Does the WTO Promote Trade? A Meta-analysis*

Rodolfo G. Campos[1] Mario Larch[2] Jacopo Timini[1] Elena Vidal[3] Yoto V. Yotov[4]

[1]Banco de España
[2]University of Bayreuth, CEPII, ifo, CESifo and GEP
[3]Banco de España and OECD
[4]Drexel University and CESifo

September 7, 2024

Abstract

The World Trade Organization (WTO) and its predecessor, the General Agreement on Tariffs and Trade (GATT), are key institutions of the multilateral trading system. While the WTO is expected to promote trade by reducing tariffs and non-tariff barriers, existing estimates of its effect on trade vary widely in magnitude, sign, and significance. We collect 2,547 estimates from 71 papers and apply meta-analysis techniques to conduct a systematic quantitative review of the literature, complementing it with established advances in gravity models to obtain estimates of the WTO's impact on trade. The meta-analysis shows that, on average, the literature finds a significant and positive trade effect of the WTO, although the estimates depend strongly on study characteristics. Moreover, we find no evidence of publication bias. Our structural gravity estimates confirm these findings: the WTO increases trade. However, the effects are heterogeneous across sectors and income levels of trading partners.

Keywords: World Trade Organization, trade, gravity model, meta-analysis.

JEL codes: C83, F13, F14, F15.

*An earlier version of this paper circulated under the title: "Star Wars - Return of the WTO trade effect". The title referred to the heated debate in the literature about the statistical significance of the coefficient, as usual indicated by one, two or three stars after the number. We thank Pablo Millan Perez for excellent research assistance and Raquel Lorenzo for her assistance in thoroughly double-checking the data used in the meta-analysis. In writing this paper we have benefited from comments by [acknowledgments to be added in the published version], and by participants at [acknowledgments to be added in the published version]. The views expressed in this paper are those of the authors and do therefore not necessarily reflect those of the Banco de España, the Eurosystem, or the OECD.

1 Introduction

January 1, 2025, will mark the 30th anniversary of the creation of the World Trade Organization (WTO), the successor to the General Agreement on Tariffs and Trade (GATT) signed almost 50 years earlier in 1947. These institutions are responsible for ensuring that there are common rules for trade in goods and services, both in terms of tariffs and non-tariff measures, and for creating fora to resolve trade disputes, discuss new trade rules, and promote transparency in trade policy matters, amongst others.

Yet, despite its long institutional history and the battery of policies it has deployed, there is no consensus on whether the GATT/WTO multilateral trading system promotes trade.[1] Indeed, since Rose (2004) first estimated a limited or null role for the WTO in promoting trade, a plethora of researchers have debated the validity of his findings, and whether they depend on the details of the methodology used. For example, the literature has argued that the size of the WTO trade effect depends on: the characteristics of the trading partners (Subramanian and Wei, 2007); whether trade policy changes other than those directly linked to WTO membership, e.g. trade agreements, are taken into account (Tomz et al., 2007); the set of fixed effects used in the estimation strategy (Roy, 2011; Cheong et al., 2014); or the use of both domestic and international trade data (Larch et al., 2019).

Against this backdrop, this paper makes two contributions. First, we analyze the literature on the WTO trade effect using meta-analysis techniques (Havránek and Irsova, 2011; Havránek et al., 2015; Havránek and Irsova, 2017; Bajzik et al., 2020). We collect 2,547 estimates of the WTO trade effect from 71 papers and construct several indicators that categorize the salient features of the research design employed for each of the estimates. Second, guided by theory and capitalizing on established developments in the empirical trade literature, we obtain an updated set of WTO effects and explore their heterogeneity across several dimensions.

The meta-analysis reveals three important insights. First, despite the doubts raised by Rose's original paper, the cumulative literature shows that the effect of WTO membership on bilateral trade—the "WTO trade effect"—is significantly positive for trade between members and between members and non-members. The average size of these trade effects is about +23% and +14%, respectively. Second, the average estimate masks considerable heterogeneity in the estimates related to study characteristics. Indeed, when controlling for study characteristics, one of the most important being domestic trade, the "WTO trade effect" increases to approximately +50%, aligning closely with our structural gravity

[1] For the sake of simplicity we will use "WTO" in the text for referring to GATT/WTO.

estimates. Third, the empirical literature estimating the WTO effect does not suffer from publication bias. Publication bias refers to the systematic distortion or omission of research results that are not considered important or statistically significant.[2] Publication bias typically occurs when studies with the expected sign or statistically significant results are more likely to be published or given prominence and studies with unexpected or statistically non-significant results are not published. The absence of publication bias in the WTO effect may be a by-product of the fact that the first study found non-significant results, leading to a lively debate with several studies supporting findings in opposite directions.

Our structural gravity analysis complements the meta-analysis by providing updated estimates of the WTO trade effect based on a theory-consistent specification of the model. In doing so, we push the boundaries of the literature along several dimensions: first, we extend the time dimension of our analysis, in some cases starting from 1948. This is important because the closer we get to the present, the fewer countries that are not GATT or WTO members, the smaller the reference group, and the less variation remains from the fixed effects included in typical gravity models. Second, we go beyond the manufacturing sector, which has been the focus of some recent papers analyzing the WTO effect using state-of-the-art techniques (Larch et al., 2019; Felbermayr et al., 2024). Indeed, we include agriculture, mining, manufacturing, and services sectors in our analysis. Third, we compare two different specifications, which can be understood as upper- and lower-bound estimates of the WTO effect. Fourth, we use 9 different databases to obtain our estimates. In this way, we show that the WTO effect is not a by-product of a particular database. Our estimates confirm the main findings of the meta-analysis: the WTO trade effect is large, positive, and significantly different from zero. Our preferred lower- and upper-bound specifications suggest that the increase in trade between WTO members is in the range of +38% to +54%. However, these effects are heterogeneous across sectors and income levels.

The WTO has a large, positive, and significant effect on trade in goods and trade in services, although the effect tends to be larger for the former. Indeed, the trade effect of the WTO in agriculture and manufacturing is about two to three times as large as its effect on services. The effects on trade between advanced economies (AEs) and emerging market and developing economies (EMDEs), or among EMDEs, tend to be larger than those on trade between AEs. Combining these two dimensions, our results suggest that in agriculture, WTO membership tends to favour more exports from EMDEs (both to AEs and to other EMDEs). In mining and energy, WTO membership tends to favour more exports from AEs to

[2] Rothstein et al. (2005) define publication bias as the "tendency to decide to publish a study based on the results of the study rather than on its theoretical or methodological quality".

EMDEs. In manufacturing, WTO membership favours exports across the board, i.e., all combinations of trade by level of development (i.e., AEs-AEs, AEs-EMDEs, EMDEs-AEs, and EMDEs-EMDEs). The results for services are qualitatively similar to those for manufacturing but of a smaller magnitude.

The remainder of the paper is organized as follows: Section 2 offers a brief review of the empirical model, which is typically used to obtain estimates of the effects of GATT and WTO; Section 3 describes how we compiled the estimates available in the literature; Section 4 characterizes the distribution of available estimates and studies whether there is evidence of publication bias; and Section 5 analyzes how various characteristics affect the magnitude of the estimates obtained. Finally, in Section 6 we estimate the impact of the WTO effect on trade using up-to-date practices in estimation. Section 7 concludes.

2 Estimates of the WTO effect

Membership in the WTO often requires the implementation of substantial trade liberalization reforms (e.g., reduction of tariffs and non-tariff measures) and implies easier access to members' markets. Thus, the GATT and WTO are expected to reduce trade barriers, especially between members, and to promote a reciprocal opening of economies to trade. However, in a very influential paper, Rose (2004) reported estimates that challenged the view that WTO membership promotes trade among members. Since then, dozens of researchers have attempted to either refute or rationalize his findings.

Rose (2004), and most subsequent studies, use a gravity model to analyze this question. Gravity models—whose theoretical foundations were laid by Anderson (1979); Bergstrand (1985, 1989) and which have become the workhorse model of trade (Eaton and Kortum, 2002; Anderson and van Wincoop, 2003; Arkolakis et al., 2012)—explain trade flows between two locations in terms of their economic size and distance, the latter being a proxy for trade costs. Researchers using gravity models often estimate regressions of the form:

$$\log X_{ijt} = \gamma_1 Bothin_{ijt} + \gamma_2 Onein_{ijt} + \beta Z_{ijt} + \varepsilon_{ijt}, \tag{1}$$

where X_{ijt} denotes bilateral trade flows from exporter i to importer j over period t, *Bothin* is an indicator that countries i and j are both GATT or WTO members at time t, *Onein* is an indicator that only i or j is a member at time t, Z_{ijt} is a vector of other characteristics included in the specification, and ε_{ijt} is a remainder error term.

Because Equation (1) is a log-linearized transformation of the equation implied by modern trade theory, many researchers, especially in more recent times, have relied on the Poisson Pseudo Maximum Likelihood (PPML) estimator (Santos Silva and Tenreyro, 2006) to estimate the following version of the same equation, where trade is measured in levels:

$$X_{ijt} = \exp(\gamma_1 Bothin_{ijt} + \gamma_2 Onein_{ijt} + \beta Z_{ijt}) + \varepsilon_{ijt}. \tag{2}$$

Regardless of whether the equation to be estimated is (1) or (2), the parameters of interest γ_1 and γ_2 have the same interpretation in both cases: they measure the semi-elasticity of bilateral trade to WTO membership.[3]

3 Literature search and number of observations included in the meta-analysis

To start with a large sample of papers, we first identified all documents that cite Rose (2004) using Google Scholar. We performed this search on Google Scholar on March 30, 2022, using the software Publish or Perish, version 8. We reduced the universe of eligible papers in a step-wise fashion. In the first step, we eliminated books, dissertations, and papers not written in English. We also eliminated papers not accessible online. This left us with 1,821 candidate papers. In a second step, we screened the papers by reading the title, abstract, and introduction to determine whether they contained estimates of the effect of WTO membership on trade. 546 papers survived this first screening step and were considered eligible. The flowchart in Figure 1 shows the information flow at each stage of the literature search, including the number of studies identified, screened, and deemed eligible.

We used three criteria to select estimates from the eligible papers. First, at least one estimate in the paper must be from a gravity equation that regresses trade on a dummy variable indicating WTO membership (either *Bothin* or *Onein*). Second, to be consistent with equations (1) or (2), the dependent

[3] When the indicators *Bothin* and *Onein* are included simultaneously in the regression and data on domestic trade is not used, as in the original article by Rose, the comparison is relative to the group of country pairs in which neither country is a member of the WTO. In later studies, the indicator *Onein* tends to be omitted while *Bothin* is included. The reason is that later papers tend to include country-time fixed effects, the theory-consistent way of accounting for multilateral resistance terms (MRTs), which are co-linear with *Onein*. In these cases, the interpretation of *Bothin* is the same as before. If there are no controls to absorb the effect of *Onein*, then the parameter of *Bothin* estimates the effect on trade of both the exporter and the importer being members of the WTO relative to the counterfactual of at least one of them not being a member. This distinction is often not made by the authors of the papers, who refer to any of the estimates as "the WTO effect". In line with this interpretation, we include all regressions in which at least one of *Bothin* or *Onein* is present.

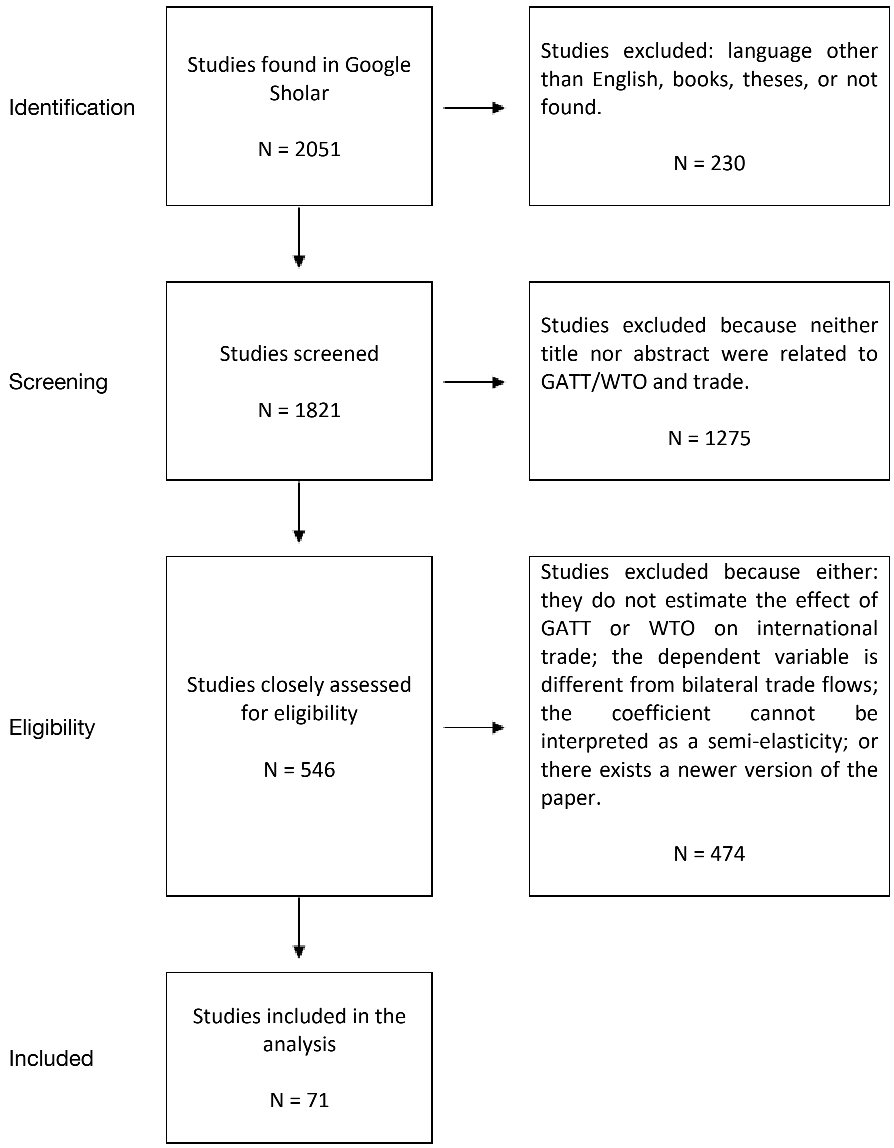


Figure 1: Identification, screening and eligibility of studies

variable must be either the level of trade flows or the log of trade flows, so that the parameter can be interpreted as a semi-elasticity. Hence, we excluded regressions using trade growth, trade volatility, or export and import ratios. Third, we excluded estimations that did not estimate the total effect of trade and focused only on the extensive or intensive margin.[4] If there were multiple versions of the paper, we used the estimates reported in the version published in a journal or, if unpublished, the most recent version.

Our final dataset contains 2,547 estimates from 71 papers. A total of 1,569 of these estimates correspond to the coefficient for the *Bothin* variable and 978 to the coefficient of the *Onein* variable. We record the estimated coefficient and the reported standard error when available. There are 2,284 cases with valid non-missing observations for the standard error. Out of these, 1,422 cases (from 62 studies) correspond to the coefficient for the *Bothin* variable and 862 cases (from 34 studies) to the coefficient of the *Onein* variable. In addition to coefficients and standard errors, we collect information on several additional variables (described in a later section) that capture characteristics of the specification used, the estimation method, or information about the papers in which the regressions are included. Most of the coefficients in the final sample come from publications in peer-reviewed articles. They make up for 2,207 coefficients (1,399 for *Bothin* and 808 for *Onein*). Out of these, 2,188 have valid non-missing observations for the standard error (1,381 for *Bothin* and 807 for *Onein*).

4 Publication bias

In Table 1, we report unweighted and weighted means of the two coefficients of interest and confidence intervals as a measure of the variation in the coefficients. These confidence intervals are obtained by regressing the coefficient on a constant and using standard errors clustered by study to construct the asymptotic distribution of the coefficients around the mean. The table shows that the literature finds point estimates of the effect of WTO membership on bilateral trade—the "WTO trade effect"—that are positive on average for trade between members and between members and non-members. Their magnitudes are centred around 0.21 and 0.13, respectively, implying an increase of trade associated with WTO membership of $\exp(0.21) - 1 = 23\%$ and $\exp(0.13) - 1 = 14\%$.

[4] We excluded estimations when the extensive margin was defined as the number of firms that participate in trade, or products that are traded, and the intensive margin was defined as the ratio of trade value to the number of firms/products. This does not mean that we automatically excluded estimations based on trade flows being strictly positive, as occurs, for example, when the logarithm of trade value is used as the dependent variable.

Table 1: Average estimates in the literature

	Studies	Mean	95% CI		Weighted mean	95% CI	
Bothin (γ_1): Both in WTO	65	0.21	[0.14	0.28]	0.20	[0.15	0.26]
Onein (γ_2): One in WTO	35	0.13	[0.05	0.21]	0.12	[0.05	0.20]
All	71	0.18	[0.12	0.24]	0.17	[0.12	0.22]

Note: The weighted means weight all studies equally, regardless of how many estimates they report. Confidence intervals are calculated by clustering standard errors at the study level.

However, while positive on average, the estimates of the "WTO trade effect" vary widely in the literature (Figure 2), with coefficient values ranging from -5.67 to $+8.01$.[5] The histogram of the empirical distribution of the estimated coefficients has a peak near but above zero for both coefficients of interest (*Bothin* and *Onein*). In the case of *Bothin*, negative coefficients are very close to zero in most cases, while for *Onein* there are several cases where negative coefficients are far from zero. The differences in the empirical distributions of *Bothin* and *Onein* may reflect their different theoretical interpretation. For *Bothin*, there is little theoretical justification for a negative effect because WTO membership reduces bilateral trade costs by reducing trade policy barriers between both trading partners, and is therefore expected to increase trade. The expected sign of *Onein*, on the other hand, can go in either direction, as WTO membership of only the exporter or the importer can, in principle, lead to trade diversion because, once a member, it is more convenient than before to trade with other members, and less convenient than before to trade with non-members.[6] It can also lead to trade creation because (at least part of) WTO members' trade reforms can be unilateral, thus opening the economy to members and non-members or because of dynamic consequences of trade that make members' exports more competitive. Interestingly, estimates vary widely not only between studies but also within studies (see Figures A.2 to A.5 in the Appendix).

The literature suggests several approaches to detect publication bias. It is common to start with a graphical analysis (Egger et al., 1997) using "funnel plots". This technique uses a scatter plot with the estimated size of an effect on the horizontal axis and the precision of the estimates (calculated as the inverse of the standard errors) on the vertical axis. Without publication bias, the scatter plot should lie within lines that look like an inverted funnel: estimates with high precision should be close to the mean or median effect, whereas those with low precision will be more dispersed. The distribution should also

[5] See Figure A.1 in the Appendix for a histogram of t-statistics.

[6] This, of course, is related to the use of country-time fixed effects. If the latter are included, they control for the traditional general equilibrium trade diversion effects. Thus, we would expect that the chance of obtaining a positive estimate on *Onein* is higher when including country-time fixed effects, which is only possible when additionally domestic trade is included in the dataset.

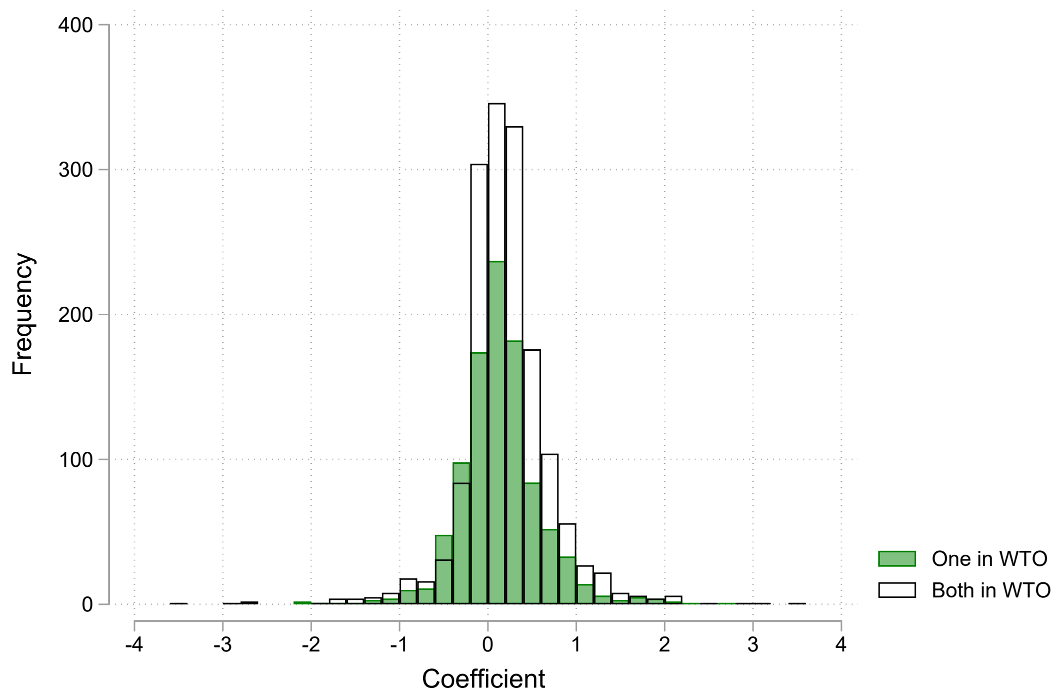
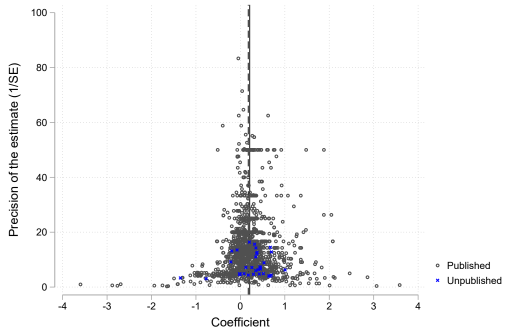
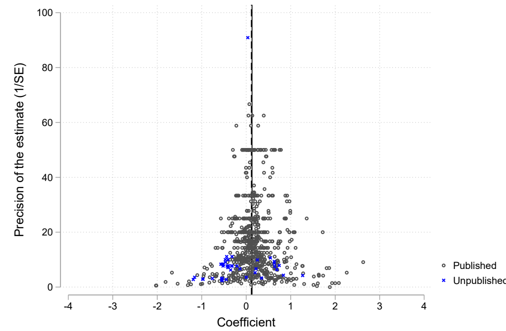


Figure 2: Histogram of coefficients

Note: The figure shows the histogram of the estimates of the WTO effect on trade reported in individual studies.



(a) “Two in WTO”



(b) “One in WTO”

Figure 3: Funnel plots

Note: The figures show scatter plots with the estimated coefficient on the horizontal axis and precision (the inverse of the standard error) on the vertical axis. Estimates that belong to published articles are shown in black and those that belong to unpublished work are shown in blue. The figure on the left shows estimates for *Bothin* and the figure on the right for *Onein*.

be symmetric. Figure 3 does not suggest the presence of publication bias, as the scatter plots for the effect of WTO membership on trade resemble inverted funnels and do not show the pattern typical of publication bias (high precision estimates far from the mean or median). Moreover, the funnel plots do not show an obvious difference between published (shown in black) and unpublished results (shown in blue).

A test of publication bias that goes beyond a visual assessment consists of regressing the estimated coefficients on the standard errors of the estimates. The idea behind this regression is that publication bias leads to a co-movement between the magnitude of estimates and standard errors, which can be detected by finding a non-zero coefficient associated with the standard errors (Stanley, 2005; Havránek, 2010; Havránek and Irsova, 2011; Bajzik et al., 2020). More specifically, the method suggests a regression specified as follows:

$$WTO\ effect_{ij} = \alpha + \beta \times \texttt{se}(WTO\ effect)_{ij} + \epsilon_{ij}, \tag{3}$$

where the left-hand side variable, $WTO\ effect_{ij}$, identifies the estimate i of the WTO trade effect in study j. On the right-hand side of the equation, the explanatory variables are a constant with coefficient α, which captures the mean effect excluding publication bias, and $\texttt{se}(WTO\ effect)_{ij}$, the reported standard error of the WTO effect, whose coefficient β identifies the size and sign of the publication bias. Intuitively, if $\beta = 0$, there is no evidence of publication bias. If $\beta \neq 0$ instead, it means that the estimates and their standard errors are correlated, a finding that suggests the presence of publication bias.

Table 2 shows the results of estimating Equation (3) using different methods. Columns (1) to (3) use the sample of *Onein* coefficients. Column (1) reports the results of a pooled set of coefficients, including all estimates of the type *Onein*. Column (2) restricts the sample to estimates published in a journal. Column (3) includes fixed effects by study. Columns (4) to (6) mimic columns (1) to (3), but using the sample of *Bothin* coefficients.

In the case of *Onein* (columns (1) to (3)), the data suggest no publication bias, as there is no conclusive evidence that the standard errors are systematically related to the coefficient estimates. In the case of *Bothin* (columns (4) to (6)), there is also no clear evidence of publication bias. The point estimates are in the range of 0.26–0.28 after controlling for standard errors.[7]

Table 2: Tests for publication bias

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Standard error	0.002	0.002	0.002	-0.314*	-0.311	-0.436
	(0.312)	(0.291)	(0.352)	(0.189)	(0.196)	(0.340)
Constant	0.138**	0.159***	0.138**	0.258***	0.257***	0.275***
	(0.058)	(0.051)	(0.062)	(0.041)	(0.042)	(0.058)
Observations	862	807	862	1,422	1,381	1,422
R-squared	0.013	0.013	0.016	0.019	0.019	0.029
Number of id			34			62

Note: The dependent variable in columns (1)–(3) is the estimate for "*Onein*". Column (1) includes all estimates of this type. Column (2) only those that were published in a journal. Column (3) includes fixed effects by study. The value reported for the constant is the average of all fixed effects. The dependent variable in columns (4)–(6) is the estimate for "*Bothin*". Column (4) includes all estimates of this type. Column (5) only those that were published in a journal. Column (6) includes fixed effects by study. The value reported for the constant is the average of all fixed effects. Standard errors (in parentheses) are clustered at the study level and bootstrapped with 10,000 replications.

[7] In further robustness tests, reported in Tables C.1 and C.2 in the Appendix, we run the same regressions on a sample winsorized at the 1% level to ensure that results are not driven by outliers, and for different sub-samples of well-identified studies, i.e., those studies that use structural gravity models including all theoretically grounded fixed effects. We also show the same regressions for the rest of the sample, i.e., those studies that do not use structural gravity models. We also use alternative tests of publication bias that relax different assumptions. First, we use the estimator proposed by Ioannidis et al. (2017), which consists of a weighted average of adequately powered (WAAP) results. The use of this estimator is likely to reduce bias due to publication selection, reporting, or small sample bias. Second, we use the Andrews and Kasy (2019) selection model, which relaxes the assumption that the bias is a linear function of the standard error. Third, we use the meta-analysis instrumental variable estimator (MAIVE) proposed by Havranek et al. (2023), which relaxes the assumption that the standard error is exogenous to the research studies and therefore reports estimates that are robust to the standard error being affected by estimation techniques or other issues of the sort. Results are in line with those reported in the main text and are reported in Tables C.3, C.4, and C.5, respectively

5 The impact of estimation characteristics

There are several potential candidates for explaining the heterogeneity in the estimates of the "WTO trade effect", related to the methodology and data used. Gravity theory has developed rapidly over the last twenty years, so many studies may have used different specifications to estimate the gravity model. In addition, studies may focus on different samples, and there are many reasons to explain why there might be different WTO trade effects.

In light of recent advances in gravity theory, the literature suggests that certain features should be included when estimating a structural gravity equation. First, in line with Anderson and van Wincoop (2003), gravity models should account for multilateral resistance terms (MRTs), i.e., the fact that trade between two countries depends on the trade costs that these two countries face when trading with the rest of the world. Eaton and Kortum (2002), Anderson and van Wincoop (2003), Redding and Venables (2004), Feenstra (2004), and Baldwin and Taglioni (2006) suggest that a theory-consistent way to do this empirically is to include exporter-time and importer-time fixed effects. However, many studies do not account for MRTs or do so using atheoretical approaches.

Second, as suggested by Santos Silva and Tenreyro (2006), it is advisable to estimate the gravity equation in its multiplicative form using the Poisson pseudo-maximum likelihood (PPML) estimator. The PPML estimator has the advantage of accounting for heteroscedasticity and zero trade flows. However, various regressions use the log-linearized transformation of the multiplicative trade equation implied by modern trade theory, and in some cases do not account for the loss of information by excluding observations with zeros, which are automatically excluded from the regressions by calculating the logarithm of the dependent variable.

Third, Baier and Bergstrand (2007) argue that trade policy variables in gravity models may be endogenous. For example, countries that trade more with each other may be more likely to sign a trade agreement. To mitigate such bias, they suggest to include pair fixed effects, a practice that has become standard in the literature. Many studies use standard gravity variables instead, such as those that measure bilateral distance, identify a common language, or identify a common border.

Fourth, Yotov (2022) summarizes many theoretical and empirical reasons why the inclusion of domestic trade flows in gravity models is advisable, primarily because it aligns the empirical specification with gravity theory. Although the rationale for using domestic trade has spread rapidly in the literature, it is not yet a standard practice shared by all researchers or practitioners. The sample of countries used is also important. Advanced economies, on the one hand, and emerging and developing economies, on the other, may benefit differently from WTO membership, given their different economic and trade structures and capacities.

To study the effect of these different choices in estimation, we code six variables that describe the characteristics of the methodology and data used and include these variables as additional explanatory variables in Equation (3). These indicators consist of a dummy variable indicating the absence of the use of multilateral resistance terms (MRTs) or the use of atheoretical ones,[8] a dummy variable for an estimation method other than PPML, a dummy variable for the absence of pair fixed effects, a dummy variable for the absence of domestic trade, and dummy variables for when the regression is on a sample in which both are AEs or one is an AE and the other is not (most coefficients come from regressions that have no restriction of this form).

Table 3: Estimation characteristics (*Bothin*)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Standard error	-0.314*	-0.295	-0.316*	-0.309	-0.313*	-0.310*	-0.291
	(0.189)	(0.189)	(0.190)	(0.193)	(0.188)	(0.184)	(0.214)
No or atheorical MTRs		0.080					0.057
		(0.077)					(0.073)
Other than PPML			0.237***				0.240***
			(0.066)				(0.069)
No pair FE				-0.056			-0.089*
				(0.055)			(0.046)
No domestic trade					-0.166***		-0.351***
					(0.051)		(0.072)
One advanced						-0.091	0.001
						(0.113)	(0.082)
Both advanced						0.065	0.057
						(0.106)	(0.089)
Constant	0.258***	0.202***	0.082	0.289***	0.422***	0.261***	0.429***
	(0.041)	(0.068)	(0.062)	(0.050)	(0.040)	(0.037)	(0.054)
Observations	1,422	1,422	1,422	1,422	1,422	1,422	1,422
R-squared	0.019	0.024	0.059	0.022	0.021	0.023	0.073

Note: The dependent variable is the estimate for "*Bothin*". Standard errors (in parentheses) are clustered at the study level and bootstrapped with 10,000 replications.

Tables 3 and 4 summarize the impact of different ways of estimating gravity equations on the WTO effect. Table 3 shows the results for the coefficient of *Bothin* and Table 4 shows the results for the coefficient of

[8] They are usually approximated by the so-called "remoteness indexes", computed as GDP-weighted distance averages.

Table 4: Estimation characteristics (*Onein*)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Standard error	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	(0.312)	(0.400)	(0.311)	(0.309)	(0.317)	(0.319)	(0.357)
No or atheorical MTRs		-0.106					0.046
		(0.228)					(0.282)
Other than PPML			-0.207				-0.036
			(0.151)				(0.107)
No pair FE				-0.160**			-0.207***
				(0.071)			(0.063)
No domestic trade					-0.486***		-0.369
					(0.048)		(0.307)
One advanced						0.355**	0.434*
						(0.173)	(0.240)
Both advanced						0.293**	0.318**
						(0.127)	(0.139)
Constant	0.138**	0.235	0.338**	0.245***	0.618***	0.088	0.575***
	(0.058)	(0.233)	(0.146)	(0.056)	(0.051)	(0.058)	(0.065)
Observations	862	862	862	862	862	862	862
R-squared	0.013	0.016	0.018	0.033	0.023	0.064	0.108

Note: The dependent variable is the estimate for "*Onein*". Standard errors (in parentheses) are clustered at the study level and bootstrapped with 10,000 replications.

Onein. The overarching conclusion from this analysis is that some of the estimation characteristics can have much larger effects than the potential publication bias in this case.

Among the variables considered, the most important negative effects on the coefficients for *Onein* and *Bothin* come from not using domestic trade or not controlling for pair fixed effects. This means that the omission of domestic trade or pair fixed effects leads to lower coefficients on average in the sample of existing estimates. In the case of *Onein*, the coefficients related to a sample consisting exclusively of AEs are also significantly larger than those from regressions without this restriction. The last column shows that the estimate of the WTO effects is significantly larger using the current best practices in estimation. The constant in the last column should be interpreted as the average effect coming from regressions in which theoretical MRTs are used, zero trade flows are not excluded, the PPML estimator is used, pair fixed effects are used to control for endogeneity, domestic trade data are available, and the sample does not select only AEs. Such a regression yields point estimates for the two coefficients of interest that are positive and substantially larger than the average estimates in the literature.[9]

[9] In further robustness tests, reported in Tables C.6 and C.7 in the Appendix, we run the same regressions on a sample winsorized at the 1% level to ensure that results are not driven by outliers. Results are in line with those reported in the main text

6 (Structural) Gravity estimates of the WTO trade effect

We now turn to estimating the impact of the WTO on bilateral trade using a theory-consistent version of the gravity specification. In this section, we focus specifically on examining the impact of the WTO among its member countries (i.e., the *Bothin* coefficient in the previous section) and relegate the effect between members and non-members (*Onein*) to the end of the section. We use a standard empirical specification (e.g., Yotov et al., 2016), which takes the form:

$$X_{ijt} = \exp(\beta_1 WTO_{ijt} + \psi Z'_{ijt} + \eta_{it} + \theta_{jt} + \omega_{ij}) + \varepsilon_{ijt}, \tag{4}$$

where the dependent variable X_{ijt} refers to gross bilateral trade flows between the exporter i and importer j in year t, also including the case where $i = j$ (i.e., domestic trade flows). The variable WTO_{ijt} is a dummy variable, equal to 1 when countries i and j are both members of the WTO at time t, and zero otherwise. Thus, the coefficient β_1 captures the semi-elasticity of bilateral trade flows for WTO membership. The vector Z'_{ijt} contains control variables that capture other time-varying trade policies. In our main specification, this vector includes dummy variables identifying membership in trade agreements (TA_{ijt}) and the European Union (EU_{ijt}). As a robustness check, we further modify and expand the vector Z'_{ijt}, for example, by allowing the effect of bilateral distance to vary over time, an econometric strategy often interpreted as a proxy for bilateral transport costs. The terms η_{it} and θ_{jt} are exporter-time and importer-time fixed effects, respectively. They are important for two main reasons: first, they represent the theory-consistent way to control for "multilateral resistance terms" (Anderson and van Wincoop, 2003); second, they control for all features having country-time variation (such as GDP, GDP per capita, population, etc.). The term ω_{ij} represents directional (i.e., depending on the direction of trade) country-pair fixed effects, controlling for all features with exporter-importer (pair) variation (such as the standard gravity variables: distance, contiguity, common language, colonial relationship, etc.). It is also a standard way of accounting for non-time-varying non-observable components of trade costs. ε_{ijt} is the error term.

When gravity is estimated with domestic trade flows, it is standard practice to include a time-varying dummy variable that distinguishes international trade flows from domestic trade flows in gravity equations that consider both domestic and international trade flows. The coefficients of this variable are usually interpreted as a proxy for trade globalization because they capture the ease of trading internationally relative to trading domestically (Bergstrand et al., 2015). This avoids attributing general changes in the ease of trading across international borders brought about by multilateral progress to other bilateral or regional trade policy variables, such as trade agreements. However, as our primary focus is estimating the WTO effect, this can be problematic.

Unlike the so-called first globalization (1870–1913) where the ease of trading internationally was mainly driven by falling transport costs (see, e.g., Jacks et al., 2010), globalization in the post-World War II period (our period of analysis) was mainly driven by falling trade policy-related costs (see, e.g., Anderson and van Wincoop, 2004), such as multilateral tariff reductions promoted by the WTO. In this case, including the trade globalization proxy will remove some of the WTO effects from the WTO variable. There is no simple solution to this problem. Therefore, we interpret the estimates from specifications that do not explicitly use time-varying border variables to control for common globalization effects as an upper-bound estimate of the WTO effect. In addition, however, we also estimate a version of this equation with the addition of a time-varying dummy variable as follows:

$$X_{ijt} = \exp(\beta_1 WTO_{ijt} + \phi_t b_{ij} + \psi Z'_{ijt} + \eta_{it} + \theta_{jt} + \omega_{ij}) + \varepsilon_{ijt}. \tag{5}$$

The only difference to the previous equation is the appearance of $\phi_t b_{ij}$, where the variable b_{ij} is a dummy variable that distinguishes international trade flows ($b_{ij} = 1$) from domestic trade flows ($b_{ii} = 0$), and the time-varying coefficients $\{\phi_t\}$ capture the semi-elasticity of bilateral trade flows to crossing an international border in different years. Following our previous discussion, we interpret the specification of equation (5) as a lower-bound estimate of the WTO effect.

Data on trade flows (1948–2019) come from the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) gravity database (Head et al., 2010), which collects and reports bilateral merchandise trade flows from the International Monetary Fund's (IMF) Direction of Trade Statistics (DOTS) database, reported by the country of destination. Trade flows are gross, nominal, and expressed in the same currency (thousands of US dollars). We construct domestic trade flows as the difference between nominal Gross Domestic Product (GDP) and nominal total exports (the latter is calculated as the sum of bilateral exports; both GDP and trade data come from the same CEPII gravity database). While using gross output instead of GDP would be a more theoretically consistent way of constructing domestic trade flows, relying on a GDP-based measure of domestic trade significantly increases our sample in terms of years and countries included. This ensures that we are best equipped to (1) study early entries into the

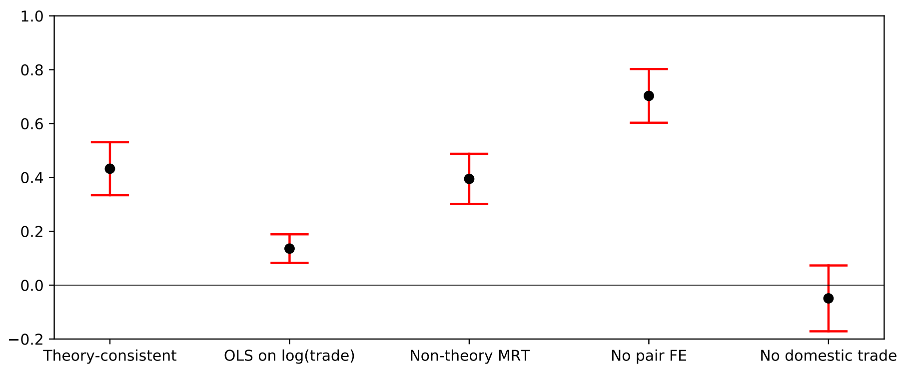
WTO system; (2) observe entries from both AEs and EMDEs; and (3) have a large group of untreated bilateral pairs. Also, the problems derived from using a GDP-based measure may be alleviated by the rich set of dummy variables. Indeed, Campos et al. (2021) compare the results of gravity models using GDP-based measures of domestic trade with those using gross output and find that the presence of country and time-fixed effects in gravity equations makes such a difference less relevant in practical applications.

As a further robustness test, we address the issue related to the measurement of domestic trade by using six alternative databases, four of them with domestic trade constructed as a GDP-based measure and two with domestic trade as a gross output-based measure (and already available in the database). These additional databases are the following: merchandise trade flows from IMF DOTS as reported by the country of origin (1948-2019), merchandise trade flows from United Nations Comtrade database as reported by the country of destination (1962-2019), merchandise trade flows from United Nations Comtrade database as reported by the country of origin (1962-2019), merchandise trade flows from CEPII BACI (1996-2019, Gaulier and Zignago, 2010), manufacturing trade flows from CEPII TradeProd (1966-2018, de Sousa et al., 2012; Mayer et al., 2023); manufacturing trade flows from the WTO Structural Gravity Database (1980-2016; Larch et al., 2019). The results are consistent with those of our main database, even though they are not strictly comparable because of the coverage of different countries, years, and sectors.[10] We report the results from this robustness exercise in the Appendix. Finally, the data on WTO membership also come from the (CEPII) gravity database (Head et al., 2010).

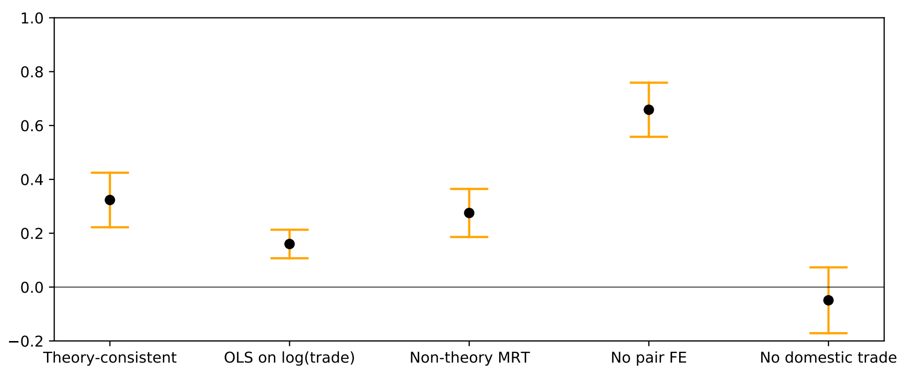
In Figure 4, we show the point estimates and confidence intervals for the two alternative specifications (with and without border-time dummies) obtained from estimating a theory-consistent structural gravity model. We also show the results obtained when one deviates from standard practice, such as using OLS regression on log trade, not controlling for MRTs, not including pair fixed effects, or not considering domestic trade flows. Both the "upper bound" and "lower bound" equations yield similar patterns across specifications.

In the preferred estimation, the upper-bound estimate yields a point estimate of 0.43 implying an increase in trade flows of $\exp(0.43) - 1 = 54\%$. The lower-bound estimate yields a point estimate of 0.32 and an increase in trade flows of $\exp(0.32) - 1 = 38\%$. We note a substantial effect of not including pair fixed effects on the WTO coefficient. This finding is consistent with Baier and Bergstrand (2007)

[10]Gross output-based databases include only manufacturing output and trade, whereas GDP-based databases include all merchandise trade.



(a) Upper bound



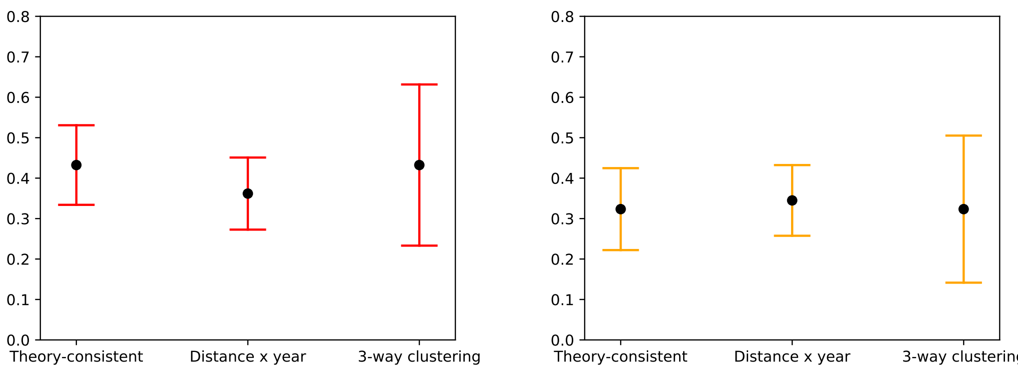
(b) Lower bound

Figure 4: Deviations from standard practice

Note: The figures show point estimates (black dots) and 90% confidence intervals (vertical lines). Standard errors are clustered by country pair. Estimations use data from the DOTS database. The top panel ("upper bound") shows results from estimations that do not control for border-year dummies. The bottom panel ("lower bound") shows results from estimations that control for border-year dummies.

highlighting the importance of using a panel approach with country-pair fixed effects to fully control for all time-invariant bilateral trade costs and mitigate endogeneity concerns, which otherwise is likely to lead to an upward bias in the coefficient estimates. We also note the key role of domestic trade in the empirical estimates, the exclusion of which can lead to a downward bias in the estimated coefficient of the trade policy variables, as demonstrated for the WTO effects by Larch et al. (2019).

We test the robustness of our preferred specification by allowing the effect of bilateral distance to vary over time or by testing alternative clustering strategies, as suggested by Egger and Tarlea (2015). The former allows to capture the effects of policy and non-policy variables, such as the construction of road or rail infrastructure connecting (or improving) two countries. The results are consistent with our main specification and are shown in Figure 5.



(a) Upper bound

(b) Lower bound

Figure 5: Robustness

Note: The figures show point estimates (black dots) and 90% confidence intervals (vertical lines). 'Distance x year' refers to a specification that adds interactions of the logarithm of distance and year dummies to the covariates. Unless otherwise indicated, standard errors are clustered by country pair. '3-way clustering' refers to clustering by exporter, importer, and time. Estimations use data from the IMF DOTS database. The three estimations in the panel on the left do not control for border-year dummies ("upper bound"). The three estimations in the panel on the right control for border-year dummies ("lower bound").

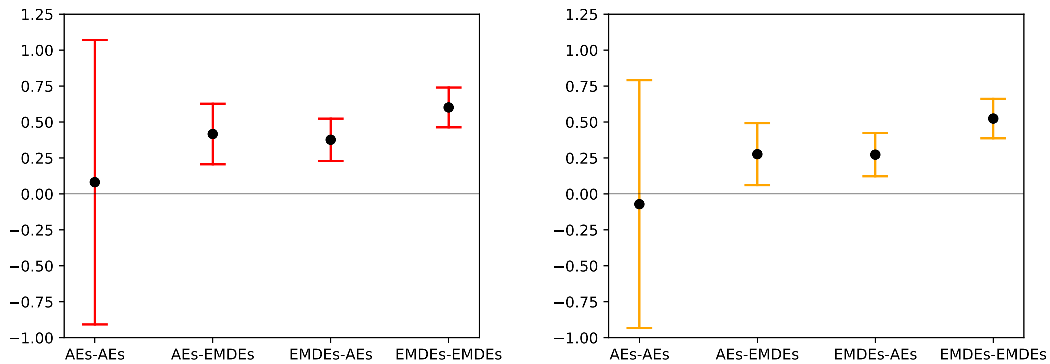
We next examine whether the large, positive, and significant effect of WTO membership is heterogeneous across country income levels. To do this, we interact our main variable of interest (WTO_{ijt}) with four different dummy variables that capture whether the trade flow is between two advanced economies (AEs-AEs), between an advanced economy and an emerging economy (also taking into account who is the exporter and who is the importer; AEs-EMDEs and EMDEs-AEs), or between emerging markets and developing economies (EMDEs-EMDEs).

The estimates in Figure 6 show that the WTO effect is largely concentrated in AE-EMDE and EMDE-EMDE trade flows. These results support the idea that the reductions in trade barriers, such as tariffs and non-tariff barriers, achieved and implemented through successive rounds of trade negotiations have promoted the integration of EMDEs into global value chains, while at the same time allowing AEs to gain access to lower-cost inputs. This mutually beneficial relationship has enabled both AEs and EMDEs to expand their exports to each other. At the same time, many EMDEs (such as China, India, or Brazil, to name a few) have joined or deepened their participation in the global trading system in recent decades. By doing so, they also promote trade among EMDEs by strengthening regional value chains based on WTO integration. On the other hand, trade among AEs does not seem to have been significantly affected by WTO membership, perhaps suggesting that other trade integration initiatives—such as the European Union—, which have further promoted integration among AEs, may have played a more important role in this case.

We also examine whether the WTO effect is heterogeneous across sectors. For this purpose, we use the OECD TiVA database, which contains consistent international and domestic trade data for 76 economies from 1995 to 2019. The OECD TiVA database provides bilateral trade data for 45 industries that can be consistently aggregated, as we do, into four main economic sectors: agriculture, mining, manufacturing, and services.[11] Figure 7 shows estimates of the WTO effect on bilateral trade flows between members for the four main economic sectors. The main message from this figure is that the WTO has had a large, positive, and significant effect on trade in goods and trade in services, although the effect tends to be larger for the former group. The effect of the WTO on agricultural and manufacturing trade flows is about two to three times larger than its effect on services.

Figure 8 shows sector estimates that allow for heterogeneity in the income levels of trading partners. According to our main estimates, WTO membership favours more exports from EMDEs in agriculture (both to AEs and other EMDEs). In mining and energy, WTO membership favours more exports from AEs to EMDEs. In manufacturing, WTO membership favours exports across the board, i.e., across all levels of development (i.e., AEs-AEs, AEs-EMDEs, EMDEs-AEs, and EMDEs-EMDEs). For services, the results are qualitatively similar to those for manufacturing, but the effects are smaller.

[11] For a more granular analysis, we exploit the ITPD-E database (The International Trade and Production Database for Estimation), which contains consistent data on international and domestic trade for industries within the agriculture, mining, energy, manufacturing, and services sectors (1986-2019 for agriculture; 1988-2019 for mining, energy and manufacturing, 2000-2019 for services; see Borchert et al. (2021) for more details). Results, reported in Figure B.1 using this database also show that the WTO trade effects are large and positive, but very heterogeneous across sectors and industries within sectors.

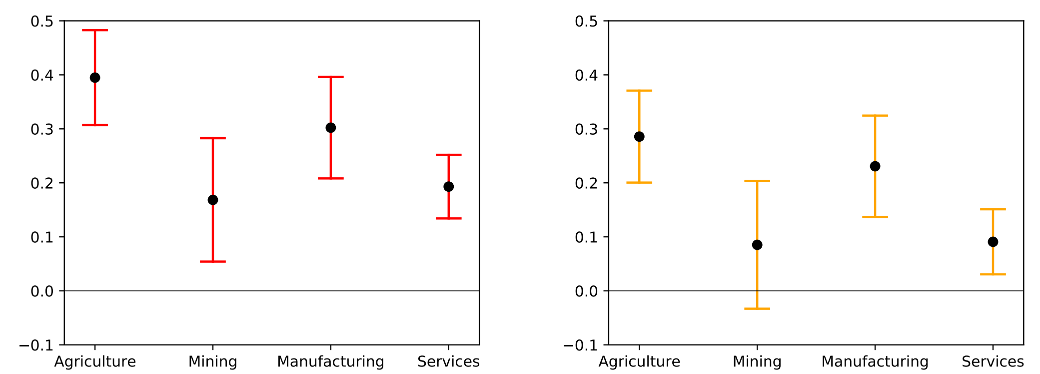


(a) Upper bound

(b) Lower bound

Figure 6: Results by country income level

Note: The figures show point estimates (black dots) and 90% confidence intervals (vertical lines). Standard errors are clustered by country pair. Estimations use data from the IMF DOTS database. The label 'AEs' refers to 'Advanced Economies'. The label 'EMDEs' refers to 'Emerging Market and Developing Economies', as classified by the IMF. The four estimations on the left do not control for border-year dummies ("upper bound"). The four estimations on the right control for border-year dummies ("lower bound").

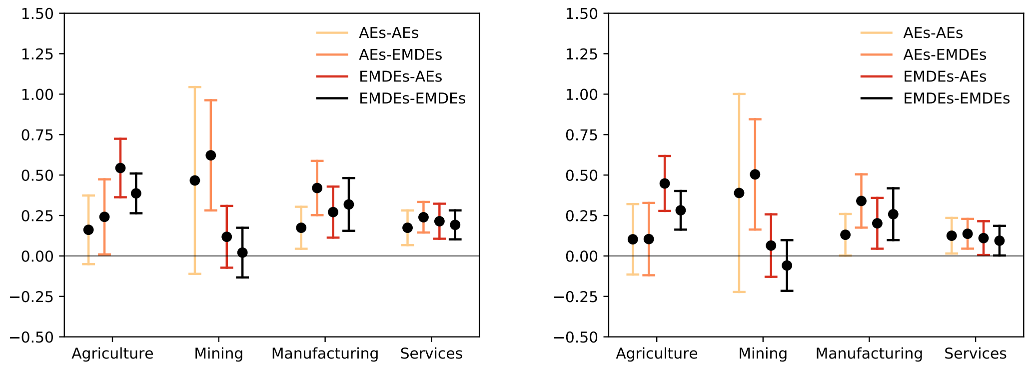


(a) Upper bound

(b) Lower bound

Figure 7: Results by sector

Note: The figures show point estimates (black dots) and 90% confidence intervals (vertical lines). Standard errors are clustered by country pair. The sector labelled mining also includes energy. Estimations use data from the OECD TiVA database. The four estimations on the left do not control for border-year dummies ("upper bound"). The four estimations on the right control for border-year dummies ("lower bound").



(a) Upper bound

(b) Lower bound

Figure 8: Results by sector and country income level

Note: The figures show point estimates (black dots) and 90% confidence intervals (vertical lines). Standard errors are clustered by country pair. The sector labelled mining also includes energy. Estimations use data from the OECD TiVA database. The label 'AEs' refers to 'Advanced Economies'. The label 'EMDEs' refers to 'Emerging Market and Developing Economies', as classified by the IMF. The four estimations on the left do not control for border-year dummies ("upper bound"). The four estimations on the right control for border-year dummies ("lower bound").

In addition, we also distinguish between GATT and WTO in our analysis. The coefficients obtained for GATT and WTO separately, as shown in Table B.4 in the Appendix, are similar in magnitude across all specifications. However, the coefficient for the WTO is estimated with greater precision, possibly due to the larger number of trade flows and entries documented in our databases for the period during which the WTO was active.

Finally, we extend our analysis to assess the impact of the WTO on trade flows between members and non-members. WTO members sometimes reduce trade barriers against all countries, not just other WTO members. Our estimates (reported in Table 5) show that the WTO trade effect tends to be positive and significant for trade flows between members and non-members. Moreover, our results consistently indicate that the magnitude of the WTO trade effect for trade flows between members and non-members is smaller than the effect observed for trade flows between members. This suggests that non-members do not take full advantage of the benefits of WTO membership and that only part of the WTO-related reductions in trade barriers are granted to all. If anything, our findings support the notion that the WTO is primarily a trade-creating institution rather than a trade-diverting one. These results are consistent with previous studies on this topic, which have also found a positive impact of the WTO on trade flows between members and non-members (Subramanian and Wei, 2007; Larch et al., 2019).

Table 5: Both in WTO versus One in

	Upper bound		Lower bound	
	Both in	One in	Both in	One in
IMF DOTS (importer)	0.71 (0.13)	0.28 (0.11)	0.55 (0.13)	0.23 (0.11)
IMF DOTS (exporter)	0.81 (0.13)	0.35 (0.11)	0.62 (0.13)	0.29 (0.11)
UN Comtrade (importer)	0.68 (0.15)	0.28 (0.13)	0.57 (0.14)	0.26 (0.12)
UN Comtrade (exporter)	0.87 (0.17)	0.49 (0.16)	0.75 (0.16)	0.46 (0.15)
BACI CEPII	0.35 (0.12)	0.10 (0.10)	0.28 (0.12)	0.08 (0.09)
TradeProd CEPII	0.31 (0.12)	-0.00 (0.09)	0.15 (0.12)	-0.02 (0.09)
WTO Structural DB	0.59 (0.20)	0.05 (0.19)	0.21 (0.20)	-0.05 (0.19)

Note: The table reports the point estimate and standard deviation for the variables *Bothin* and *Onein*. Standard errors are clustered by country pair. Results in the first two columns do not control for border-year dummies ("upper bound"). Results in the last two columns control for border-year dummies ("lower bound"). The rows correspond to various databases used in the estimation.

7 Conclusions

In this paper, we perform a quantitative review of the literature using meta-analysis techniques, complemented by direct empirical analysis of trade data based on theory-consistent structural gravity models, to assess the trade effect of the WTO. Specifically, we first collect 2,483 estimates of the WTO trade effect from 71 papers and construct variables characterizing the salient features of the methodology used for each estimate. We then assess the presence of publication bias and the impact of estimation characteristics on the estimated WTO trade effect. We also implement a theory-consistent gravity model to provide a theory-based update of the estimated WTO trade effects. In our preferred specification, we use both domestic and international trade flows, three-way fixed effects (exporter-time, importer-time, exporter-importer) to account for MRTs and trade policy endogeneity, and apply a Poisson pseudo-maximum likelihood estimator.

Our meta-analysis suggests that the WTO trade effect estimated in the literature is, on average, significantly positive. We also find no evidence of publication bias, a finding that may be related to the heated debate in the literature. However, the reported estimates vary considerably and depend on the characteristics of the studies. In other words, how gravity equations are specified matters for trade cost estimates such as WTO membership. This conclusion should not be surprising given recent theoretical and empirical advances in the literature. Our structural gravity estimates confirm these findings: the WTO matters for trade. Our preferred lower- and upper-bound specifications indicate that the increase in trade between WTO members is large, positive, and significant, ranging from +38% to +54%. However, these effects are heterogeneous across sectors and income levels of trading partners.

Looking ahead, our evidence of a large but heterogeneous increase in trade among WTO members does not shed light on within-country effects associated with WTO membership (e.g., firm- or household-level effects). This issue merits further research and consideration.

References

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

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

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

Andrews, I. and M. Kasy (2019): "Identification of and Correction for Publication Bias," *American Economic Review*, 109, 2766–2794.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012): "New Trade Models, Same Old Gains?" *American Economic Review*, 102, 94–130.

Baier, S. L. and J. H. Bergstrand (2007): "Do Free Trade Agreements Actually Increase Members' International Trade?" *Journal of International Economics*, 71, 72–95.

Bajzik, J., T. Havranek, Z. Irsova, and J. Schwarz (2020): "Estimating the Armington Elasticity: The Importance of Study Design and Publication Bias," *Journal of International Economics*, 127.

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

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

——— (1989): "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade," *Review of Economics and Statistics*, 71, 143–153.

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

Borchert, I., M. Larch, S. Shikher, and Y. V. Yotov (2021): "The International Trade and Production Database for Estimation (ITPD-E)," *International Economics*, 166, 140–166.

Campos, R. G., J. Timini, and E. Vidal (2021): "Structural Gravity and Trade Agreements: Does the Measurement of Domestic Trade Matter?" *Economics Letters*, 208.

Cheong, J., D. W. Kwak, and K. K. Tang (2014): "The WTO Puzzle, Multilateral Resistance Terms and Multicollinearity," *Applied Economics Letters*, 21, 928–933.

de Sousa, J., T. Mayer, and S. Zignago (2012): "Market Access in Global and Regional Trade," *Regional Science and Urban Economics*, 42, 1037–1052.

Eaton, J. and S. Kortum (2002): "Technology, Geography and Trade," *Econometrica*, 70, 1741–1779.

Egger, M., G. D. Smith, M. Schneider, and C. Minder (1997): "Bias in Meta-analysis Detected by a Simple, Graphical Test," *British Medical Journal*, 315, 629–634.

Egger, P. H. and F. Tarlea (2015): "Multi-way Clustering Estimation of Standard Errors in Gravity Models," *Economics Letters*, 134, 144–147.

Feenstra, R. (2004): *Advanced International Trade: Theory and Evidence*, Princeton, New Jersey: Princeton University Press.

Felbermayr, G., M. Larch, E. Yalcin, and Y. V. Yotov (2024): "On the Heterogeneous Trade and Welfare Effects of GATT/WTO Membership," *Review of World Economics*.

Gaulier, G. and S. Zignago (2010): "BACI: International Trade Database at the Product-Level. The 1994-2007 Version," Working Papers 2010-23, CEPII.

Havranek, T., Z. Irsova, P. R. D. Bom, and H. Rachinger (2023): "Spurious Precision in Meta-Analysis," CEPR Discussion Papers 17927, C.E.P.R. Discussion Papers.

Havránek, T. (2010): "Rose Effect and the Euro: Is the Magic Gone?" *Review of World Economics*, 146, 241–261.

Havránek, T., R. Horvath, Z. Irsova, and M. Rusnak (2015): "Cross-country heterogeneity in intertemporal substitution," *Journal of International Economics*, 96, 100–118.

Havránek, T. and Z. Irsova (2011): "Estimating Vertical Spillovers from FDI: Why Results Vary and What the True Effect Is," *Journal of International Economics*, 85, 234–244.

——— (2017): "Do Borders Really Slash Trade? A Meta-Analysis," *IMF Economic Review*, 65, 365–396.

Head, K., T. Mayer, and J. Ries (2010): "The Erosion of Colonial Trade Linkages After Independence," *Journal of International Economics*, 81, 1–14.

Ioannidis, J. P. A., T. D. Stanley, and H. Doucouliagos (2017): "The Power of Bias in Economics Research," *Economic Journal*, 127, 236–265.

Jacks, D. S., C. M. Meissner, and D. Novy (2010): "Trade Costs in the First Wave of Globalization," *Explorations in Economic History*, 47, 127–141.

Larch, M., J.-A. Monteiro, R. Piermartini, and Y. V. Yotov (2019): "On the Effects of GATT/WTO Membership on Trade: They are Positive and Large After All," WTO Staff Working Papers ERSD-2019-09, World Trade Organization (WTO), Economic Research and Statistics Division.

Mayer, T., G. Santoni, and V. Vicard (2023): "The CEPII Trade and Production Database," Working Papers 2023-01, CEPII research center.

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

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

Rothstein, H., A. Sutton, and M. Borenstein (2005): *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustment*, New York: Wiley.

Roy, J. (2011): "Is the WTO Mystery Really Solved?" *Economics Letters*, 113, 127–130.

Santos Silva, J. and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88, 641–658.

Stanley, T. D. (2005): "Beyond Publication Bias," *Journal of Economic Surveys*, 19, 309–345.

Subramanian, A. and S.-J. Wei (2007): "The WTO Promotes Trade, Strongly but Unevenly," *Journal of International Economics*, 72, 151–175.

Tomz, M., J. L. Goldstein, and D. Rivers (2007): "Do We Really Know That the WTO Increases Trade? Comment," *American Economic Review*, 97, 2005–2018.

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

YOTOV, Y. V., R. PIERMARTINI, J. MONTEIRO, AND M. LARCH (2016): *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, Geneva, Switzerland: United Nations and World Trade Organization.

A Meta-analysis

References of papers surveyed in the meta-analysis

Aichele, R. and G. Felbermayr (2015): "Kyoto and Carbon Leakage: An Empirical Analysis of the Carbon Content of Bilateral Trade," *The Review of Economics and Statistics*, 97, 104–115.

Allee, T. L. and J. E. Scalera (2012): "The Divergent Effects of Joining International Organizations: Trade Gains and the Rigors of WTO Accession," *International Organization*, 66, 243–276.

Balding, C. (2010): "Joining the World Trade Organization: What is the Impact?" *Review of International Economics*, 18, 193–206.

Benedictis, L. D. and L. Tajoli (2011): "The World Trade Network," *The World Economy*, 34, 1417–1454.

Bista, R. (2015): "Reconciling the WTO Effects on Trade at the Extensive and Intensive Margins," *International Economic Journal*, 29, 231–257.

Buono, I. and G. Lalanne (2012): "The Effect of the Uruguay Round on the Intensive and Extensive Margins of Trade," *Journal of International Economics*, 86, 269–283.

Carnegie, A. (2013): "Instruments of Coercion: International Institutions and the Sites of Power in International Relations," Tech. rep., APSA 2013 Annual Meeting Paper.

——— (2014): "States Held Hostage: Political Hold-Up Problems and the Effects of International Institutions," *American Political Science Review*, 108, 54–70.

Chang, P.-l. and T. Fujii (2012): "Country Image and International Trade," Working Papers 29-2012, Singapore Management University, School of Economics.

Chang, P.-l. and M.-J. Lee (2011): "The WTO trade effect," *Journal of International Economics*, 85, 53–71.

Cheong, J., D. W. Kwak, and K. K. Tang (2014): "The WTO Puzzle, Multilateral Resistance Terms and Multicollinearity," *Applied Economics Letters*, 21, 928–933.

——— (2015): "Can Trade Agreements Curtail Trade Creation and Prevent Trade Diversion?" *Review of International Economics*, 23, 221–238.

Christodoulopoulou, S. (2010): "The Effects of Multilateral Trade Liberalization on the Extensive and the Intensive Margins of Trade," MPRA Paper 29169, University Library of Munich, Germany.

Correia Leite, J. (2008): "The Determinants of Colombian Exports: An Empirical Analysis Using the Gravity Model," *Revista Desarrollo y Sociedad*.

Cristea, A. D. and A. Miromanova (2022): "Firm-level Trade Effects of WTO Accession: Evidence from Russia," *Review of International Economics*, 30, 237–281.

De Santis, R. (2012): "Impact of Environmental Regulations on Trade in the Main EU Countries: Conflict or Synergy?" *The World Economy*, 35, 799–815.

Dutt, P. (2020): "The WTO is Not Passé," *European Economic Review*, 128, 103507.

Dutt, P. and D. Traca (2010): "Corruption and Bilateral Trade Flows: Extortion or Evasion?" *The Review of Economics and Statistics*, 92, 843–860.

Eicher, T. S. and C. Henn (2011): "In Search of WTO Trade Effects: Preferential Trade Agreements Promote Trade Strongly, but Unevenly," *Journal of International Economics*, 83, 137–153.

Engelbrecht, H.-J. and C. Pearce (2007): "The GATT/WTO Has Promoted Trade, But Only in Capital-intensive Commodities!" *Applied Economics*, 39, 1573–1581.

Esteve-Pérez, S., S. Gil-Pareja, and R. Llorca-Vivero (2020): "Does the GATT/WTO Promote Trade? After All, Rose Was Right," *Review of World Economics (Weltwirtschaftliches Archiv)*, 156, 377–405.

Felbermayr, G. and W. Kohler (2010): "Modelling the Extensive Margin of World Trade: New Evidence on GATT and WTO Membership," *The World Economy*, 33, 1430–1469.

Felbermayr, G., M. Larch, E. Yalcin, and Y. V. Yotov (2022): "On the Heterogenous Trade and Welfare Effects of GATT/WTO Membership," FIW Working Paper series 195, FIW.

Felbermayr, G. J. and W. Kohler (2006): "Exploring the Intensive and Extensive Margins of World Trade," *Review of World Economics (Weltwirtschaftliches Archiv)*, 142, 642–674.

Ge, Y., Y. He, Y. Jiang, and X. Yin (2014): "Border Trade and Regional Integration," *Review of Development Economics*, 18, 300–312.

Gil-Pareja, S., R. Llorca-Vivero, and J. A. Martínez-Serrano (2014): "Do Nonreciprocal Preferential Trade Agreements Increase Beneficiaries' Exports?" *Journal of Development Economics*, 107, 291–304.

——— (2016): "A Re-Examination of the Effect of GATT/WTO on Trade," *Open Economies Review*, 27, 561–584.

Goldstein, J. L., D. Rivers, and M. Tomz (2007): "Institutions in International Relations: Understanding the Effects of the GATT and the WTO on World Trade," *International Organization*, 61, 37–67.

Gowa, J. (2010): "Alliances, Market Power, and Postwar Trade: Explaining the GATT/WTO," *World Trade Review*, 9, 487–504.

Gowa, J. and S. Kim (2005): "An Exclusive Country Club: The Effects of the GATT on Trade, 1950–94," *World Politics*, 57, 453–478.

Grant, J. H. and K. A. Boys (2012): "Agricultural Trade and the GATT/WTO: Does Membership Make a Difference?" *American Journal of Agricultural Economics*, 94, 1–24.

Gurevich, T., P. R. Herman, F. Toubal, and Y. V. Yotov (2021): "One Nation, One Language? Domestic Language Diversity, Trade and Welfare," CEPR Discussion Paper 15701, CEPR.

Hanh, P. T. H. (2011): "Does WTO Accession Matter for the Dynamics of Foreign Direct Investment and Trade?" *The Economics of Transition*, 19, 255–285.

Helpman, E., M. Melitz, and Y. Rubinstein (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics*, 123, 441–487.

Herz, B. and M. Wagner (2011): "The 'Real' Impact of GATT/WTO – A Generalised Approach," *The World Economy*, 34, 1014–1041.

Johnson, J. C., M. Souva, and D. L. Smith (2013): "Market-Protecting Institutions and the World Trade Organization's Ability to Promote Trade1," *International Studies Quarterly*, 57, 410–417.

Kim, M. H. (2010): "Does the WTO Promote Trade? Further Evidence," *The Journal of International Trade & Economic Development*, 19, 421–437.

——— (2011): "Do We Really Know That the WTO Increases Trade? Revisited," *Global Economy Journal*, 11, 1–21.

Kim, M. H., D. Kim, and Y. Han (2008): "The Impact of International Institutions on Bilateral Trade: Economic and Political," *Global Economy Journal*, 8, 1–22.

Kuenzel, D. J. (2019): "Do Trade Flows Respond to Nudges? Evidence from the WTO's Trade Policy Review Mechanism," *Review of International Economics*, 27, 735–764.

Larch, M., J.-A. Monteiro, R. Piermartini, and Y. V. Yotov (2019): "On the Effects of GATT/WTO Membership on Trade: They are Positive and Large After All," WTO Staff Working Papers ERSD-2019-09, World Trade Organization (WTO), Economic Research and Statistics Division.

Lee, H., C. M. Koo, and E. Park (2008): "Are Exports of China, Japan and Korea Diverted in the Major Regional Trading Blocs?" *The World Economy*, 31, 841–860.

Lee, H.-H., D. Park, and J. Wang (2013): "Different Types of Firms, Different Types of Products, and Their Dynamics: An Anatomy of China's Imports," *China Economic Review*, 25, 62–77.

Lissovolik, B. and Y. Lissovolik (2006): "Russia and the WTO: The 'Gravity' of Outsider Status," *IMF Staff Papers*, 53, 1–27.

Liu, B., J. Devine, and D. Hudson (2022): "A Gravity Model Analysis of Chinese Apparel Export Flows," Texas Tech University, mimeo.

Liu, X. (2009): "GATT/WTO Promotes Trade Strongly: Sample Selection and Model Specification," *Review of International Economics*, 17, 428–446.

Lohmann, J. (2011): "Do Language Barriers Affect Trade?" *Economics Letters*, 110, 159–162.

Mujahid, I. and M. Kalkuhl (2016): "Do Trade Agreements Increase Food Trade?" *The World Economy*, 39, 1812–1833.

Mölders, F. and U. Volz (2011): "Trade Creation and the Status of FTAs: Empirical Evidence from East Asia," *Review of World Economics (Weltwirtschaftliches Archiv)*, 147, 429–456.

Nitsch, V. (2007): "Does the G7/G8 Promote Trade?" *Economics Letters*, 94, 136–140.

Ornelas, E. and M. Ritel (2020): "The Not-so-generalised Effects of the Generalized System of Preferences," *The World Economy*, 43, 1809–1840.

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

Pusterla, F. (2006): "Regional Integration Agreements: Impact, Geography and Efficiency," Bocconi University, mimeo.

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

——— (2005): "Which International Institutions Promote International Trade?" *Review of International Economics*, 13, 682–698.

——— (2007): "Do We Really Know That the WTO Increases Trade? Reply," *American Economic Review*, 97, 2019–2025.

Roy, D., A. Munasib, and X. Chen (2014): "Social Trust and International Trade: The Interplay Between Social Trust and Formal Finance," *Review of World Economics (Weltwirtschaftliches Archiv)*, 150, 693–714.

Roy, J. (2011): "Is the WTO Mystery Really Solved?" *Economics Letters*, 113, 127–130.

Serrano, R. and V. Pinilla (2012): "The Long-run Decline in the Share of Agricultural and Food Products in International Trade: A Gravity Equation Approach to its Causes," *Applied Economics*, 44, 4199–4210.

——— (2014a): "Changes in the Structure of World Trade in the Agri-Food Industry: The Impact of the Home Market Effect and Regional Liberalization From a Long-Term Perspective, 1963–2010," *Agribusiness*, 30, 165–183.

——— (2014b): "New Directions of Trade for the Agri-food Industry: A Disaggregated Approach for Different Income Countries, 1963–2000," *Latin American Economic Review*, 23, 1–22.

——— (2016): "The Declining Role of Latin America in Global Agricultural Trade, 1963–2000," *Journal of Latin American Studies*, 48, 115–146.

Sheng, Y., H. C. Tang, and X. Xu (2014): "The Impact of the ACFTA on ASEAN–PRC Trade: Estimates Based on an Extended Gravity Model for Component Trade," *Applied Economics*, 46, 2251–2263.

Soukar, L. (2019): "Natural Resources Endowment and WTO," *Journal of Economic Integration*, 34, 546–589.

Subramanian, A. and S.-J. Wei (2007): "The WTO Promotes Trade, Strongly but Unevenly," *Journal of International Economics*, 72, 151–175.

Suvankulov, F. and Y. Guc (2012): "Who is Trading Well in Central Asia? A Gravity Analysis of Exports from the Regional Powers to the Region," *Eurasian Journal of Business and Economics*, 5, 21–43.

Tomz, M., J. L. Goldstein, and D. Rivers (2007): "Do We Really Know That the WTO Increases Trade? Comment," *American Economic Review*, 97, 2005–2018.

Vandenbussche, H. and M. Zanardi (2010): "The chilling trade effects of antidumping proliferation," *European Economic Review*, 54, 760–777.

Wei, S.-J. and Z. Zhang (2007): "Collateral Damage: Exchange Controls and International Trade," *Journal of International Money and Finance*, 26, 841–863.

——— (2010): "Do External Interventions Work? The Case of Trade Reform Conditions in IMF Supported Programs," *Journal of Development Economics*, 92, 71–81.

Zhou, M. and C. Park (2012): "The Cohesion Effect of Structural Equivalence on Global Bilateral Trade, 1948-2000," *International Sociology*, 27, 502–523.

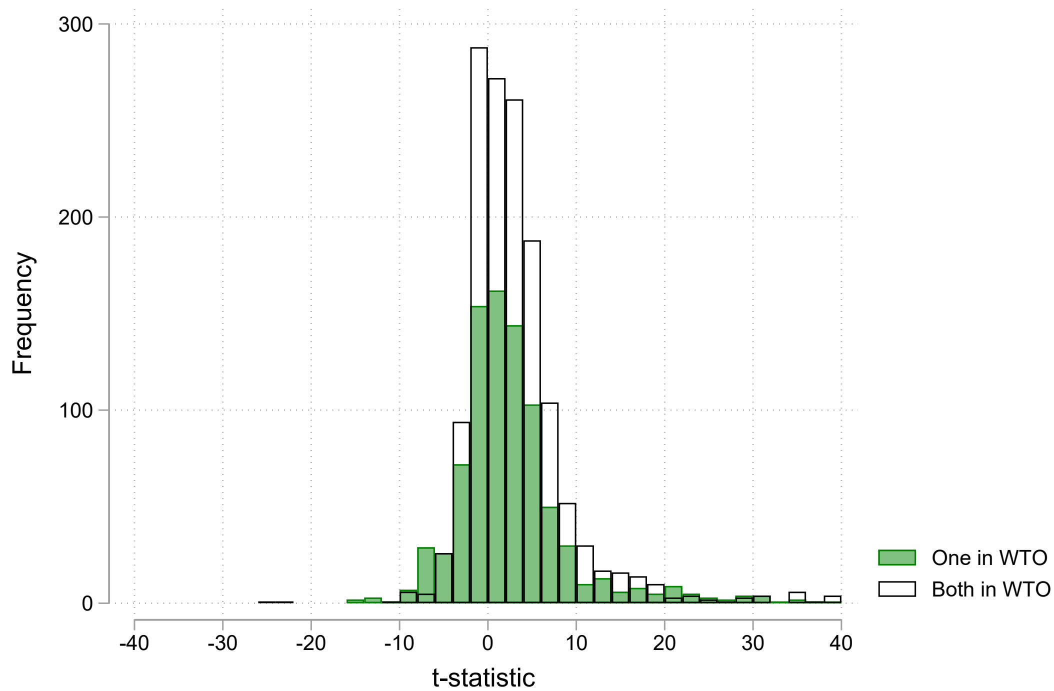


Figure A.1: Histogram of t-statistics

Note: The figure shows the histogram of the t-statistic of the WTO effect on trade in individual studies.

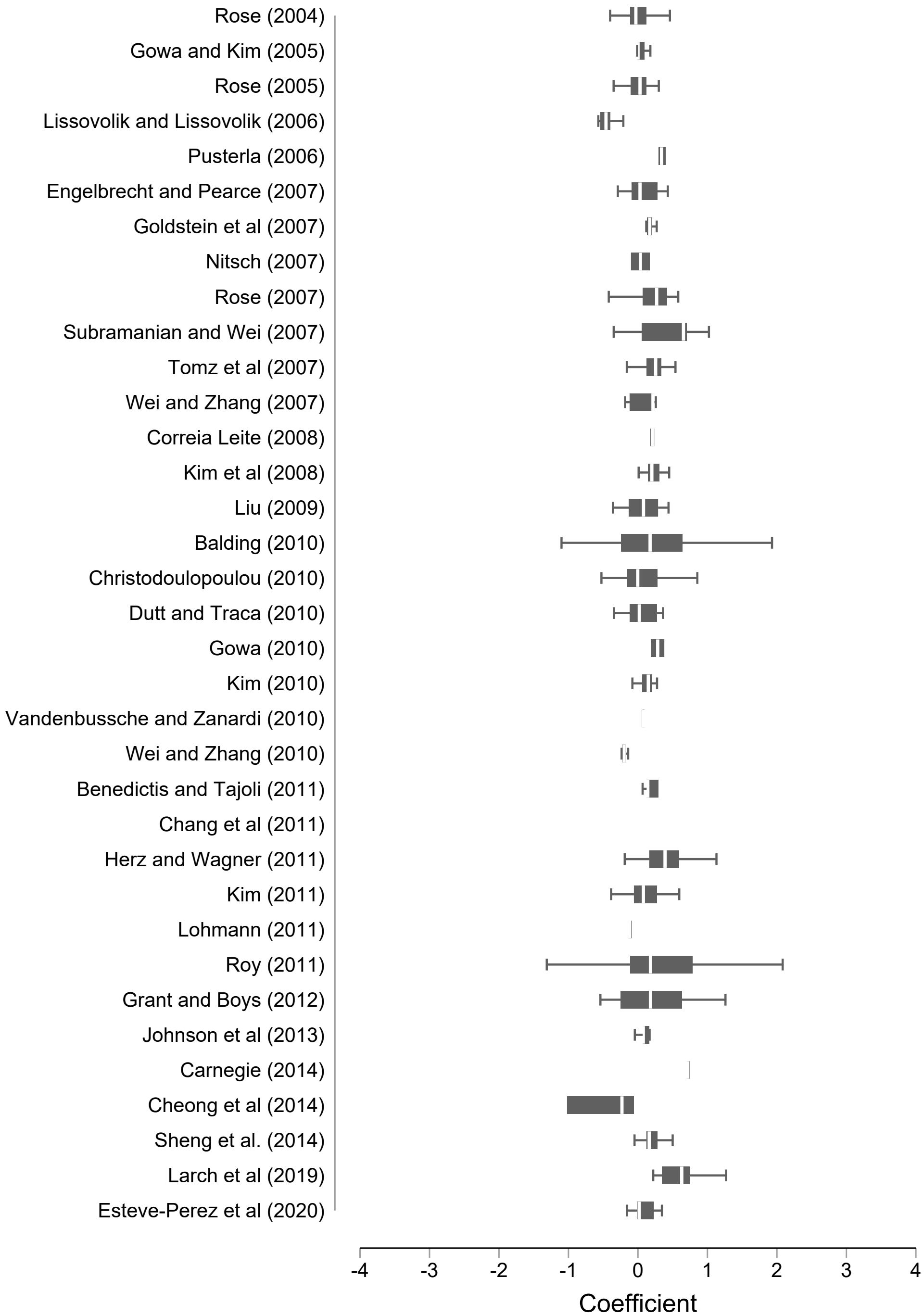


Figure A.2: Box plots for estimates for "One in WTO"

Note: Papers listed in chronological order. Outside values are not shown. Papers with too few observations to calculate percentiles are not removed from the list.

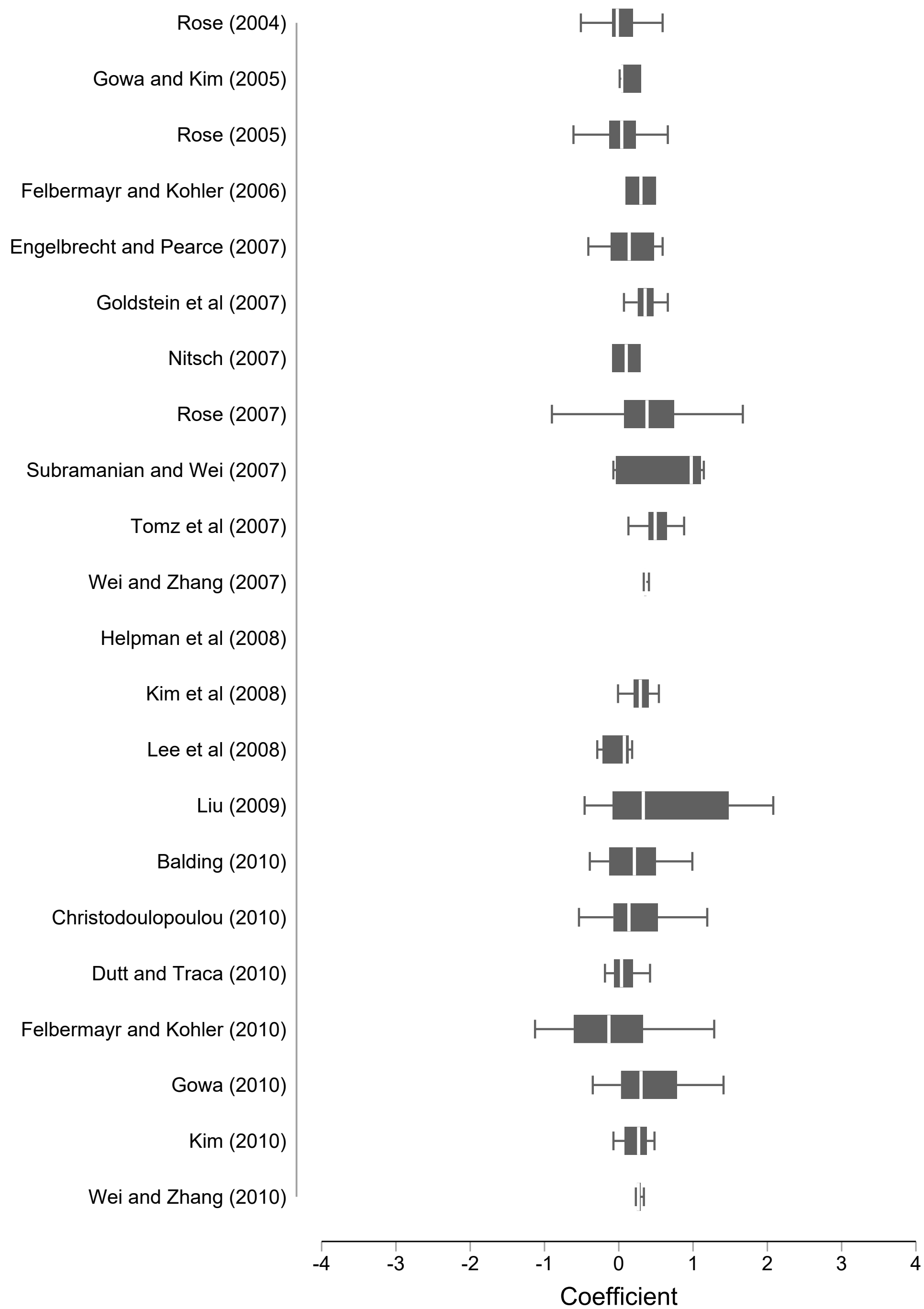


Figure A.3: Box plots for estimates for "Both in WTO" (2004–2010)

Note: Papers listed in chronological order. Outside values are not shown. Papers with too few observations to calculate percentiles are not removed from the list.

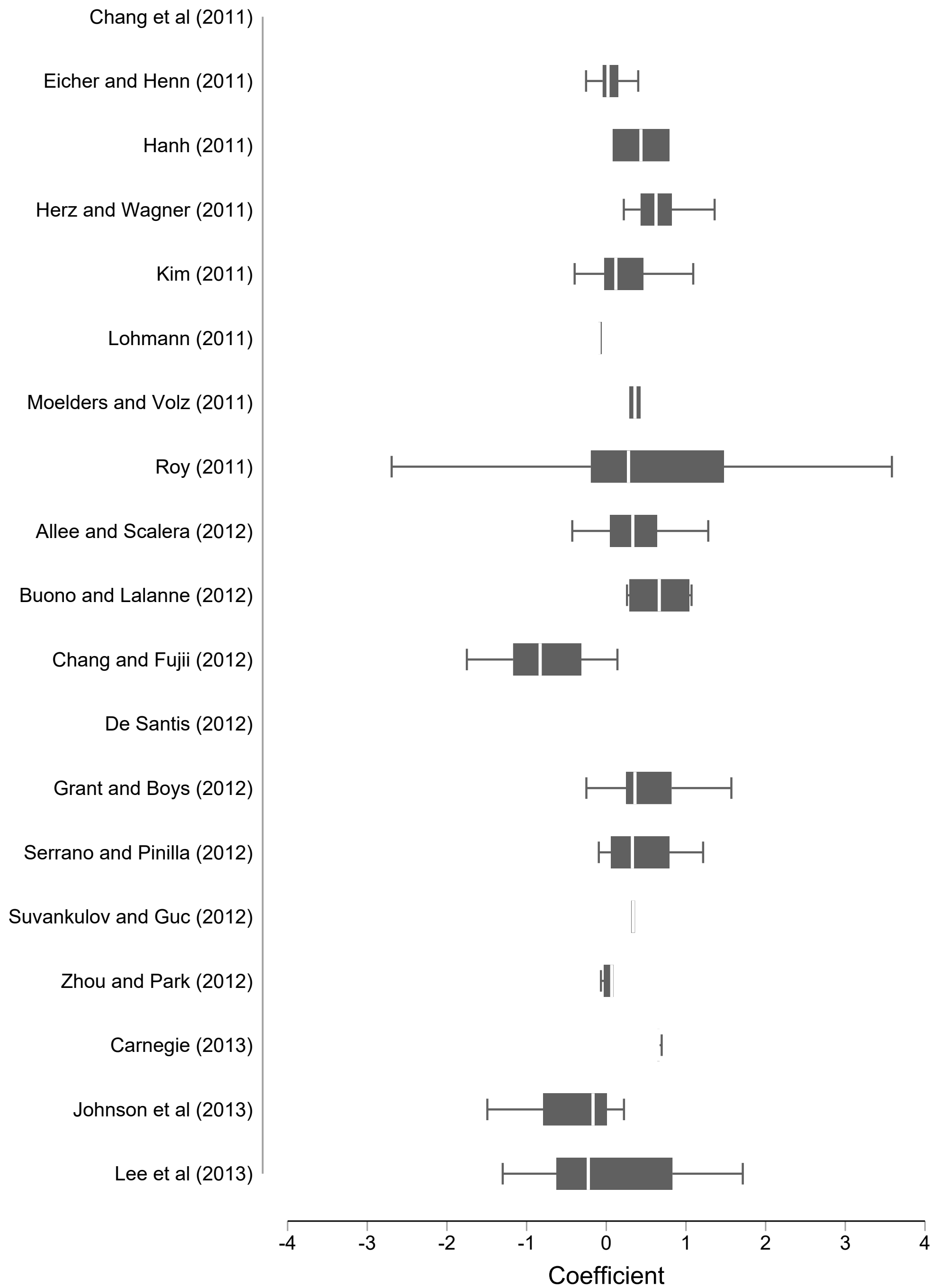


Figure A.4: Box plots for estimates for "Both in WTO" (2010–2013)

Note: Papers listed in chronological order. Outside values are not shown. Papers with too few observations to calculate percentiles are not removed from the list.

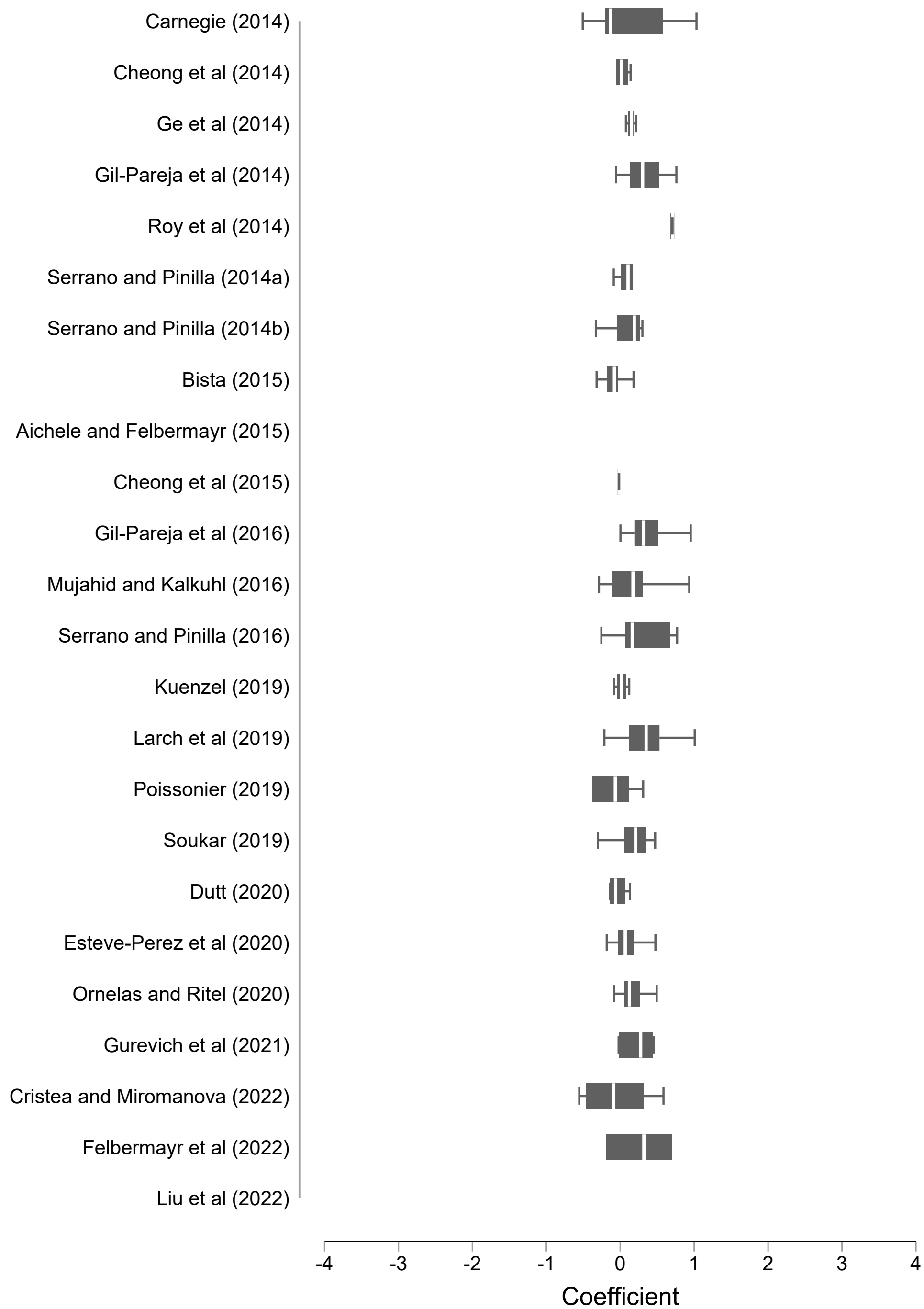


Figure A.5: Box plots for estimates for “Both in WTO” (2014–2022)

Note: Papers listed in chronological order. Outside values are not shown. Papers with too few observations to calculate percentiles are not removed from the list.

B Other trade databases

Table B.1: Different specifications in different databases

	Theory	OLS	Non-theory MRT	No pair FE	No domestic trade
Upper bound					
IMF DOTS (importer)	0.43 (0.06)	0.14 (0.03)	0.39 (0.06)	0.70 (0.06)	-0.05 (0.07)
IMF DOTS (exporter)	0.47 (0.06)	0.14 (0.03)	0.41 (0.05)	0.65 (0.07)	-0.11 (0.07)
UN Comtrade (importer)	0.41 (0.06)	0.11 (0.03)	0.39 (0.06)	0.56 (0.06)	-0.06 (0.07)
UN Comtrade (exporter)	0.40 (0.06)	0.12 (0.04)	0.38 (0.06)	0.52 (0.07)	-0.40 (0.14)
BACI CEPII	0.25 (0.05)	0.11 (0.05)	0.28 (0.04)	0.65 (0.07)	0.08 (0.06)
TradeProd CEPII	0.32 (0.05)	0.18 (0.04)	0.62 (0.15)	0.95 (0.07)	-0.02 (0.08)
WTO Structural DB	0.55 (0.08)	0.13 (0.04)	0.76 (0.24)	1.35 (0.10)	0.02 (0.07)
Lower bound					
IMF DOTS (importer)	0.32 (0.06)	0.16 (0.03)	0.28 (0.05)	0.66 (0.06)	—
IMF DOTS (exporter)	0.34 (0.06)	0.17 (0.03)	0.28 (0.05)	0.60 (0.07)	—
UN Comtrade (importer)	0.32 (0.06)	0.09 (0.03)	0.29 (0.06)	0.55 (0.07)	—
UN Comtrade (exporter)	0.30 (0.06)	0.10 (0.04)	0.28 (0.05)	0.51 (0.07)	—
BACI CEPII	0.20 (0.05)	0.10 (0.05)	0.19 (0.04)	0.65 (0.07)	—
TradeProd CEPII	0.17 (0.05)	0.15 (0.04)	0.41 (0.09)	0.96 (0.07)	—
WTO Structural DB	0.26 (0.06)	0.10 (0.04)	0.61 (0.12)	1.24 (0.09)	—

Table B.2: Robustness in different databases

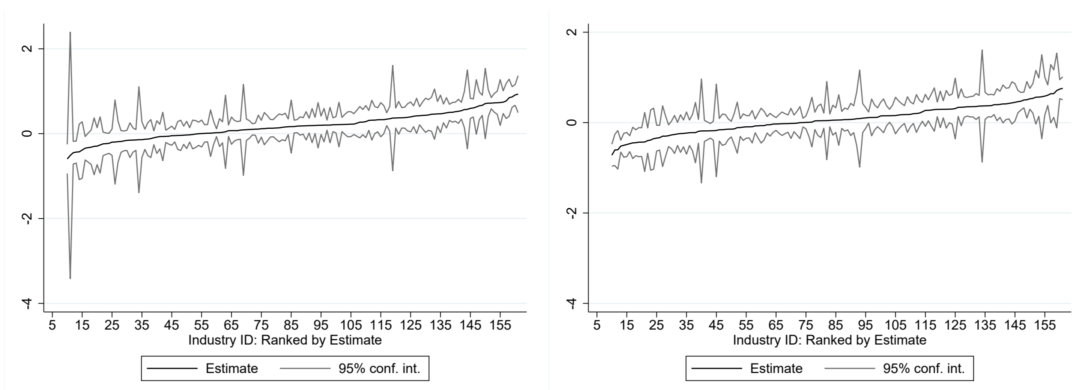
	Theory-consistent	Dist x year	3-way clustering
Upper bound			
IMF DOTS (importer)	0.43 (0.06)	0.36 (0.05)	0.43 (0.12)
IMF DOTS (exporter)	0.47 (0.06)	0.37 (0.05)	0.47 (0.13)
UN Comtrade (importer)	0.41 (0.06)	0.34 (0.05)	0.41 (0.11)
UN Comtrade (exporter)	0.40 (0.06)	0.32 (0.06)	0.40 (0.13)
BACI CEPII	0.25 (0.05)	0.20 (0.05)	0.25 (0.13)
TradeProd CEPII	0.32 (0.05)	0.19 (0.04)	0.32 (0.07)
WTO Structural DB	0.55 (0.08)	0.34 (0.06)	0.55 (0.13)
Lower bound			
IMF DOTS (importer)	0.32 (0.06)	0.34 (0.05)	0.32 (0.11)
IMF DOTS (exporter)	0.34 (0.06)	0.35 (0.05)	0.34 (0.12)
UN Comtrade (importer)	0.32 (0.06)	0.33 (0.05)	0.32 (0.10)
UN Comtrade (exporter)	0.30 (0.06)	0.30 (0.06)	0.30 (0.12)
BACI CEPII	0.20 (0.05)	0.20 (0.05)	0.20 (0.11)
TradeProd CEPII	0.17 (0.05)	0.17 (0.04)	0.17 (0.06)
WTO Structural DB	0.26 (0.06)	0.26 (0.05)	0.26 (0.10)

Table B.3: Level of development in different databases

	AEs-AEs	AEs-EMDEs	EMDEs-AEs	EMDEs-EMDEs
Upper bound				
IMF DOTS (importer)	0.08 (0.13)	0.42 (0.09)	0.38 (0.08)	0.60 (0.09)
IMF DOTS (exporter)	0.33 (0.10)	0.45 (0.09)	0.45 (0.09)	0.55 (0.08)
UN Comtrade (importer)	-0.07 (0.13)	0.37 (0.10)	0.35 (0.08)	0.57 (0.09)
UN Comtrade (exporter)	0.24 (0.11)	0.44 (0.09)	0.33 (0.10)	0.50 (0.09)
BACI CEPII	0.04 (0.09)	0.27 (0.08)	0.22 (0.06)	0.31 (0.08)
TradeProd CEPII	0.59 (0.18)	0.37 (0.06)	0.18 (0.10)	0.42 (0.08)
WTO Structural DB	0.37 (0.08)	0.62 (0.08)	0.49 (0.12)	0.55 (0.09)
Lower bound				
IMF DOTS (importer)	-0.07 (0.13)	0.28 (0.09)	0.27 (0.08)	0.52 (0.09)
IMF DOTS (exporter)	0.18 (0.11)	0.30 (0.08)	0.32 (0.08)	0.44 (0.08)
UN Comtrade (importer)	-0.20 (0.13)	0.27 (0.09)	0.27 (0.08)	0.52 (0.09)
UN Comtrade (exporter)	0.12 (0.12)	0.35 (0.08)	0.24 (0.09)	0.44 (0.08)
BACI CEPII	-0.03 (0.09)	0.21 (0.07)	0.18 (0.06)	0.29 (0.08)
TradeProd CEPII	0.32 (0.14)	0.20 (0.06)	0.02 (0.09)	0.33 (0.07)
WTO Structural DB	0.08 (0.06)	0.30 (0.07)	0.18 (0.09)	0.33 (0.08)

Table B.4: GATT and WTO estimated separately

	Upper bound		Lower bound	
	GATT	WTO	GATT	WTO
IMF DOTS (importer)	0.27 (0.12)	0.27 (0.03)	0.28 (0.13)	0.22 (0.04)
IMF DOTS (exporter)	0.26 (0.13)	0.31 (0.04)	0.26 (0.14)	0.20 (0.04)
UN Comtrade (importer)	0.30 (0.13)	0.25 (0.03)	0.34 (0.14)	0.28 (0.04)
UN Comtrade (exporter)	0.34 (0.15)	0.27 (0.03)	0.38 (0.16)	0.25 (0.04)
BACI CEPII	—	0.25 (0.05)	—	0.20 (0.05)
TradeProd CEPII	0.47 (0.08)	0.40 (0.03)	0.35 (0.07)	0.12 (0.04)
WTO Structural DB	0.15 (0.09)	0.62 (0.03)	0.03 (0.09)	0.12 (0.04)



(a) Upper bound

(b) Lower bound

Figure B.1: Results by industry

Note: The figures show the distribution of point estimates (black lines) and 95% confidence intervals (grey lines), for industries in agriculture, mining, manufacturing, and services. Standard errors are clustered by country pair. Estimations use data from the ITPD-E database. The estimations on the left do not control for border-year dummies ("upper- bound"). The estimations on the right control for border-year dummies ("lower-bound").

C Robustness checks

Table C.1: Tests for publication bias

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Standard error	-0.106	-0.0330	-0.0391	-0.251	-0.245	-0.339
	(0.255)	(0.237)	(0.301)	(0.204)	(0.211)	(0.391)
Constant	0.160***	0.172***	0.151***	0.249***	0.247***	0.260***
	(0.0455)	(0.0451)	(0.0518)	(0.0402)	(0.0409)	(0.0619)
Observations	862	807	862	1,422	1,381	1,422
R-squared	0.002	0.000	0.000	0.007	0.007	0.009
Number of id			34			62

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: The dependent variable in columns (1)–(3) is the estimate winsorized at 1% for "*Onein*". Column (1) includes all estimates of this type. Column (2) only those that were published in a journal. Column (3) includes fixed effects by study. The value reported for the constant is the average of all fixed effects. The dependent variable in columns (4)–(6) is the estimate winsorized at 1% for "*Bothin*". Column (4) includes all estimates of this type. Column (5) only those that were published in a journal. Column (6) includes fixed effects by study. The value reported for the constant is the average of all fixed effects. Standard errors (in parentheses) are clustered at the study level and bootstrapped with 10,000 replications.

Table C.2: Tests for publication bias with sub-samples

VARIABLES	(1)	(2)	(3)	(4)
Standard error	-0.626	0.00832	-0.192	-0.145
	(0.842)	(0.372)	(0.476)	(0.360)
Constant	0.435	0.0649	0.134**	0.192***
	(0.270)	(0.0502)	(0.0641)	(0.0518)
Observations	33	527	317	629
R-squared	0.053	0.000	0.012	0.001

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable in columns (1) and (2) is the estimate for "*Onein*". The dependent variable in columns (3) and (4) is the estimate for "*Bothin*". Columns (1) and (3) are based on those studies that include theoretical multilateral resistances and pair fixed effects. Columns (2) and (4) are based on those studies that include neither theoretical multilateral resistances nor pair fixed effects. Standard errors (in parentheses) are clustered at the study level and bootstrapped with 10,000 replications.

Table C.3: Nonlinear model: WAAP

VARIABLES	(1)	(2)	(3)
precision	0.317***	0.310***	0.332***
	(0.00712)	(0.00734)	(0.0141)
Observations	1,520	595	934
R-squared	0.565	0.750	0.372

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: WAAP is computed as an unrestricted weighted average of the adequately powered. This model is based on the funnel plot. It discards estimates with retrospective power below 80% and computes an inverse-variance-weighted mean of the remaining estimates. This is used to reduce the publication bias in meta-analysis. The dependent variable in column (1) is the estimate for "*All*". In column (2) is for *Onein*. Column (3) for *Bothin*

Table C.4: Selection model: Andrews-Kasy

	μ	τ	$(-\infty, -1.96)$	$(-1.96, 0]$	$[0, 1.96)$
Estimate	0.049	0.448	0.312	0.894	0.839
Standard error	0.025	0.016	0.047	0.085	0.067

Note: The selection model due to Andrews and Kasy (2019) where P2 and P3 denote the probability that estimates insignificant at the 5% level are published relative to the probability that significant estimates are published. Standard errors are reported in parentheses.

Table C.5: MAIVE: Meta-Analysis Instrumental Variable Estimator

MAIVE coefficient	0.082
MAIVE standard error	0.238
F-test of first step in IV	0.474
Hausman-type test	0.966
Critical Value of Chi2(1)	3.841
AR Confidence interval	$[-0.781, 0.999]$

Note: The dependent variable is the estimate for All WTO effect. MAIVE takes the inverse of the sample size of primary studies as an instrument for reported squared standard errors. Standard errors in parentheses are clustered at the study level. The Hausman test statistic consists of a weighted squared difference between the MAIVE and a standard point estimate, which is the same method as MAIVE but without instrumenting the standard errors and including inverse-variance weights.

Table C.6: Estimation characteristics (*Bothin*)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Standard error	-0.251	-0.229	-0.235	-0.236	-0.251	-0.241	-0.192
	(0.204)	(0.206)	(0.199)	(0.206)	(0.204)	(0.196)	(0.219)
No or atheorical MTRs		0.0747					0.0538
		(0.0764)					(0.0721)
Other than PPML			0.236***				0.240***
			(0.0658)				(0.0689)
No pair fe				-0.0629			-0.0955**
				(0.0548)			(0.0459)
No domestic trade					-0.168***		-0.350***
					(0.0519)		(0.0716)
One advanced						-0.0996	-0.00808
						(0.109)	(0.0779)
Both advanced						0.0661	0.0578
						(0.106)	(0.0889)
Constant	0.249***	0.197***	0.0707	0.282***	0.414***	0.251***	0.420***
	(0.0402)	(0.0696)	(0.0624)	(0.0514)	(0.0419)	(0.0373)	(0.0549)
Observations	1,422	1,422	1,422	1,422	1,422	1,422	1,422
R-squared	0.007	0.013	0.057	0.012	0.009	0.013	0.075

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable is the estimate winsorized at 1% for "*Bothin*". Standard errors (in parentheses) are clustered at the study level and bootstrapped with 10,000 replications.

Table C.7: Estimation characteristics (*Onein*)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Standard error	-0.106	-0.205	-0.0989	-0.0761	-0.112	-0.188	-0.197
	(0.255)	(0.322)	(0.254)	(0.277)	(0.260)	(0.227)	(0.268)
No or atheorical MTRs		-0.201					-0.0670
		(0.178)					(0.248)
Other than PPML			-0.197				-0.0297
			(0.149)				(0.105)
No pair fe				-0.157**			-0.198***
				(0.0705)			(0.0616)
No domestic trade					-0.482***		-0.275
					(0.0490)		(0.284)
One advanced						0.353**	0.409*
						(0.173)	(0.222)
Both advanced						0.271**	0.294**
						(0.116)	(0.124)
Constant	0.160***	0.356**	0.348**	0.261***	0.636***	0.123***	0.609***
	(0.0455)	(0.169)	(0.143)	(0.0438)	(0.0415)	(0.0445)	(0.0489)
Observations	862	862	862	862	862	862	862
R-squared	0.002	0.016	0.008	0.027	0.015	0.063	0.118

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: The dependent variable is the estimate winsorized at 1% for "*Onein*". Standard errors (in parentheses) are clustered at the study level and bootstrapped with 10,000 replications.