic trade flows/costs. Panel C implements all improvements except for the theoretically motivated size controls. Finally, the results in Panel D are obtained with all other adjustments but with the a-theoretical remoteness indexes instead of the structural multilateral resistances. See text for further details.

Figure 3: Who Did It? What Solved the Mystery?
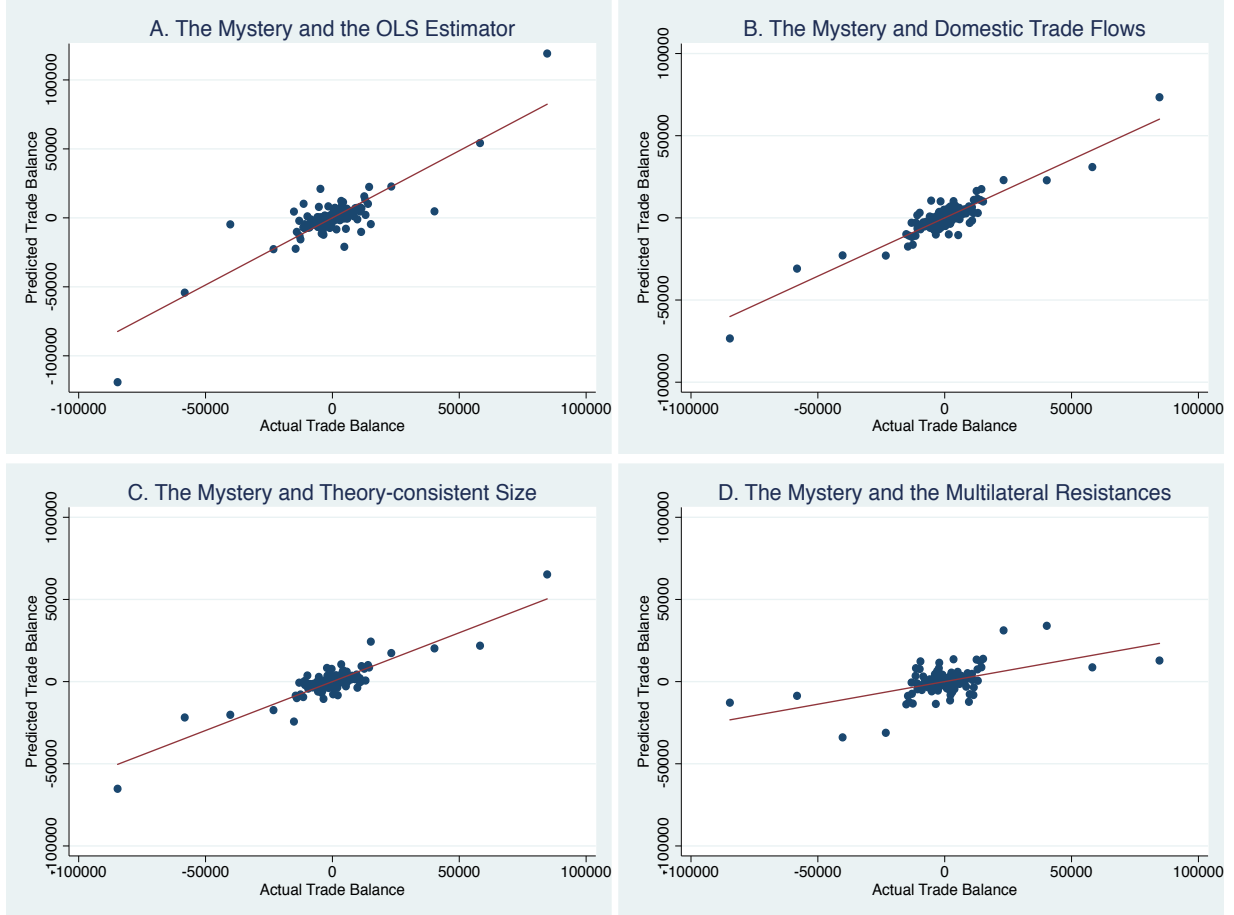
Note: This figure sheds light on the relative importance of the alternative improvements that we implement to resolve the mystery of the trade balances. Panel A implements all improvements except the PPML estimator. Instead, the results in this panel are obtained with the OLS estimator. Panel B implements all improvements except the addition of domestic trade flows/costs. Panel C implements all improvements except for the theoretically motivated size controls. Finally, the results in Panel D are obtained with all other adjustments but with the a-theoretical remoteness indexes instead of the structural multilateral resistances. Alternative statistical tests for importance of each adjustment appear in Table 4. See text for further details.

Figure 4: On the Importance of the MRs
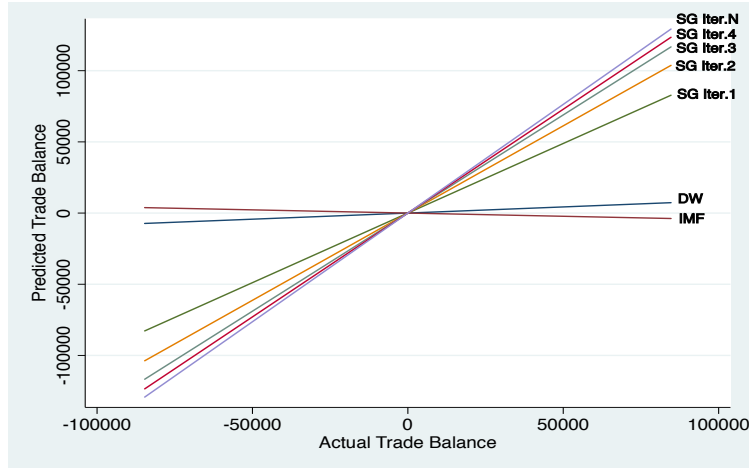
Note: This figure sheds light on the importance of the multilateral resistances for solving the mystery of the trade balances. Specifically, the figure plots the fitted lines between the actual and the predicted bilateral trade balances depending on the construction of the multilateral resistances. All results are obtained from underlying OLS regressions with data on international trade flows only, i.e., no domestic trade flows. Lines DW and IMF are constructed based on the remoteness indexes from Davis and Weinstein (2002) and IMF (2019), respectively. Lines whose labels start with 'SG Iter.' are obtained from an iterative procedure following Anderson and van Wincoop (2003). See text for further details.

Figure 5: On the Importance of the Aggregate Imbalances
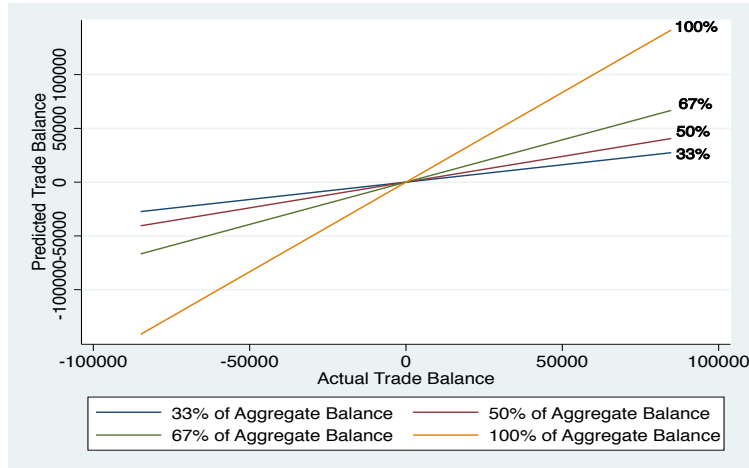
Note: This figure sheds light on the importance of the aggregate trade balances for solving the mystery of the trade balances. Specifically, the figure plots the fitted lines between the actual and the predicted bilateral trade balances depending on how far are the aggregate trade balances used in the model from the actual aggregate trade balances. All results are obtained from underlying OLS regressions with data on international trade flows only, i.e., no domestic trade flows. We start with a specification where all aggregate trade imbalances in our model are set to 33 percent of the corresponding actual aggregate trade imbalances, then we increase the aggregate imbalances in the model to 50 percent, then to 66 percent of the actual imbalances, then we go to 100 percent. See text for further details.

Figure 6: 'Direct' vs 'Total' Bilateral Trade Costs

A. Symmetric Direct/Partial Trade Costs

B. Asymmetric Total Trade Costs
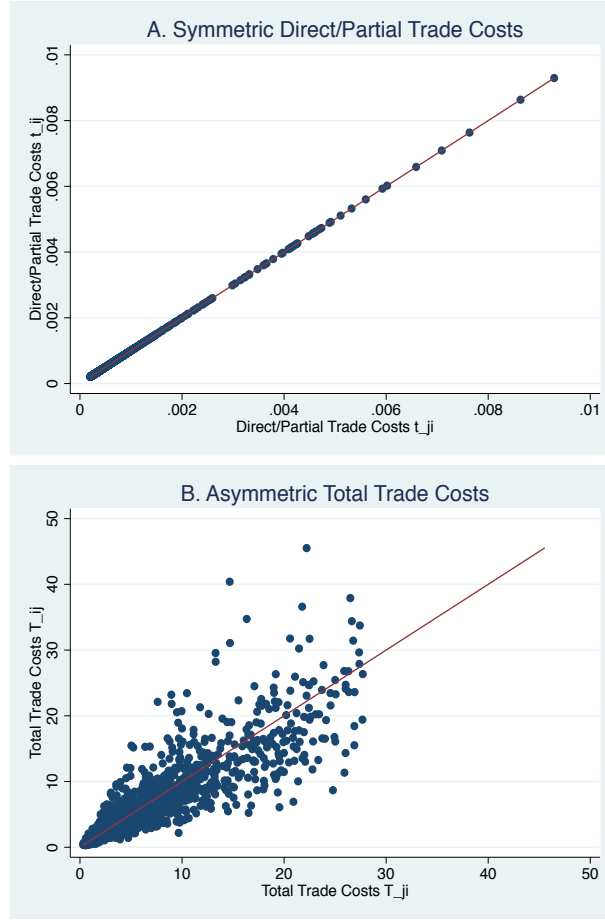
Note: This figure compares the direct/partial and the total bilateral trade costs that are obtained from the same gravity specification with symmetric trade costs. Panel A reports estimates of the partial/direct bilateral trade costs in each direction of trade flows, as proxied by the estimates of the standard gravity variables. Panel B reports the corresponding total trade costs, which also take into account the GE trade cost components as captured by the multilateral resistances. For clarity, we have dropped the largest 5 percent of the total trade cost estimates. See text for further details.

Supplementary Appendix

This Appendix accompanies manuscript *"From Theory to Policy with Gravitas: A Solution to the Mystery of the Excess Trade Balances,"* (2021, by Felbermayr and Yotov) and it is not intended for publication. Appendix A offers details on the data used to obtain the main results. Appendix B presents and describes the results from some additional sensitivity experiments.

Appendix A: Data

We use this Appendix to offer a brief description of the main dataset employed in our analysis. As noted in the main text, the trade data for our analysis come from the dataset constructed by Baier et al. (2016). The main variable that we employ from the original dataset of Baier et al. (2016) is their variable for trade flows. A very important and specific feature of the trade variable from Baier et al. (2016) is that it includes consistently constructed international and domestic trade flows for the countries, which are listed in Table 5.[20] As discussed in the main text, availability of data on domestic trade flows is potentially important for our purposes because (i) the use of domestic trade flows is consistent with structural gravity theory, (ii) on a related note, this is one of the main improvements that we propose in this paper to the original specification of Davis and Weinstein (2002); and (iii) because the use of domestic trade flows can enable us to introduce country-specific trade cost asymmetries in our econometric model.

The primary source for international bilateral trade flows in Baier et al. (2016) is the standard source for such data, i.e., the United Nations' COMTRADE database. As noted above, an important feature of their dataset is that it includes values for "internal trade" flows (a.k.a. "domestic sales"). To construct domestic trade observations, Baier et al. (2016) combine data on industry-level gross output from two main sources: the CEPII TradeProd

---

[20]For our analysis, we do not distinguish between West Germany and Germany. Our main analysis are for the year 2000, while the panel results are for the period 1991-2006. Thus, our sample includes 68 countries. The sensitivity experiments show that the results are robust to alternative sets of countries and data.

database and UNIDO IndStat, and they select both the sample of countries and the period of study in order to achieve the widest possible use of the available production data from these sources. Since production values in TradeProd are largely taken from earlier versions of UNIDO IndStat - and further augmented using the World Bank "Trade, Production, and Protection" database, Baier et al. (2016) generally use TradeProd to provide production data for earlier years and data from UNIDO to fill in later years where needed. To ensure maximum number of production data observations, Baier et al. (2016) also cross-check against the World Bank data to fill in some additional missing values from the beginning of the period.

Baier et al. (2016) then construct domestic trade values as the difference between the value of a country's gross output and the value of its total exports to other markets. In some isolated cases, it is not possible to calculate domestic trade values because the production data is either missing or implies a negative value for domestic trade. Baier et al. (2016) address these issues in a series of steps. First, they apply linear interpolation between non-missing values whenever possible. Second, if values are negative or missing only for a particular industry, we apply the average share of expenditure spent on domestic output by that country on other (non-missing) industries. Finally, they extrapolate any remaining missing production values at the beginning or end of the sample using the evolution of that country's industry-level exports.[21] We refer the reader to Baier et al. (2016) and the accompanying data appendix for further details on the original data.

---

[21]Baier et al. (2016) also experimented with using the U.S. GDP deflator as an alternate basis for extrapolating missing output, following the procedure used in Anderson and Yotov (2016). This method makes virtually no difference for their results.

Table 5: List of Countries, Baier et al. (2016)

3-letter ISO Country Code	Country Name
ARG	Argentina
AUS	Australia
AUT	Austria
BEL	Belgium
BGR	Bulgaria
BOL	Bolivia
BRA	Brazil
CAN	Canada
CHE	Switzerland
CHL	Chile
CHN	China
CMR	Cameroon
COL	Colombia
CRI	Costa Rica
CYP	Cyprus
DEU	Germany
DNK	Denmark
ECU	Ecuador
EGY	Egypt, Arab Rep.
ESP	Spain
FIN	Finland
FRA	France
GBR	United Kingdom
GRC	Greece
HUN	Hungary
IDN	Indonesia
IND	India
IRL	Ireland
IRN	Iran
ISL	Iceland
ISR	Israel
ITA	Italy
JOR	Jordan
JPN	Japan
KEN	Kenya
KOR	Korea, South
KWT	Kuwait
LKA	Sri Lanka
MAC	Macao
MAR	Morocco
MEX	Mexico
MLT	Malta
MMR	Myanmar
MUS	Mauritius
MWI	Malawi
MYS	Malaysia
NER	Niger
NGA	Nigeria
NLD	Netherlands
NOR	Norway
NPL	Nepal
PAN	Panama
PHL	Philippines
POL	Poland
PRT	Portugal
QAT	Qatar
ROM	Romania
SEN	Senegal
SGP	Singapore
SWE	Sweden
THA	Thailand
TTO	Trinidad and Tobago
TUN	Tunisia
TUR	Turkey
TZA	Tanzania
URY	Uruguay
USA	United States
ZAF	South Africa

Notes: This table lists the countries that are included in the sample used to obtain our main analysis. The original data are from the dataset of Baier et al. (2016)

Appendix B: Additional Sensitivity Experiments

We use this Appendix to report the results from some additional sensitivity experiments that we performed in order to test the robustness of our main findings, and which we mention in the main text.

- First, in Table 6, we confirm our main results with data for alternative years (1991-2006) of the manufacturing dataset of Baier et al. (2016). Overall, the estimates across years are similar with moderate variation but no systematic patterns across the years in our main dataset, thus demonstrating that the structural gravity model performs well in each year of the sample.

- Second, we test the robustness of our results to the presence of outliers. Specifically, we drop the countries with the largest actual bilateral trade imbalances, e.g, USA, China, Germany, Canada, and Japan. The resulting statistics for the performance of the model are as follows: Sign – 72%; Variance Ratio – 54%; R^2 – 0.53; Coefficient Estimate – 0.53. While the performance of the gravity model is weaker without the countries with large trade imbalances, the test statistics are still an order of magnitude larger as compared to those from Davis and Weinstein (2002).

- Third, we study the influence of asymmetric trade costs. Staring with the panel specification with symmetric pair fixed effects from Panel C of Table 3, first we show that introduction country-specific asymmetries through interactions between the border dummy variable, $BRDR_{cc'}$, and GDP or GDP per capita (GDPPC) do not improve the fit of the model. To ease comparison, Panel A of Table 7 reproduces our main estimates from Panel C of Table 3, which were based on specification (6) from the main text and are obtained with symmetric pair fixed effects. To obtain the results in panels B and C from Table 7, we add to this specification the interactions $BRDR\_GDP_c$ and $BRDR\_GDPPC_c$, respectively. The results are not affected significantly. Finally, the results in panel D of Table 7 are also obtained from specification (6) but this time,

we replace the symmetric pair fixed effects with asymmetric pair fixed effects, thus introducing and controlling for time-invariant bilateral asymmetries. This adjustment leads to almost perfect fit.

- Finally, we also obtained estimates with an alternative dataset. Specifically, we employ the WIOD dataset, which has several advantages and some caveats. On the positive side, WIOD offers complete sectoral coverage for the countries in the data and this database includes consistently constructed domestic trade flows. The downside of WIOD is that country coverage is limited (to 43 countries) and that the trade data has been adjusted to match the underlying IO linkages. With these caveats in mind, we obtain sectoral trade balance estimates with WIOD and report them for the first and for the last year of the sample (2000 and 20014, respectively) in Tables 8 and 9. Overall, the results confirm our main findings and they are similar across the two years. We also document some differences across sectors (e.g., Mining and Services perform worse on average), which, as noted above, can be used to identify sectors where the modeling of bilateral trade costs can be improved.

Table 6: Robustness: Manufacturing, 1988-2006

	R^2	Reg.Coeff	Var.Ratio	Sign (%)
1991	0.82	0.69	0.57	0.65
1992	0.85	0.73	0.62	0.68
1993	0.87	0.73	0.62	0.67
1994	0.85	0.73	0.63	0.68
1995	0.79	0.69	0.60	0.68
1996	0.76	0.68	0.60	0.69
1997	0.78	0.68	0.59	0.71
1998	0.82	0.73	0.65	0.72
1999	0.85	0.76	0.68	0.70
2000	0.86	0.79	0.72	0.70
2001	0.86	0.77	0.69	0.70
2002	0.85	0.77	0.71	0.73
2003	0.79	0.73	0.67	0.72
2004	0.78	0.71	0.64	0.73
2005	0.79	0.69	0.60	0.73
2006	0.80	0.69	0.58	0.73

Notes: This table reproduces the main results from Table 1 for every year between 1991 and 2006 from the manufacturing dataset of Baier et al. (2016). See text for further details.

Table 7: Asymmetric Direct/Partial Bilateral Trade Costs

Regression Fit (R^2)	Regression Coefficient	Variance Ratio	Sign (%)
A. Symmetric Trade Costs. Symmetric Pair Fixed Effects			
0.88	0.87	0.86	0.74
B. Symmetric Pair Fixed Effects & $BRDR\_GDP_c$			
0.88	0.88	0.88	0.74
c. Symmetric Pair Fixed Effects & $BRDR\_GDPPC_c$			
0.88	0.86	0.85	0.73
D. Asymmetric Trade Costs. Directional Pair Fixed Effects			
0.99	1.01	1.04	0.88

Notes: This table offers statistics that compare predicted vs. actual trade balances from four underlying gravity specifications. In each case, we implement all four adjustments that we propose. All comparisons are for the year 2000. The difference between the four specifications is in the definition of the bilateral trade costs that are used to predict trade flows in a structural gravity specification. For comparison purposes, the results in Panel A replicate our findings from Panel C of Table 3, which are based on specification (6) from the main text and are obtained with symmetric pair fixed effects. To obtain the results in panels B and C, we add to this specification the interactions $BRDR\_GDP_c$ and $BRDR\_GDPPC_c$, respectively. Finally, the results in panel D are also obtained from specification (6) but this time we replace the symmetric pair fixed effects with asymmetric pair fixed effects, thus introducing and controlling for time-invariant bilateral asymmetries.

Table 8: Robustness: WIOD, 2000

Sector	R^2	Reg.Coeff	Var.Ratio	Sign (%)
Crop and animal production	0.64	0.67	0.70	0.76
Forestry and logging	0.95	0.85	0.77	0.77
Fishing and aquaculture	0.81	0.86	0.92	0.73
Mining and quarrying	0.41	0.37	0.34	0.76
Manufacture of food beverages, tobacco	0.80	0.76	0.71	0.73
Manufacture of textiles, apparel, leather	0.73	0.59	0.47	0.72
Manufacture of wood and cork;	0.89	1.14	1.45	0.71
Manufacture of paper and paper products	0.58	0.61	0.64	0.71
Printing and reproduction of recorded media	0.73	0.69	0.65	0.78
Manufacture of coke and refined petroleum	0.84	0.72	0.62	0.66
Manufacture of chemicals and chemical products	0.45	0.42	0.39	0.73
Manufacture of basic pharmaceutical products	0.37	0.31	0.25	0.74
Manufacture of rubber and plastic products	0.71	0.69	0.67	0.75
Manufacture of other non-metallic minerals	0.66	0.64	0.62	0.71
Manufacture of basic metals	0.54	0.53	0.51	0.69
Manufacture of fabricated metal products	0.74	0.69	0.65	0.73
Manufacture of computer, electronic and optical	0.78	0.73	0.68	0.67
Manufacture of electrical equipment	0.66	0.72	0.79	0.73
Manufacture of machinery and equipment n.e.c.	0.76	0.74	0.72	0.68
Manufacture of motor vehicles, trailers and semi-trailers	0.88	0.74	0.62	0.80
Manufacture of other transport equipment	0.74	0.72	0.69	0.76
Manufacture of furniture; other manufacturing	0.85	0.83	0.80	0.73
Repair and installation of machinery and equipment	0.85	0.80	0.74	1.01
Electricity, gas, steam and air conditioning supply	0.67	0.64	0.61	0.67
Water collection, treatment and supply	0.02	-0.24	3.91	0.76
Sewerage; waste collection, disposal;	0.61	0.56	0.51	0.89
Construction	0.43	0.50	0.58	0.70
Wholesale, repair of vehicles and motorcycles	0.98	1.00	1.02	0.80
Wholesale trade, except of vehicles and motorcycles	0.53	0.56	0.60	0.71
Retail trade, except of motor vehicles and motorcycles	0.72	0.65	0.58	0.73
Land transport and transport via pipelines	0.86	0.75	0.66	0.71
Water transport	0.61	0.58	0.55	0.69
Air transport	0.49	0.49	0.48	0.64
Warehousing and support activities for transportation	0.66	0.68	0.69	0.74
Postal and courier activities	0.39	0.45	0.52	0.84
Accommodation and food service activities	0.88	0.81	0.75	0.75
Publishing activities	0.85	0.72	0.62	0.71
Motion picture, video and television, sound	0.46	0.41	0.37	0.81
Telecommunications	0.79	0.81	0.83	0.72
Computer programming, consultancy; information	0.37	0.23	0.14	0.73
Financial services, except insurance and pension	0.51	0.42	0.34	0.73
Insurance, reinsurance and pension funding	0.43	0.29	0.20	0.62
Auxiliary to financial and insurance activities	0.68	0.53	0.42	0.76
Real estate activities	0.75	0.66	0.57	0.78
Legal and accounting, management, consultancy	0.70	0.57	0.47	0.76
Architectural, engineering, technical testing	0.66	0.56	0.48	0.86
Scientific research and development	0.85	0.76	0.68	0.69
Advertising and market research	0.01	0.36	9.97	0.55
Other professional, scientific, veterinary activities	0.79	0.83	0.87	0.79
Administrative and support service activities	0.61	0.65	0.70	0.74
Public administration and defense	0.94	0.93	0.92	0.72
Education	0.56	0.63	0.71	0.72
Human health and social work activities	0.85	0.82	0.78	0.77
Other service activities	0.62	0.54	0.47	0.73
Undifferentiated goods- and services activities	0.04	-0.69	12.15	0.23
Activities of extraterritorial organizations	0.97	0.63	0.41	0.98

Notes: This table reproduces our main results for each sector in WIOD for the year 2000. See text for further details.

Table 9: Robustness: WIOD, 2014

Sector	R^2	Reg.Coeff	Var.Ratio	Sign (%)
Crop and animal production	0.55	0.55	0.54	0.71
Forestry and logging	0.94	0.88	0.82	0.66
Fishing and aquaculture	0.70	0.84	1.00	0.70
Mining and quarrying	0.49	0.50	0.51	0.77
Manufacture of food beverages, tobacco	0.64	0.58	0.52	0.67
Manufacture of textiles, apparel, leather	0.72	0.60	0.51	0.72
Manufacture of wood and cork;	0.39	0.61	0.94	0.62
Manufacture of paper and paper products	0.41	0.41	0.40	0.71
Printing and reproduction of recorded media	0.67	0.52	0.41	0.79
Manufacture of coke and refined petroleum	0.87	0.81	0.76	0.67
Manufacture of chemicals and chemical products	0.23	0.28	0.35	0.68
Manufacture of basic pharmaceutical products	0.24	0.27	0.30	0.63
Manufacture of rubber and plastic products	0.84	0.66	0.52	0.75
Manufacture of other non-metallic minerals	0.70	0.64	0.59	0.60
Manufacture of basic metals	0.63	0.63	0.63	0.65
Manufacture of fabricated metal products	0.71	0.78	0.86	0.69
Manufacture of computer, electronic and optical	0.77	0.60	0.47	0.72
Manufacture of electrical equipment	0.82	0.74	0.68	0.70
Manufacture of machinery and equipment n.e.c.	0.64	0.58	0.53	0.68
Manufacture of motor vehicles, trailers and semi-trailers	0.65	0.60	0.56	0.83
Manufacture of other transport equipment	0.58	0.51	0.44	0.79
Manufacture of furniture; other manufacturing	0.75	0.83	0.93	0.74
Repair and installation of machinery and equipment	0.84	0.77	0.70	0.96
Electricity, gas, steam and air conditioning supply	0.58	0.56	0.53	0.59
Water collection, treatment and supply	0.08	0.36	1.70	0.79
Sewerage; waste collection, disposal;	0.38	0.38	0.39	0.81
Construction	0.79	0.75	0.71	0.71
Wholesale, repair of vehicles and motorcycles	0.97	0.97	0.98	0.82
Wholesale trade, except of vehicles and motorcycles	0.40	0.48	0.58	0.70
Retail trade, except of motor vehicles and motorcycles	0.58	0.54	0.51	0.66
Land transport and transport via pipelines	0.80	0.70	0.61	0.68
Water transport	0.74	0.66	0.58	0.71
Air transport	0.48	0.52	0.57	0.69
Warehousing and support activities for transportation	0.55	0.54	0.53	0.72
Postal and courier activities	0.56	0.57	0.58	0.86
Accommodation and food service activities	0.73	0.76	0.78	0.76
Publishing activities	0.69	0.55	0.44	0.70
Motion picture, video and television, sound	0.79	0.61	0.47	0.82
Telecommunications	0.83	0.66	0.53	0.70
Computer programming, consultancy; information	0.51	0.29	0.16	0.67
Financial services, except insurance and pension	0.50	0.33	0.22	0.66
Insurance, reinsurance and pension funding	0.25	0.19	0.13	0.63
Auxiliary to financial and insurance activities	0.67	0.49	0.36	0.60
Real estate activities	0.54	0.51	0.49	0.73
Legal and accounting, management, consultancy	0.67	0.55	0.44	0.74
Architectural, engineering, technical testing	0.65	0.53	0.43	0.79
Scientific research and development	0.86	0.83	0.80	0.71
Advertising and market research	0.64	0.68	0.73	0.82
Other professional, scientific, veterinary activities	0.85	0.71	0.60	0.82
Administrative and support service activities	0.70	0.62	0.55	0.78
Public administration and defense	0.89	0.87	0.86	0.77
Education	0.69	0.59	0.50	0.69
Human health and social work activities	0.74	0.75	0.76	0.75
Other service activities	0.80	0.66	0.54	0.71
Undifferentiated goods- and services activities	0.02	0.14	1.06	0.38
Activities of extraterritorial organizations	0.98	0.64	0.42	1.26

Notes: This table reproduces our main results for each sector in WIOD for the year 2014. See text for further details.