WIFO ■ WORKING PAPERS
703/2025

Everything You Always Wanted to Know About EU Membership Trade Effects but Were Afraid to Ask

Harald Oberhofer
Zhenyi Wang

Everything You Always Wanted to Know About EU Membership Trade Effects but Were Afraid to Ask

Harald Oberhofer (WIFO), Zhenyi Wang (WU Vienna)

WIFO Working Papers 703/2025
April 2025

Abstract

This paper studies the role of dataset choices for trade policy estimates in structural gravity models. Using twelve publicly available data sources and applying alternative and commonly used dataset restrictions, we obtain 586 estimates for the EU membership trade effects from a standard structural gravity model specification. These estimates are used in a meta-regression analysis to shed light on potential sources for heterogeneity in the obtained EU membership trade effects. The meta study reveals a crucial role of domestic trade flows for the effect size of the EU membership trade effect estimate. The estimated EU trade effect is on average 20 percentage points larger in datasets that include domestic trade flows. The effect size of the EU membership trade effect estimates also vary with the time- and country coverage, across time interval and consecutive year panel data, alternative product classifications, the level of sectoral disaggregation in the data and the use of imports as alternative measure for trade flow. Alternative estimation packages do not significantly alter the effect size of our estimates for the EU membership trade effects. The paper concludes with some recommendations on data source selection and dataset restrictions.

E-Mail: harald.oberhofer@wifo.ac.at

2025/1/W/0

Media owner (publisher), producer: Austrian Institute of Economic Research
1030 Vienna, Arsenal, Objekt 20 | Tel. (43 1) 798 26 01 0 | https://www.wifo.ac.at
Place of publishing and production: Vienna

Free download: https://www.wifo.ac.at/publication/pid/58075956

Everything You Always Wanted to Know About EU Membership Trade Effects But Were Afraid to Ask*

Harald Oberhofer† Zhenyi Wang‡

Abstract

This paper studies the role of dataset choices for trade policy estimates in structural gravity models. Using twelve publicly available data sources and applying alternative and commonly used dataset restrictions, we obtain 586 estimates for the EU membership trade effects from a standard structural gravity model specification. These estimates are used in a meta-regression analysis to shed light on potential sources for heterogeneity in the obtained EU membership trade effects. The meta study reveals a crucial role of domestic trade flows for the effect size of the EU membership trade effect estimate. The estimated EU trade effect is on average 20 percentage points larger in datasets that include domestic trade flows. The effect size of the EU membership trade effect estimates also vary with the time- and country coverage, across time interval and consecutive year panel data, alternative product classifications, the level of sectoral disaggregation in the data and the use of imports as alternative measure for trade flow. Alternative estimation packages do not significantly alter the effect size of our estimates for the EU membership trade effects. The paper concludes with some recommendations on data source selection and dataset restrictions.

JEL-Codes: F13, F14, F15, C80.

Keywords: Trade policy, EU membership trade effects, Gravity models, meta study, meta-regression analysis.

*We are thankful to Jesus Crespo Cuaresma, Peter Egger, Gabriel Felbermayr, Mario Larch, Michael Pfaffermayr, João Santos Silva, Yoto Yotov and participants at research seminars at the Vienna University of Economics and Business, Universities of Salzburg and Würzburg, the FIW research retreat and annual conferences of the European Trade Study Group (2023) and the Austrian Economic Assoication (2023) for very helpful and constructive comments and suggestions on previous versions of the manuscript.

†Department of Economics, Vienna University of Economics and Business (WU), Welthandelsplatz 1, 1020 Vienna, Austria; Austrian Institute of Economic Research (WIFO); CESifo Munich. Email: harald.oberhofer@wu.ac.at.

‡Department of Economics, Vienna University of Economics and Business (WU), Welthandelsplatz 1, 1020 Vienna, Austria. Email: zhenyi.wang@wu.ac.at.

1 Introduction

Over the last decades various multilateral, bilateral and even unilateral trade policy measures have substantially contribute to the shaping of the world economy. The formation of the World Trade Organization (WTO) in 1995, China's accession to it in 2001, the establishment, enlargement and deepening of the European Common Market and the increasing number of bilateral free trade agreements have attracted a lot of interest among economic policy makers, the broader public, and economists working in the field of International Economics. An important strand of the academic literature studies the trade and welfare implications of alternative trade policy measures, thereby providing important scientific contributions to the theoretical modeling and econometric estimation of trade policy effects.

The seminal contributions of Anderson (1979) and Anderson and Van Wincoop (2003) emphasize the need for theory-consistent estimation of structural gravity models and call for a close connection between the econometric identification of trade policy effects and the calculation of general equilibrium trade and welfare effects in quantitative trade models. Allen et al. (2020) show that a large class of international trade theories result in similar empirical specifications of the structural gravity model. Due to an increasing use of a "dual gravity" approach, the partial equilibrium (direct) trade policy effect estimates from econometric specifications of structural gravity model typically serve as inputs for new quantitative multi-country, multi-sector general equilibrium trade models in the spirit of e.g., Eaton and Kortum (2002) or Caliendo and Parro (2015).[1]

The scientific claim to closely link international trade theories with the empirical estimation of the parameters of interest lead to a series of methodological and empirical contributions on how to best estimate trade effects in structural gravity models.[2] This literature discusses potential reasons for differences in the obtained estimates and provides solutions to overcome previous mistakes and methodological limitations. Important contributions in this strand paid particular attention to e.g., the econometric and empirical specification of the gravity model (Fally, 2015; Bergstrand et al., 2015; Larch et al., 2018), the applied estimators (Santos Silva and Tenreyro, 2006; Egger and Staub, 2016; Pfaffermayr, 2020; Weidner and Zylkin, 2021; Larch and Yotov, 2024; Bergstrand et al., 2025) and the sample coverage (Glick and Rose, 2016).

The selection of a particular dataset for a specific trade policy question at hand has not yet received similar interest in the literature and is mostly treated as an innocent and less important decision to be made.[3]. This paper seeks to challenge this view by examining whether the choice of a particular data source, along with commonly imposed dataset restrictions, significantly influences the estimated trade policy effects in structural gravity models. We aim at exploring the sensitivity in the estimated effect size of a trade policy across different data sources applying a standard three-way fixed effects structural gravity model specification. We apply the standard Poisson pseudo-maximum-likelihood (PPML) estimator from Santos Silva and Tenreyro (2006) and its bias-correct version from Weidner and Zylkin (2021) across all different data sources and and alternative

[1] Felbermayr et al. (2022), for example, offer a recent application of the the dual gravity approach to quantify the trade and welfare effects of a hypothetical undoing of European goods and services markets integration.

[2] Overviews on the advancements in this literature and recommendations for empirical work are provided by e.g., Head and Mayer (2014), Yotov (2012), Larch and Yotov (2024) and Larch et al. (2025).

[3] To the best of our knowledge, Egger and Wolfmayr (2018) is the only contribution that provides a discussion of the conceptual principles of trade data collection, examines different data cleaning approaches applied to nine alternative trade data sources, and offers a descriptive comparison of the resulting trade data

samples.[4]

We collect trade data from twelve different publicly accessible data sources that are frequently used for trade policy analysis using structural gravity models. These datasets vary with respect to the time and country coverage and the level of disaggregation. Five data sources report aggregated bilateral trade flow data for country-pairs, while seven data sources additionally or exclusively provide trade data at the industry- or product-level. Nine data sources include domestic trade flows, while three particularly popular data sources only contain cross-country bilateral trade flows. In the former nine data sources, domestic trade is compiled as the difference between total domestic production and total exports to all foreign trading partners (Campos et al., 2021). Three datasets are based on trade data inferred from input-output tables. Nine data sources rely on bilateral trade data provided by national statistical offices.

For each of the twelve data sources we construct up to nine data samples by imposing a set of plausible and commonly applied dataset restrictions and modifications. These restrictions concern the country and time coverage, the level of data aggregation, alternative available industry classifications, the usage of consecutive annual panel data versus interval data, the replacement of exports with imports, the limitation of sectoral coverage to manufacturing industries, and the availability of domestic trade flows. The alternative data sources and sample restrictions together with the application of two popular alternative estimation packages available in Stata (*ppmlhdfe* and *ppml_fe_bias*) enable us to estimate 586 EU membership trade effects. These estimates serve as the dependent variable in a subsequent meta-regression analysis that empirically studies the impacts of data source choices and sample restrictions on the effect size and statistical significance of the estimated EU membership trade effects.

A descriptive analysis of the 586 estimates reveals some first insights into the role of dataset choices for the effect size of trade policy estimates in structural gravity models. The average estimates across the twelve different data sources differ in economically non-negligible magnitudes. In datasets with domestic trade flows, the EU membership trade effects are statistically highly statistically significant and economically sizable ranging from 10% to around 71%. The variation in the effect size, however, also indicates that other dataset restrictions matter quantitatively. Estimates from samples including only bilateral manufacturing trade flows are mostly close to zero and usually statistically not different from zero. Two data sources that only include bilateral total trade at the country-level without further industry-level disaggregation, by contrast, reveal sizable positive EU membership trade effects.

The meta-regression analysis sheds a more systematic light on the impact of data source characteristics for the effect size of EU membership trade effect estimates. Unlike in standard meta-analyses known in the literature, we obtain all estimates using the same specification and various data sources by ourselves. For the different data set characteristics and restrictions we define dummy variables that are used as covariates. The obtained parameter estimates from the meta-regression can be interpreted as average differences in the effect size of the estimated EU membership trade effects holding other data source characteristics and sample restrictions

[4] Hou (2020) studies the sensitivity of the OLS and the PPML estimators across two different and popular data sources, the IMF DOTS and UN Comtrade, and alternative sample restrictions. She finds that the OLS estimates are more sensitive to data source choices than the PPML estimator.

fixed.

The main findings from the meta-regression analysis are: The inclusion of domestic trade flows substantially increases the effect size of EU membership trade effect estimates in standard structural gravity model specifications. This finding is robust across alternative meta-regression estimators. In datasets with domestic trade flows, the estimated EU membership trade effect is on average 20 percentage points larger. In addition, restricting the country coverage and using only trade data for manufacturing reduces the effect size of the estimated EU trade effects. Applying interval data instead of an annual panel data, exploiting disaggregated industry-level data, and replacing exports by imports as the trade flow variable of interest leads to larger quantitative estimates for the EU membership trade effect. The application of the two alternative Stata packages does not result in a systematically different effect size of the estimates for the EU membership trade effect. The impact of using input-output data versus data based on official data from nation statistical offices remains inconclusive across alternative meta-regression estimators.

The remainder of the paper is organized as follows. Section 2 describes the twelve databases in more detail. By discussing some properties of the alternative data sources, we show the necessity to analyze the effect of choosing one over the other on the effect size of the EU membership trade effect. Section 3 discusses the standard empirical specifications of the gravity model from which we obtain the estimates that will serve es outcome of interest in the meta-regression analysis. Section 4.1 descriptively summarizes our 586 estimates for the EU membership trade effect. Section 4.2 introduces the meta-regression analysis. Section 4.3 presents the results of the meta-regression analysis and discusses further findings for differences in the effect size of the EU membership trade effect effects. We conclude in Section 5 and provide recommendations on data source selection and data restrictions for empirical research on trade policy effects in structural gravity models.

2 Data description and descriptive statistics

2.1 Trade data sources

Many international organizations and research collaborations collect and publish official trade data independently from each other, supporting the rising interest in data- and evidence-based trade policy debates. These trade data sources differ in key aspects such as the compilation methods, time and country coverage, and the level of sectoral disaggregation. As a result, trade economists enjoy the convenience of selecting the most suitable data source for their specific research question. At the same time, the choice of different data sources for similar research questions might hinder direct comparability of the obtained findings across scientific contributions. Moreover, the impact of selecting a specific data source on the estimated effect size of trade policy variables within structural gravity models has not yet been systematically investigated in the international trade literature.

This paper aims to fill this gap by estimating a standard structural econometric gravity model using various data sources and dataset restrictions. In a meta-study, we then regress the estimated EU-membership trade effects on data source characteristics, including dataset restrictions typically applied in the literature. This approach allows us to systematically examine the impact of a researcher's choices on the quantitative magnitude of the trade effect estimates of interest.

The paper uses twelve of the most popular sources for trade data, including: the Comtrade Database compiled by the United Nations; IMF's Direction of Trade Statistics Database (DOTS); CEPII's BACI Database; two versions of the Trade and Production Database (TradeProd 2012 and 2023); the International Trade and Production Database for Estimation (ITPD-E); a dataset collected by Oberhofer and Pfaffermayr (2021) (OP (2021)); WTO's Structural Gravity Manufacturing Database (SGMD); Robert Feenstra's World Trade Flows Database (WTF); the Eora26 database; OECD's Trade in Value Added Database (TiVA); and the World Input-Output Database (WIOD). In the following, we describe and compare the twelve different data sources. Among these, five are most frequently used in the trade policy evaluation literature. Three other ones are based on input-output tables, and four have only recently become available.

The UN Comtrade and IMF's DOTS are two of the most commonly used and largely comparable data sources for bilateral trade data. Both rely on official trade data reported by national statistical offices and supplement it with information from various international organizations. UN Comtrade is used as an official source for updating DOTS, which provides supplementary data accounting for 6% of world imports and 2.3% of world exports in 2015 (Marini et al., 2018). Both data sources provide estimates for missing trade flows, albeit using different methods, and report both import and export data. The main difference is that DOTS reports data that are aggregated at the bilateral country level, while Comtrade offers data at the disaggregated product level. Data from DOTS is available for the years from 1970 to 2020, and from Comtrade for the period 1988 to 2021, although the last two years need to be excluded from the analysis due to a substantial amount of missing data. The number of countries covered by these two data sources ranges from 183 to 246, depending on the bilateral trade flow variable used.

BACI and TradeProd are two other widely used databases using officially reported bilateral trade data, both compiled by CEPII and described in detail in Gaulier and Zignago (2010), De Sousa et al. (2012), and Mayer et al. (2023). BACI provides bilateral trade flow data at the 6-digit HS product level, using the UN Comtrade database as source. It uses both import and export data to create a single measure of trade flows. The BACI database covers the period from 1996 to 2021, but due to missing observations in Comtrade, we again need to exclude the last two years from our analysis. We use two versions of the TradeProd database. The 2023 version expands the time coverage from 1980-2006 to 1966-2018, and aggregates the 26 manufacturing sectors from the 2012 version to bilateral trade flows for nine broad sectors based on the (2-digit) ISIC industry classification. Both versions of TradeProd calculate domestic trade flows using production data from the World Bank, OECD, and UNIDO.

Eora26[5], OECD's TiVA[6], and WIOD[7] are based on national and international input-output tables that can be used to calculate international and domestic trade flows. Eora26 aggregates the MRIO database into 26

[5]Lenzen et al. (2013) discusses the methodology used to construct Eora26 from the Global Multi-Region Input–Output (MRIO) database.

[6]Guilhoto et al. (2022) provides a detailed description of TiVA and the available data.

[7]Timmer et al. (2015) offers a detailed description of the data sources used and provides an illustrative example on how to use WIOD.

sectors. We derive bilateral trade flows from Eora26 for 190 countries spanning the years from 1990 to 2016. The 2021 edition of the TiVA database provides trade data for 66 economies, including all OECD, EU, and G20 countries, as well as several East and Southeast Asian and South American countries. The WIOD 2016 release offers trade data for 56 industries across 43 countries from 2000 to 2014, while the 2013 version covers 35 industries and 40 countries from 1995 to 2011. To maximize the time coverage in WIOD, we combine both versions using concordance tables, resulting in a somewhat unbalanced dataset that includes 13 manufacturing industries for 43 countries from 1995 to 2014.[8]

The remaining four databases have been developed recently and provide information on both domestic and international trade flows. ITPD-E, as discussed in Borchert et al. (2021), includes data on bilateral and domestic trade flows for 120 manufacturing industries across 243 countries from 2000 to 2016. This database relies solely on officially reported data, excluding estimates, and is thus particularly useful for the econometric estimation of structural gravity models. The SGMD database, introduced by Monteiro (2020), covers bilateral and domestic trade flows for manufacturing goods, spanning 186 countries over a 37-year period from 1980 to 2016. It utilizes free on board (FOB) export data from UN Comtrade, complemented by mirrored import data adjusted for cost, insurance, and freight (CIF).

The WTF database converts product-level data from UN Comtrade to the SITC Rev.2 industry classification and provides aggregated data for the broad sectors of manufacturing, mining (including oil), and agriculture.[9] To minimize potential measurement errors, WTF excludes observations where the CIF/FOB ratio is less than 0.1 or greater than 10, and the CIF value is below 50,000 USD. Domestic trade flows are derived from bilateral trade data by subtracting total exports from current GDP (see also Yotov, 2012). GDP values are taken from the World Bank's World Development Indicators (WDI) and total exports from the IMF's DOTS database. Lastly, Oberhofer and Pfaffermayr (2021) provide a unique dataset covering 65 economies with data at three-year intervals from 1994 to 2012. This dataset integrates trade data from the OECD's STAN database, Nicita and Olarreaga (2007)'s database with additional data on gross production, total exports, and total imports from OECD–STAN, UNIDO, CEPII, and WIOD. Missing export data are substituted with mirrored imports, and domestic trade flows are derived as the difference between gross production and total exports.

A key distinction among the twelve data sources is the level disaggregation. The BACI database provides detailed trade data at the product level, whereas other Comtrade-based sources, including Comtrade itself, offer data at alternative industry levels. In line with approaches used by other sources such as WTF, we convert BACI's product-level data to the 2-digit industry level according to SITC Rev.3. As a result, out of the twelve data sources, seven provide trade flows at a disaggregated industry level, while the remaining five offer aggregated trade flow data either at the country level or for broad sectors, such as agriculture, manufacturing, and services. For large parts of this paper, we will focus on data for manufacturing industries. The aggregation of more detailed 4-digit industry-level data to 2-digit levels is comparatively simple for this sector.

Additionally, since most studies on the EU membership trade effects focus exclusively on manufacturing industries, our approach allows for more meaningful comparisons. Many services are either non-tradable or

[8]For a detailed account of the integration of the two WIOD waves, see Wolfmayr et al. (2019).

[9]WTF is part of a larger initiative to develop quality-adjusted trade prices as discussed in Feenstra and Romalis (2014).

rely on alternative trade modes, such as the foreign presence of service providers. Agricultural markets are subject to stronger government intervention. As a result, the estimated trade effects of EU membership may differ significantly for these sectors. This would complicate a fair comparison of estimates across different data sources, as not all of them include data on agriculture and services trade. In Section 4.3.4, we address the issue of heterogeneous EU membership trade effect estimates across broad sectors and provide some evidence on this matter.

2.2 Some descriptive comparisons

Table 1 summarizes the main characteristics of the twelve data sources, including country and time coverage, domestic trade flow coverage, sectoral disaggregation and the number of observations. We also add simple summary statistics for the trade flows including sample averages together with minimum and maximum values.

There is notable variation in the time and country coverage across the twelve data sources.[10] ITPD-E, TradeProd (2012), and WTF offer the widest country coverage, including more than 230 countries. They are followed by BACI and SGMD, which include over 220 countries. Comtrade contains more than 200 countries, but accounts for export and import flows asymmetrically across countries. OP (2021), TiVA, and WIOD collect data only for the most important economies in Europe, Northern America, and Eastern Asia. Eora26, DOTS, SGMD, TradeProd (2023) and Comtrade comprise the longest time span, whereas TradeProd (2012) has not been updated to the most recent years and ends already in 2006. All other data sources span time periods from the mid-1990s to recent years. OP (2021) is available from 1994 to 2012 but only contains three year interval data.

Disaggregated industry-level data are available in BACI, ITPD-E, TiVA, TradeProd (2012, 2023), Comtrade, and WIOD.[11] The level of disaggregation varies, though: ITPD-E provides data for 120 manufacturing industries, BACI and Comtrade each report data for 35 industries, while TradeProd (2023) covers only nine broader sectors. TradeProd (2012) covers 26 manufacturing industries but spans a shorter time period. TiVA and WIOD, which are based on input-output data, provide information for 17 and 13 manufacturing industries, respectively. The other data sources offer information on single sectors: SGMD and OP (2021) include overall manufacturing trade; DOTS and WTF cover trade in both agricultural and manufacturing industries; and Eora26 includes trade data across all three broad sectors—manufacturing, agriculture, and services.

DOTS and Comtrade provide import data in addition to the commonly reported export data. Countries may have an incentive to collect import data more accurately and comprehensively due to tariffs and other restrictive non-tariff trade policy measures. In the subsequent analysis of EU membership trade effects, we will also use import data to assess whether using them instead of exports impacts the estimated effect size. The descriptive statistics reported in Table 1 further highlight non-negligible differences across the alternative data sources, underscoring the need for a systematic analysis of how dataset choices impact the quantitative magnitudes of the EU membership trade effect estimates.

[10] Detailed country breakdowns for each data source are provided in Table A1 in Appendix A.
[11] Table A2 in Appendix A provides details on the covered industries.

Table 1: Main data source characteristics

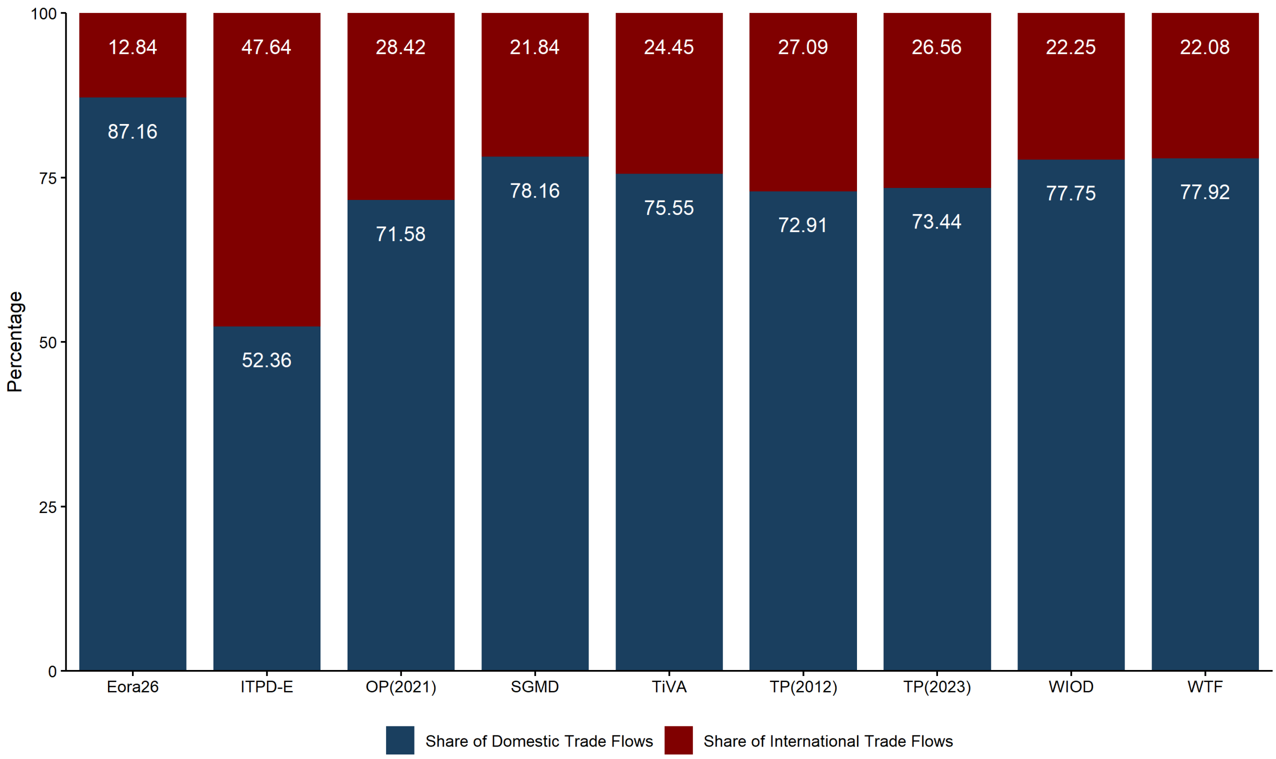
Datasets	Time and country coverage	Units	Obs.	Mean	S.D.	Min	Max
BACI (HS96)	228 countries; 1996-2019.	Thousands of US dollars.	616,897	3.33e+05	3.98e+06	0.001	4.90e+08
	Industry-level: 35 industries.		9,270,570	22183.83	3.36e+05	0.001	1.01e+08
DOTS	183 countries; 1970-2020.	US dollars.	716,062	4.80e+08	5.00e+09	1	4.81e+11
	Imports.		780,992	4.44e+08	4.72e+09	1	5.40e+11
Eora26	190 countries; 1990-2016.	US dollars.	974,700	2.64e+06	1.61e+08	0	6.28e+10
ITPD-E	237 countries; 2000-2016.	Millions US dollars.	714,951	534.26	24877.110	0	9.21e+06
	Industry-level: 120 industries.		34,619,387	11.03	663.829	0	9.65e+05
OP (2021)	65 countries; every three years in 1994-2012.	Millions of US dollars.	29,575	4819.46	91137.19	0	7776997
SGMD	229 countries; 1980-2016.	US dollars.	972,692	1.02e+09	2.99e+11	0	2.92e+14
TiVA	67 countries; 1995-2018.	Millions of US dollars.	107,736	6750.23	1.38e+05	0	1.42e+07
	Industry-level: 17 industries.		1,831,512	397.07	9,603.779	0	1.75e+06
TradeProd (2012)	231 countries; 1980-2006.	Thousands of US dollars.	702,322	5.24e+05	2.34e+07	0	4.14e+09
	Industry-level: 26 industries.		15,459,569	23802.32	1383990	0	5.85e+08
TradeProd (2023)	164 countries; 1966-2018.	Millions US dollars.	1,105,301	766.44	44224.4	0	1.38e+07
	Industry-level: 9 industries.		9,941,281	85.21	5918.541	0	3198207
UN Comtrade	202 exporters and 246 importers; 1988-2019.	US dollars.	510,999	1.29e+09	1.39e+10	0	1.40e+12
	Imports; 204 exporters and 246 importers.		599,846	1.09e+09	1.29e+10	0	1.64e+12
	Industry-level: 35 industries.		8,583,220	7.66e+07	1.17e+09	0	5.15e+11
WIOD	43 countries; 1995-2014.	Millions of US dollars.	35,735	13445.25	193644.7	0	1.39e+07
	Industry-level: 13 industries.		464,555	1034.25	17547.23	0	2279768
WTF	253 countries; 1984-2016.	Thousands of US dollars.	783,315	1.78e+06	8.16e+07	0.001	1.73e+10

Notes: ITPD-E is the International Trade and Production Database for Estimation collected by Borchert et al. (2021). SGMD is the Structural Gravity Manufacturing Database collected by Monteiro (2020). Eora26 is an aggregated input-output database from the Eora's global multi-regional input-oupt (MRIO) project. TiVA is the Trade in Value Added Database from OECD. OP (2021) is taken from Oberhofer and Pfaffermayr (2021). TradeProd (2012, 2023) are two versions of the Trade and Production Database from CEPII. WTF is the World Trade Flows Database collected by Feenstra and Romalis (2014). DOTS is the Direction of Trade Statistics Database from IMF. BACI (HS96) is BACI Database from CEPII. UN Comtrade is the Comtrade Database from the United Nations.

We are particularly interested in examining how the inclusion of domestic trade flows affects trade policy estimates. Due to recent advancements in the collection of trade data, there are now nine data sources that include domestic trade flows, utilizing three different methods for their construction.[12] ITPD-E, OP (2021), SGMD, TiVA, and TradeProd (2012, 2023) calculate domestic trade flows as the difference between gross production and total exports. WTF aggregates all trade flows and uses GDP instead of gross production for calculating domestic trade. Eora26 and WIOD rely on information from national input-output tables to derive domestic trade flows.

Figure 1 provides an overview of the share of domestic and cross-border trade flows in total sales across all data sources that include information on domestic trade. The figure shows significant variation in the share of domestic trade across the different data sources. ITPD-E records the smallest share of domestic trade, though it still amounts to more than half of total sales. The high level of disaggregation across 120 manufacturing industries leads to a larger number of missing domestic trade flows for certain countries and industries. This may have implications for the estimation results, which we will address when discussing the findings from the meta-regression analysis. Eora26 exhibits the largest share of domestic trade, accounting for around 87% of total sales. Its input-output data structure allows the inclusion of all trade flows across industries, including service trade, which is predominantly consumed domestically. Despite differences in time and country coverage, the share of domestic trade flows in the other data sources varies between 71.58% and 78.16% of total sales.

Figure 1: Share of domestic trade flows across data sources

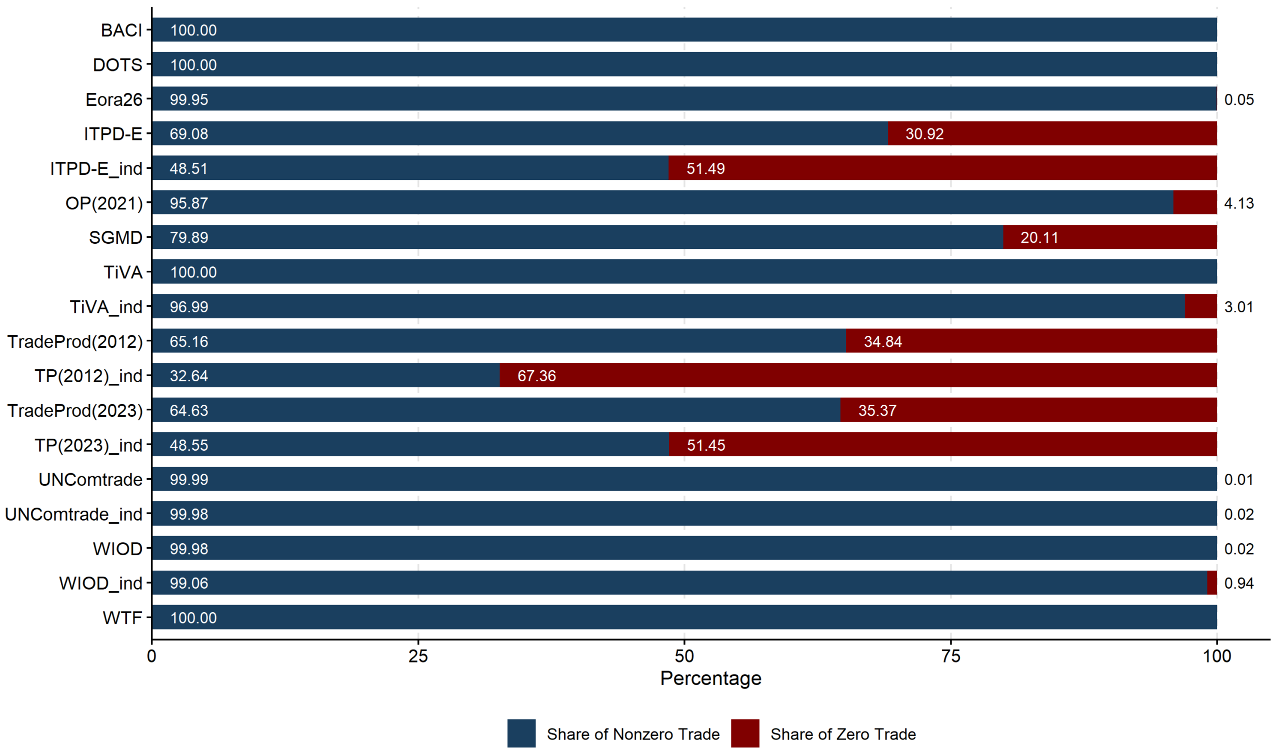


[12] Campos et al. (2021) empirically demonstrate that different methods for calculating domestic trade lead to similar estimates in structural gravity models. Consistent with this finding, we will primarily compare estimates that include domestic trade with more traditional approaches focusing solely on cross-border trade flows.

Figure 1 highlights that trade data sources lacking domestic trade flows miss out on a substantial amount of relevant transactions for analyzing the impact of trade policy on cross-border trade. Domestic trade is several times larger than international trade (with the exception of the ITPD-E), suggesting a significant potential for trade diversion from domestic to cross-border trade when trade policies reduce the relative costs of international trade. EU membership trade effect estimates based solely on international trade flows can only capture trade creation through increased trade among EU member states, which may, at least partially, result from trade diversion from non-member states. Including domestic trade flows allows us to derive theory-consistent estimates that account for both potential margins of trade diversion—the international and the domestic ones. Additionally, domestic trade flows help to control for globalization trends over the observed time periods (Bergstrand et al., 2015; Borchert and Yotov, 2017). A systematic comparison of gravity model based trade policy estimates from datasets that include or exclude domestic trade flows allows us to provide some evidence on the relative importance of substitution of domestic sales by international trade

The extensive and intensive margins of trade may respond heterogeneously to trade policy measures (see, e.g., Helpman et al., 2008; Baier et al., 2014; Adão et al., 2024; French and Zylkin, 2024). A comprehensive analysis of all trade effects from EU membership may therefore require the inclusion of zero trade flows in the underlying data sources. Figure 2 provides a descriptive overview of zero trade flows across all data sources. The largest shares of zeros are found in the disaggregated versions of TradeProd (2012) and ITPD-E. In TradeProd (2012), zero trade flows account for more than 60% of all potential bilateral combinations, while in ITPD-E, the share of zeros still makes up approximately half of all observations. As expected, this share decreases once individual sectors are aggregated to total manufacturing trade. SGMD, which only reports total manufacturing trade data, still includes about 20% zero trade flows. This is followed by OP (2021) and the disaggregated version of TiVA. In the eight remaining data sources, almost no zeros are reported. Notably, BACI, DOTS, and WTF do not include any zero trade flows at all. The meta-regression analysis will also shed light on the role of zero trade flows for the overall effect size of the estimates for the EU membership trade effect.

Figure 2: Share of zero trade flows



Notes: The ending _ind denotes disaggregated data from the original sources. The same data sources without this ending indicate country-level aggregated trade flow data.

3 Econometric specification, estimation and data samples

In this section, we introduce the standard structural gravity model, the chosen empirical specification, and address some estimation issues related to the estimation of EU membership trade trade effects using the twelve different data sources. We also discuss the alternative data samples constructed for the estimation, which are based on different dataset restrictions.

3.1 The econometric structural gravity model

The structural gravity framework is the workhorse model in the international trade literature. Following Anderson and Van Wincoop (2003), the empirical specification of the model is based on a system of structural gravity equations. At the country level, the panel data version is expressed as:

$$X_{ijt} = \frac{Y_{it}E_{jt}}{Y_t}\left(\frac{\tau_{ijt}}{\Pi_{it}P_{jt}}\right)^{1-\sigma}. \tag{1}$$

X_{ijt} denotes exports from exporter country i to importer country j in year t. Y_{it} represents the value of output of exporter i in year t, while E_{jt} depicts the total expenditures of importer j in the same year. Y_t expresses total world output in year t. τ_{ijt} captures (time-varying) bilateral trade costs between countries i and

j in year t. Assuming that each country specializes in producing only one single product, $\sigma > 1$ measures the elasticity of substitution between different goods from different countries. The structural components Π_{it} and P_{jt}, which Anderson and Van Wincoop (2003) label as "multilateral resistance", can be expressed as

$$\Pi_{it}^{1-\sigma} = \sum_{j} \left(\frac{\tau_{ijt}}{P_{jt}} \right)^{1-\sigma} \frac{E_{jt}}{Y_t}, \tag{2}$$

and

$$P_{jt}^{1-\sigma} = \sum_{i} \left(\frac{\tau_{ijt}}{\Pi_{it}} \right)^{1-\sigma} \frac{Y_{it}}{Y_t}. \tag{3}$$

A useful feature of the gravity model is its separability property, which allows for a straightforward transformation of the aggregated model into a system of sectoral gravity equations. As discussed above, the majority of data sources include industry-level trade data, for which the sectoral version of the gravity model is appropriate. Following Anderson and Van Wincoop (2004), we denote the industry dimension by k and express the sectoral gravity system as follows:

$$X_{ijt}^{k} = \frac{Y_{it}^{k} E_{jt}^{k}}{Y_t^{k}} \left(\frac{\tau_{ijt}^{k}}{\Pi_{it}^{k} P_{jt}^{k}} \right)^{1-\sigma}. \tag{4}$$

Traditionally, trade economists used to log-linearize the multiplicative structural gravity model and then proceeded with simple OLS estimation of the resulting empirical specification. However, as trade data are usually plagued by heteroscedasticity and might include a non-negligible share of zero trade flows, the OLS estimator provides biased and inconsistent estimates. Estimating the gravity equation in its multiplicative form by applying the PPLM estimator solves this problem (Santos Silva and Tenreyro, 2006).

Equations (1) to (4) illustrate another important technical aspect: the need to account for the multilateral resistance terms in empirical specifications of the gravity model. An early strand of the literature includes a series of country- and country-pair-specific observables to address this issue (Rose, 2004, among others). However, this approach is prone to omitted variable bias, leading the literature to shift toward using sets of fixed effects to control for the multilateral resistance terms. Anderson and Yotov (2010) propose a two-way gravity model with (time-varying) exporter and importer fixed effects, along with a series of (time-invariant) "gravity-style" bilateral trade cost variables to capture multilateral resistances. Egger and Nigai (2015) generalize the econometric specification to a three-way fixed effects gravity model and demonstrate that country-pair fixed effects are more effective in controlling for unobservable time-invariant trade cost variables.

The inclusion of country-pair fixed effects delivers an additional econometric benefit. The decision to join the EU is not random; the accession process considers various economic and political factors, including the level of economic development and geographic distance — two important drivers of bilateral trade. Consequently, the formation of the EU (similar to free trade agreements, FTAs) is not exogenous, which raises endogeneity concerns when estimating EU membership trade effects. By incorporating bilateral fixed effects alongside exporter-time and importer-time effects, we are able to control for a significant number of sources of a potential endogeneity issue. As a result, we can treat EU membership as conditionally exogenous, as discussed in the context of FTAs by Baier and Bergstrand (2007).

Against this backdrop, we estimate three-way gravity models that build on the theoretical foundations and incorporate recent econometric advancements. Due to distinctive features in the data sources, the empirical specifications differ in two dimensions: the treatment of domestic trade flows and the level of industrial disaggregation available in the underlying datasets. In country-level datasets without domestic trade flows, the empirical specification of the structural gravity model reads as:

$$X_{ij,t} = \exp[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \beta EU_{ij,t} + \theta FTA_{ij,t}] \times \epsilon_{ij,t} \ , \tag{5}$$

where $X_{ij,t}$ denotes trade from exporter country i to importer country j in year t. $\epsilon_{ij,t}$ is a multiplicative error term with mean one and $\exp[\ldots]$ denotes the exponential conditional mean function. $\pi_{i,t}$ and $\chi_{j,t}$ are exporter- and importer-time fixed effects that account for multilateral resistances. μ_{ij} is a country-pair fixed effect that captures time-invariant bilateral trade costs such as geographic distance, common language, and allows to treat EU membership (and FTAs) as conditionally exogenous (Baier and Bergstrand, 2007). $EU_{ij,t}$ and $FTA_{ij,t}$ are our sole time-varying bilateral trade costs measures. $FTA_{ij,t}$ is equal to one whenever a country-pair ij has a common bi- or plurilateral free trade agreement in force at year t, and zero else.[13] The EU membership dummy $EU_{ij,t}$ is equal to one whenever the exporter and the importer countries are both members of the EU in year t, and zero else. θ is a common coefficient for all free trade agreements capturing the average direct trade effect of such agreements. β is the coefficient of interest measuring the direct trade effect from common EU membership. The exponential transformation of the linear part of the model specification ensure a multiplicative structure of the gravity model which can be estimated with PPML (Santos Silva and Tenreyro, 2006) even for data sets with large shares of zero trade flows (Santos Silva and Tenreyro, 2011). In relative terms, the quantitative magnitude of the direct EU membership trade effect can be approximated by $[\exp(\beta) - 1] \times 100$.

For data sets with domestic and international trade flows the specification additionally includes a border dummy $Border_{ij}$, which takes on a value of one for cross-border trade flows and is zero else. Following Larch et al. (2018) the specification of the gravity model for cross-border and domestic trade reads as:

$$X_{ij,t} = \exp[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \beta EU_{ij,t} \times Border_{ij} + \theta FTA_{ij,t} \times Border_{ij} + \gamma Border_{ij} \times t] \times \epsilon_{ij,t} \ , \tag{6}$$

where $Border_{ij} \times t$ is an interaction term between the border dummy and a time trend. γ captures a common globalization effect and informs about the substitution of domestic trade by cross-border trade flows over time (Bergstrand et al., 2015; Oberhofer and Pfaffermayr, 2021). $Border_{ij}$ is further interacted with both trade policy measures, the FTA dummy and the common EU membership indicator. This specification follows the best practice recommendations discussed in Yotov et al. (2016) and should enable a cleaner identification of the EU membership trade effect. The interaction between EU membership and the border dummy additionally

[13] Data on FTAs are mainly sourced from the Dynamic Gravity Dataset (DGD). We augment the data with additional information provided in Mario Larch's Regional Trade Agreements Database (Egger and Larch, 2008) and from the Database on Economic Integration Agreements compiled by the NSF-Kellogg Institute (https://kellogg.nd.edu/nsf-kellogg-institute-data-base-economic-integration-agreements).

allows for substitution of domestic sales by cross-border trade within the EU and the $Border_{ij} \times t$ interaction helps to avoid a potential spurious correlation between increasing international trade over time and the successive enlargement steps of the EU.

For data sets with industry-level trade data our specification, the applied gravity model specifications simply add the industry-dimension k as an additional source of variation. The specification reads as:

$$X^k_{ij,t} = \exp[\pi^k_{i,t} + \chi^k_{j,t} + \mu^k_{ij} + \beta EU_{ij,t} \times Border_{ij} + \theta FTAs_{ij,t}] \times \epsilon^k_{ij,t} \;, \tag{7}$$

for data sets with only international trade flows and as:

$$X^k_{ij,t} = \exp[\pi^k_{i,t} + \chi^k_{j,t} + \mu^k_{ij} + \beta EU_{ij,t} \times Border_{ij} + \theta FTAs_{ij,t} \times Border_{ij} + \gamma Border_{ij} \times t] \times \epsilon^k_{ij,t} \;, \tag{8}$$

for data sets that additionally cover domestic trade. $\pi^k_{i,t}$ and $\chi^k_{j,t}$ are exporter-industry-time and importer-industry-time fixed effects. μ^k_{ij} denotes ij bilateral fixed effects that may vary over industries k. Our trade policy measures are only ij-specific and therefore are not varying over the included industries k. In all of the four alternative empirical specifications of the structural gravity model, we cluster standard errors at the exporter-importer pair level, as recommended by Larch et al. (2025) as one possible approach.

3.2 Sample definitions

In the first step, we are estimating Equations (5) to (8) for the twelve different data sources and select the respective specification that fits to the individual properties of each data source. The first set of estimates $\widehat{\beta}$ for the EU membership trade effect is based on the full samples available in all twelve data sources (Sample I). As discussed in detail above, the data sources differ in a variety of dimensions. In a next step, we are introducing alternative data set restrictions that are commonly applied to all data sources and re-estimate the empirical specifications of the structural gravity model. The upper part of Table 2 summarizes the alternatively imposed dataset restrictions.

Table 2: Sample restrictions and data sources

Samples	Description	Data sources
I	Full samples.	
II	Restricted time coverage from 2000 to 2014.	BACI (HS92), BACI (HS96), BACI (HS02), DOTS,
III	Restricted country coverage to “WIOD countries”.	Eora26, ITPD-E, OP (2021), SGMD, TiVA,
IV	Restrictions II and III.	TradeProd (2012), TradeProd (2023),
V	Three year interval data.	UN Comtrade, WIOD, WTF
VI	Restrictions III and V.	
VII	Imports as outcome of interest.	DOTS, UN Comtrade
VIII	Alternative HS industry classifications of BACI from 2002 to 2019.	BACI (HS92), BACI (HS96), BACI (HS02)
IX	Total goods trade instead of manufacturing trade.	BACI (HS92), BACI (HS96), BACI (HS02), ITPD-E, TiVA, UN Comtrade

In Sample (II) we restrict all data sets to the shortest time span available across all twelve data sources. The first year available in ITPD-E is 2000. WIOD ends in 2014. By restricting all data sources to the years 2000 to 2014, we obtain a comparable time coverage of 15 years across all data sets, for which we estimate a second set of EU membership trade effects $\widehat{\beta}$. WIOD includes the smallest country coverage across all available data sources. The first version of this input-output database runs from 1995 to 2011 and covers 40 countries. The second release from 2016 adds three more countries covering the years from 2000 to 2014. The combined version of WIOD provides unbalanced data with up to 43 countries, which cover more than 85% of world GDP (Timmer et al., 2016). These 43 countries are available in all other eleven data sources and thus we are restricting these to the WIOD coverage for Sample restriction (III). Sample (IV) combines dataset restrictions (II) and (III), yielding an estimation sample that includes only 43 countries and is limited to the period from 2000 to 2014.

One popular practice in gravity model estimation is the use of interval data instead of consecutive annual panel data (see, e.g., Baier and Bergstrand, 2007; Olivero and Yotov, 2012; Oberhofer and Pfaffermayr, 2021). Cheng and Wall (2005) argue that in a fixed effects estimation setting the dependent and independent variables may not be able to fully adjust within a single year. As a consequence, interval data might be better suited to capture the overall effect of trade policies on trade flows. Egger et al. (2022) challenge this view argue for the use of consecutive years whenever a (full) three-way fixed effects estimation framework is applied. Larch and Yotov (2024) compare the estimates from consecutive data with three-year interval data for one data source and do not find systematically different RTA trade effects. We aim at a more systematic analysis for the role of consecutive annual panel data versus interval data in gravity model estimation, and in Sample (V) construct three year interval data for all available data sources.[14] In Sample restriction (VI) we use three year interval data and only include the 43 countries covered by WIOD. These sample restrictions (V) and (VI) allow us to obtain two additional sets of estimates for the EU membership trade effect.

Samples (VII) and (VIII) make use of alternative bilateral trade measures available in some of our data sources. DOTS and UN Comtrade not only include bilateral trade reported from the exporting economies but also include the (mirrored) import figures recorded by the destinations. The destination economies might have an incentive to more accurately collect bilateral trade data to maximize tariff revenues. As a result, import data might be less prone to measurement errors. In sample restriction (VII) we replace exports by the available import figure and and re-estimate the EU membership trade effects for all sample restrictions from (I) to (VI) only for data from DOTS and UN Comtrade.

BACI is the only data source for which detailed product-level data for cross-border trade flows is publicly available. For our purpose, we aggregate the product-level data to a comparable industry-level covered by the other data sources. The mapping of products to industries is regulated by Harmonized system (HS), a product nomenclature developed by the World Customs Organization (WCO).[15] Over the course of our sample period, the HS system has been revised three times in 1992, 1996 and 2002. The revisions introduced some changes in

[14] The data from Oberhofer and Pfaffermayr (2021) are already only available as three year interval data. For the other data sources we also provide a robustness check varying the interval length to four and five years as in Olivero and Yotov (2012). The meta analysis estimation results for the alternative interval lengths are documented in Table D1 in the appendix.

[15] A description of HS can be accessed only via `https://www.wcoomd.org/en/topics/nomenclature/overview/what-is-the-harmonized-system.aspx`.

the mapping of products into industries. To assess the potential effect of these alternative mappings for the EU membership trade effect estimates, we separately apply sample restrictions (I) to (VI) to the three alternative industry-aggregates for bilateral trade flows.

Some widely used data sources, such as DOTS, SGMD and WTF, provide only aggregated trade flows at the country level, with some covering total goods trade (e.g., DOTS) and others focusing exclusively on manufacturing goods (e.g., SGMD). Historically, gravity models have often been estimated using this level of aggregation. The growing availability of multi-sector trade models has brought increased attention to industry-level heterogeneity (Eaton and Kortum, 2002; Caliendo and Parro, 2015). Recent econometric studies have begun to quantify the aggregation bias that arises from ignoring industry-level heterogeneity when estimating gravity models of trade (Breinlich et al., 2024; French and Zylkin, 2024). In our Sample (IX), we address this issue by aggregating bilateral trade data and domestic trade flows into total goods trade at the country level whenever detailed industry-level information is available. This approach enables a direct comparison of the estimates from (artificially) aggregated data with the ones from data sources that inherently offer only country-level data. Additionally, "within" data source comparisons allow us to quantify the average aggregation bias across all data sources.

Domestic trade is a core component of all theoretical trade models. (see, e.g., Anderson, 1979; Anderson and Van Wincoop, 2003; Arkolakis et al., 2012). Anderson (1979) and Anderson and Van Wincoop (2003) develop a microfounded gravity model incorporating domestic trade flows, demonstrating that a decline in "multilateral resistance" increases relative resistance for domestic trade. Consequently, a bilateral trade policy that reduces trade costs between country pairs may not only divert trade away from excluded partners but also reduce domestic trade, as cheaper imports substitute domestic production.

In empirical studies of trade policies, this channel has long been overlooked, as traditional data sources such as DOTS, BACI, and Comtrade typically only cover cross-border trade flows. Addressing this gap, Yotov (2022) provides 15 arguments in favor of including domestic trade data in structural gravity model estimation.[16] In our meta analysis, we provide systematic evidence on the quantitative effects of including domestic trade in empirical structural gravity models. To this end, we re-estimate all sample restrictions from (I) to (IX), excluding domestic trade data from data sources that actually include domestic trade.

We estimate the EU membership trade effects across all nine sample restrictions and datasets, both with and without domestic trade wherever applicable, using two commonly applied Stata packages: *ppmlhdfe* and *ppml_fe_bias*. We employ these packages for their distinct advantages. *ppmlhdfe* addresses computational challenges associated with the large number of country-pair fixed effects required to consistently identify the effects of time-varying trade policies (Larch et al., 2019). *ppml_fe_bias* extends *ppmlhdfe* by correcting various incidental parameter problems that arise when multiple fixed effects are included in PPML models, and adjusts biased standard errors, as demonstrated in Weidner and Zylkin (2021). The comparison of the two packages allows to evaluate whether the choice of estimation approach affects the quantitative magnitudes of the EU

[16] Incorporating domestic trade flows in the gravity model estimation allows, for instance, to resolve the border and distance puzzles (Bergstrand et al., 2015; Yotov, 2012), yield more accurate estimates for WTO membership impacts (Larch et al., 2025), currency union and Euro effects (Larch et al., 2018, 2019) and for trade diversion (Dai et al., 2014). Domestic trade data further enable the identification of unilateral trade policies within a structural gravity framework (Heid et al., 2021; Oberhofer et al., 2021).

membership trade effects.[17]

Overall, the nine different sample definitions yield a total of 586 EU membership trade effect estimates. Of these, 436 are based on country-level data, while the remaining 150 are derived from gravity specifications at the industry level. 165 estimates are obtained from trade data that includes domestic trade flows, whereas 421 are based exclusively on cross-border trade flows. Table B1 in the appendix provides a breakdown of the total number of estimates by data source and sample restriction.

4 Estimation results, meta-analysis and discussion

4.1 Descriptive analysis of the EU membership trade estimates

Figure 3 and Table 3 provide a descriptive overview of the 586 EU trade effect estimates based obtained from different data sources and sample restrictions. Figure 3 seperately displays the estimates for each of the twelve different data sources, including the alternative revisions of the HS classification in BACI. Circles represent estimates based solely on cross-border trade flows, whereas triangles indicate estimates from datasets that additionally incorporate both domestic sales. Faded circles or triangles indicate statistically insignificant estimates, applying the rather generous 10% significance level as the threshold.

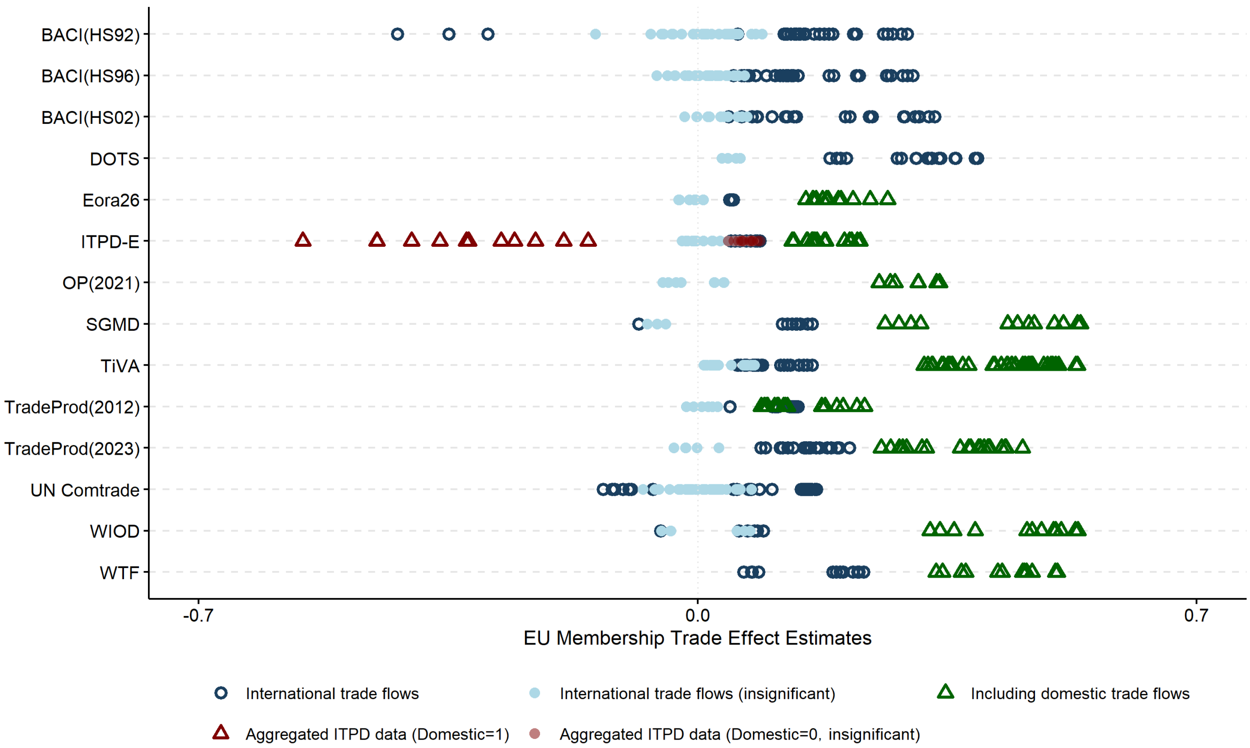
Figure 3 reveals key stylized facts regarding the distribution of the EU membership trade effects. First, a large majority of the estimates lie to the right of the zero line, supporting previous findings of a positive EU membership trade effect. The statistically significant average EU membership trade estimate of 6.54% ($[\exp(0.0634) - 1] \times 100$) across all data sources and sample restrictions, reported in the first line of Table 3, confirms the graphical impression. Second, many of the estimates based solely on international trade flows are small in magnitude and statistically indistinguishable from zero, as indicated by the faded circles close to the zero line. Third, the green triangles suggest that the effect size of the EU membership trade estimate increases once domestic trade flows are included in the data, and these estimates are all statistically significant at the 10% threshold.

Finally, some samples from the ITPD-E and BACI (HS92) data sources lead to large negative and statistically highly significant EU membership trade effect estimates. The three instances from BACI (HS92) arise from samples that use three-year intervals and include disaggregated industry-level data, while all negative estimates from the ITPD-E dataset are based on data samples where industry-level data has been aggregated to total goods trade at the country-level. This approach appears to be prone to substantial aggregation bias due to missing domestic trade data for certain industries, which are treated as zeros in standard aggregation methods. In our meta-regression analysis below, we revisit this issue and assess the sensitivity of our overall findings to the negative estimates obtained from aggregated ITPD-E data.

Table 3 reports the (weighted) average EU membership trade effect estimates, calculated across different data sources and the nine alternative dataset restrictions. The table distinguishes between estimates based solely on international trade flows and those that additionally incorporate domestic trade. The weighted averages are

[17] The bias correcting PPML estimator does not converge for some samples with a larger number of disaggregated industries and exporter-industry-time and importer-industry-time fixed effects. Overall, we are losing 33 estimates from samples I to VI.

Figure 3: EU membership trade effect estimates across data sources



Notes: EU membership trade effect estimates are reported. The Figure includes all twelve different data sources, and three revisions of the HC classification in the BACI database. Faded circles or triangles indicate statistically insignificant estimates at the 10% level.

derived using the standard errors associated with the EU membership trade effect estimates. The first row provides a direct comparison of the overall averages across data sources and sample restrictions, differentiating between estimates with and without domestic trade flows. The reported figures align with the previously presented descriptive evidence on the role of domestic trade. Specifically, the weighted average point estimate for the EU membership trade effect is 0.0637 (6.58%) when considering only international trade flows. This estimate increases substantially to 0.2925 (33.98%) when domestic trade flows are added to the data. The average estimate from data incorporating domestic trade is 5.2 times larger than the one based solely on international trade flows. This difference may be crucial not only for a quantitative assessment of the trade benefits associated with EU membership but also for deriving policy implications.

Table 3 further illustrates that the inclusion of domestic trade flows has a heterogeneous impact on the average effect size of the estimated EU membership trade effect across different data sources. In the ITPD-E database, the weighted average of all parameter estimates increases from 0.029 to 0.185, while in WIOD, it rises from 0.0312 to 0.4406 when domestic trade flows are included in the estimation sample. Estimates based solely on international trade data from Eora26, OP (2021), and WIOD indicate statistically insignificant EU trade effects. However, once domestic trade is added to the data, the estimates become statistically significant, amounting to approximately 20%, 33.7% and 55.4%, respectively.

Table 3: Summary statistics: Estimated EU membership trade effects by data sources and sample restriction

	Panel A: International trade flows				*Panel B: Including domestic trade flows*			
	Mean	SE	WM	SE	Mean	SE	WM	SE
All	0.0634	0.0060	0.0637	0.0057	0.3130	0.0115	0.2925	0.0121
BACI (HS92)	0.0263	0.0197	0.0268	0.0182				
BACI (HS96)	0.0070	0.0078	0.0145	0.0085				
BACI (HS02)	0.1114	0.0242	0.0987	0.0210				
DOTS	0.2158	0.0362	0.2336	0.0327				
Eora26	0.0102	0.0086	0.0116	0.0074	0.1921	0.0101	0.1825	0.0100
ITPD-E	0.0226	0.0081	0.0290	0.0080	0.1779	0.0073	0.1851	0.0083
OP (2021)	-0.0030	0.0128	0.0004	0.0127	0.2936	0.0121	0.2902	0.0118
SGMD	0.0715	0.0291	0.0800	0.0272	0.4217	0.0298	0.4066	0.0325
TiVA	0.0537	0.0056	0.0518	0.0056	0.4163	0.0120	0.4078	0.0122
TradeProd (2012)	0.0871	0.0115	0.0926	0.0114	0.1234	0.0133	0.1219	0.0130
TradeProd (2023)	0.1168	0.0164	0.1265	0.0148	0.3686	0.0127	0.3621	0.0134
UN Comtrade	-0.0029	0.0169	0.0071	0.0174				
WIOD	0.0332	0.0169	0.0312	0.0180	0.4501	0.0218	0.4406	0.0242
WTF	0.1659	0.0197	0.1695	0.0196	0.4269	0.0169	0.4154	0.0179
Sample I	0.0989	0.0141	0.0986	0.0146	0.3489	0.0268	0.3312	0.0338
Sample II	0.0016	0.0082	0.0047	0.0078	0.2733	0.0154	0.2535	0.0208
Sample III	0.0785	0.0115	0.0720	0.0121	0.3263	0.0287	0.3090	0.0299
Sample IV	-0.0051	0.0092	0.0067	0.0099	0.2479	0.0215	0.2345	0.0221
Sample V	0.1177	0.0143	0.1158	0.0145	0.3596	0.0319	0.3419	0.0342
Sample VI	0.0918	0.0164	0.0842	0.0162	0.3209	0.0338	0.3049	0.0311
Sample VII	0.1064	0.0278	0.0995	0.0256				
Sample VIII (HS92)	0.0148	0.0574	0.0672	0.0310				
Sample VIII (HS96)	0.1438	0.0268	0.1283	0.0258				
Sample VIII (HS02)	0.1575	0.0289	0.1451	0.0284				
Sample IX	0.1372	0.0076	0.1281	0.0062	0.0585	0.0834	0.1896	0.0728
Sample IX (exclude ITPD-E)	0.1459	0.0081	0.1355	0.0066	0.4472	0.0210	0.4389	0.0209

Notes: SE = standard error, WM = weighted mean using the inverse of the standard error of the estimates as weights. Sample I includes the original datasets, Sample II is the sub-sample from 2000 to 2014, Sample III only includes WIOD countries, sample IV includes WIOD countries from 2000 to 2014. Sample V uses three-year intervals and Sample VI additionally restricts the country sample to WIOD countries. Sample VII uses imports as dependent variables. Sample VIII uses the three different revisions of the HC classifications in the BACA dataset. Sample IX uses total trade from BACI and UN Comtrade databases instead of total manufacturing trade. Please see Table 2 for more details.

Estimates based solely on international trade flows tend to be quantitatively larger when using WTF and DOTS compared to the other ten data sources. This suggests that databases aggregating trade data across all goods yield larger positive estimates. This pattern is further supported by the fact that the mean estimate from Sample IX exceeds the average estimates derived from earlier versions of the BACI or UN Comtrade databases.

Restricting the analysis to the same time and country coverage in Sample VIII reveals that estimates based on the 1992 revision of the HS classification are statistically insignificant. In contrast, estimates from the BACI Rev.1996 and Rev.2002 indicate positive and significant EU membership trade effects. These findings highlight the importance of examining the impact of different revisions of the HS classification.

Table C2 in the appendix provides further descriptive evidence that the estimated effect size of the EU membership trade effect varies "within" data sources when alternative sample restrictions are applied. The Cochran Q test rejects the null hypothesis of effect size homogeneity at the 1% significance level for the whole distribution of 586 EU membership trade effect estimates, as well as within each of the twelve data sources. This finding is further supported by the provided I^2 and H^2 test statistics.

Additional Cochran Q tests, reported in Table C2, indicate statistically significant differences in the estimated effect size across various sample restrictions. Specifically, significant heterogeneity is detected between samples that include only international trade flows and those that also account for domestic trade, between samples covering all goods trade and those restricted to manufacturing trade, as well as between different revisions of the HC classification used in BACI. In contrast, replacing exports with imports and using aggregate country-level trade flows or detailed industry-level data do not appear to result in substantially different effect size estimates for the EU membership trade effect.

Overall, the descriptive evidence indicates that the estimated EU trade effect is substantially larger when domestic trade is incorporated than when only cross-border trade flows are considered. Moreover, estimates vary significantly across different data sources and sample restrictions, underscoring the need for further and more systematic empirical investigation.

4.2 Meta-regression analysis

In economics, a meta-analysis typically consists of three steps: synthesizing the collected estimates, conducting a meta-regression analysis, and testing for publication bias (see, e.g., Stanley, 2001; Kasy, 2021). In this paper, we examine the impacts of data source choices and sample restrictions on the effect size of the estimated EU membership trade effect. Since all estimates are derived from the same standard specifications of the structural gravity model of trade, we only need to adopt the second step and apply a meta-regression analysis.

In its generic form, a meta-regression model is written as:

$$\hat{\beta}_m = \iota_0 + \sum_{k=1}^{K} \iota_k X_{mk} + \mu_m, \tag{9}$$

where $\hat{\beta}_m$ is m-th estimate of β, which in our case is an estimate of the EU membership trade effect from the gravity model. X_{mk} are dummy variables either indicating alternative data sources ($k = 1, \ldots, 11$) or imposed sample restrictions and estimation procedures ($k = 1, \ldots, 9$). μ_m is the is the remainder error term. Across the full sample of estimates for the EU membership trade effect, $m = 1, \ldots, 586$.

Table B1 summarizes the number of estimates collected from each of the twelve data sources and for which we construct individual dataset dummy variables. Table 4 presents a descriptive overview of the empirical distribution of the obtained estimates for the nine sample restrictions discussed in Section 3.2. 28.2% of all estimates in our sample are based on data that incorporate domestic sales alongside bilateral cross-border trade flows, while 54.6% are based on restricted samples that include only countries covered in the WIOD database.

Table 4 further shows that nearly two-thirds (65.2%) of all estimates rely exclusively on trade data collected for manufacturing industries, while approximately one-quarter (25.6%) are based on disaggregated industry trade flows. A three-year time interval is used for about 37.4% of all estimates, whereas the remaining 63.6% are based on consecutive annual data. In nearly 30% of all estimation samples, the time span is restricted to the years 2000 to 2014. Data sources based on (international) input-output tables account for 20.5% of all estimates, whereas import data replace exports in only 7.7% of all samples. A 46.4% share of the estimates is obtained

Table 4: Meta-regression analysis: Summary statistics for the sample restriction dummy variables

Variables	Obs.	Mean	Std. dev.	Min	Max
Domestic	586	0.2816	0.4501	0	1
WIODcountries	586	0.5461	0.4983	0	1
Industry	586	0.2560	0.4368	0	1
Year0014	586	0.2986	0.4581	0	1
3yearintervals	586	0.3737	0.4842	0	1
Imports	586	0.0768	0.2665	0	1
Manu	586	0.6519	0.4768	0	1
Package	586	0.4641	0.4991	0	1
IO	586	0.2048	0.4039	0	1

using the bias-corrected Stata package *ppml_fe_bias*, while the remaining 53.6% are estimated with the standard *ppmlhdfe* routine. The smaller share of estimates for the EU membership trade effect from *ppml_fe_bias* is due to convergence issues in samples that include the largest number of countries and industries, as explained in Footnote 17.

4.3 Meta-regression results

To estimate Equation (9), we employ four different estimators that are commonly used in quantitative meta-analysis (Stanley and Jarrell, 2005). We adapt these estimators to our specific context, taking into account the way we derived the 586 different EU membership trade effects. All applied estimators use heteroskedasticity-robust standard errors.

The first estimator is ordinary least squares (OLS). The second is a weighted least squares (WLS) estimator, where effect size estimates are weighted by the inverse of their standard errors. These two estimators are used to analyze the impact of data source selection choice on the effect size of the EU membership trade effect estimates. When investigating the role of dataset restrictions, we introduce two additional estimators. The third estimator is a fixed effects model that exploits only within-data source variation in the EU membership trade effects for all twelve different data sources and applies OLS to the within-transformed data. The fourth estimator applies WLS while incorporating data source dummy variables instead of applying a within-transformation to all variables. The latter two estimators do not allow to identify the impact of using trade data from official statistics versus input-output tables, as the data compilation method does not vary within data sources. The effect of using input-output tables is thus absorbed by the fixed effects.

4.3.1 Data source effects

Table 5 presents the results of our first set meta-regressions, in which we regress the estimated EU membership trade effects solely on data source dummy variables. We define dummy variables for all twelve different data sources and separately distinguish three different HS classifications used for compiling BACI. The TradeProd (2012) version is omitted, meaning that the average estimated EU membership trade effect from this data source is captured by the constant.

Table 5: Meta-regression analysis: Data source effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: All estimates		*Panel B: International trade flows*		*Panel C: Including domestic trade*	
Variable	OLS	WLS	OLS	WLS	OLS	WLS
Eora26	-0.0037	-0.0255	-0.0769***	-0.0809***	0.0687***	0.0606***
	(0.0218)	(0.0204)	(0.0142)	(0.0135)	(0.0166)	(0.0164)
ITPD-E	-0.0973***	-0.0476**	-0.0466***	-0.0506***	-0.1488***	-0.0450
	(0.0263)	(0.0190)	(0.0132)	(0.0128)	(0.0506)	(0.0378)
OP (2021)	0.0404	0.0410	-0.0902***	-0.0922***	0.1702***	0.1683***
	(0.0395)	(0.0382)	(0.0167)	(0.0165)	(0.0177)	(0.0173)
SGMD	0.1417***	0.1037***	-0.0157	-0.0126	0.2983***	0.2848***
	(0.0424)	(0.0378)	(0.0306)	(0.0288)	(0.0322)	(0.0346)
TiVA	0.1458***	0.1425***	-0.0125	-0.0236*	0.3032***	0.2939***
	(0.0238)	(0.0243)	(0.0134)	(0.0133)	(0.0172)	(0.0170)
TradProd (2023)	0.1379***	0.1271***	0.0297	0.0339*	0.2453***	0.2402***
	(0.0236)	(0.0218)	(0.0199)	(0.0185)	(0.0185)	(0.0188)
WIOD	0.1368***	0.1270***	-0.0540***	-0.0614***	0.3268***	0.3188***
	(0.0460)	(0.0462)	(0.0200)	(0.0209)	(0.0253)	(0.0272)
WTF	0.1916***	0.1619***	0.0788***	0.0769***	0.3035***	0.2935***
	(0.0311)	(0.0284)	(0.0223)	(0.0222)	(0.0213)	(0.0220)
BACI (HS92)	-0.0361*	-0.0216	-0.0183	-0.0089		
	(0.0212)	(0.0156)	(0.0224)	(0.0173)		
BACI (HS96)	-0.0134	-0.0274**	0.0043	-0.0148		
	(0.0160)	(0.0139)	(0.0175)	(0.0157)		
BACI (HS02)	0.0483***	0.0278*	0.0660***	0.0405**		
	(0.0177)	(0.0163)	(0.0191)	(0.0179)		
DOTS	0.1539***	0.1648***	0.1716***	0.1775***		
	(0.0244)	(0.0222)	(0.0255)	(0.0235)		
UNComtrade	-0.0644***	-0.0655***	-0.0466***	-0.0528***		
	(0.0137)	(0.0127)	(0.0154)	(0.0147)		
Constant	0.1048***	0.1053***	0.0871***	0.0926***	0.1234***	0.1219***
	(0.0091)	(0.0087)	(0.0114)	(0.0113)	(0.0133)	(0.0131)
Obs.	586	586	421	421	165	165
R^2	0.2972	0.2950	0.2909	0.3730	0.6697	0.6751

Notes: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Heteroskedasticity robust standard errors in parentheses. The TradProd(2012) data source serves as the baseline. The average EU membership trade effect estimate for TradProd(2012) is captured by the constant.

Panel A of Table 5 pools all estimates together. Panel B includes only estimates derived from samples excluding domestic trade flows, while Panel C focuses on the 165 estimates obtained from datasets that additionally account for domestic trade. Columns (1), (3) and (5) apply the simple OLS estimator with heteroskedasticity-robust standard errors, while Columns (2), (4) and (6) report the estimation results from WLS, with the inverse of the standard errors of the parameter estimates as weights. In Panels B and C, the WLS estimations yield higher R^2 values than the OLS estimates, while in Panel A, OLS provides a slightly better fit. To account for differences in the precision of the EU membership trade effect estimates, the subsequent discussion focuses on the WLS estimation results.

The bottom row of Table 5 shows that, according to the WLS estimates, the average effect size for the EU membership trade effect ranges from 9.26% to 12.19% in the baseline category, consisting of TradeProd (2012). The other parameter estimates capture differences in in percentage points relative to this basline. Column (2) of Panel A further illustrates that using other data sources, such as SGMD, TiVA, WIOD, WTF, and DOTS, results in statistically and quantitatively significantly larger effect size estimates for the EU membership trade effect in our full sample of all estimates. The difference is most pronounced for DOTS and WTF, with more than a 16 percentage point larger average effect size. Interestingly, the newer 2023 version of TradeProd yields average estimates that are 12.7 percentage points larger than those from the earlier 2012 version. In contrast, ITPD, UnComtrade and BACI with the 1996 HS classification deliver smaller average EU membership trade effect estimates, with the difference being statistically significant.

The pattern changes when we limit the analysis to estimates based solely on international trade flows. Column (4) shows that, in this case, the estimates from TradeProd (2012) are among the largest, while for some other data sources, the difference becomes negative and statistically significant. For input-output table-based data sources, such as TiVA, WIOD and Eora26, the difference now turns out to be statistically significant.

On the other hand, once we add domestic trade to the analysis, the average EU membership trade effect estimate from TiVA and WIOD exceeds that of TradeProd (2012) by more than 30 percentage points. In general, the effect size differences are amplified by the inclusion of domestic trade. The ITPD database is the only data source, with domestic trade included, that does not show a statistically significant difference in the average EU membership trade effect when compared to TradeProd (2012). However, the effect size of the average EU membership trade effect estimate also differs across other data sources that account for domestic trade. For example, the effect size from WTF is approximately 23 percentage points larger than that from Eora26 ($= 0.2935 - 0.0606$) when we account for domestic trade in both data sources.

The results from the first meta-regression analysis reported in Table 5 reveal that the choice of a particular data source can have significant implications for the estimated effect size of the EU membership trade effect, regardless of any data source restrictions imposed (see Panel A). Furthermore, Panels B and C show that the inclusion or exclusion of domestic trade flows has heterogeneous effects on the estimated effect size across different data sources. The numerous statistically significant differences highlight that the quantitative assessment of a trade policy measure, such as EU membership, can vary considerably due to seemingly minor decisions regarding the choice for a particular data source. As a next step, we assess whether dataset restrictions also influence the effect size of the EU membership trade effect estimate, building on the first indications from our previous analysis regarding the importance of domestic trade.

4.3.2 Dataset restriction effects

Table 6 presents the main meta-regression results for the imposed dataset restrictions from Table 2. The parameter estimates are qualitatively consistent across all four estimators and quantitatively similar in magnitudes. This suggests that the choice of a particular estimator does not substantially affect the main findings of the meta-regression.

The bottom row of Table 6 shows that the full set of dataset restriction dummy variables explains a considerable share of the variation in effect size estimates for the EU membership trade effects. Moreover, the WLS estimator outperforms both the unweighted OLS and unweighted fixed-effects estimators in terms of model fit and explicitly accounts for the precision of the EU membership trade effect estimates. Column (4), which adds data source dummy variables to the WLS estimator, reports the highest R^2, indicating that more than two-thirds of the total variation in EU membership trade effects can be explained by the specified meta-regression model. Based on this standard goodness-of-fit measure and the results for differences across data sources documented in Table 5, we select Column (4) as our preferred meta-regression model specification. Most of the meta-regression estimates are, however, both qualitatively and quantitatively robust across the four applied estimators.

Table 6: Meta-regression analysis: Main results for dataset restrictions

Variable	(1) OLS	(2) WLS	(3) FE	(4) WLS
Domestic	0.1856***	0.2090***	0.2052**	0.2137***
	(0.0199)	(0.0142)	(0.0688)	(0.0104)
WIODcountries	-0.0114	-0.0121	-0.0146**	-0.0083
	(0.0117)	(0.0094)	(0.0053)	(0.0068)
Industry	0.0079	0.0221**	0.0085	0.0198**
	(0.0121)	(0.0101)	(0.0115)	(0.0086)
Year0014	-0.0802***	-0.0763***	-0.0759***	-0.0716***
	(0.0118)	(0.0101)	(0.0169)	(0.0080)
3yearintervals	0.0251*	0.0243**	0.0258	0.0250***
	(0.0142)	(0.0119)	(0.0226)	(0.0083)
Imports	0.0333*	0.0241	0.0272	0.0223**
	(0.0188)	(0.0181)	(0.0231)	(0.0105)
Manu	-0.0335**	-0.0409***	-0.0369	-0.0575***
	(0.0164)	(0.0119)	(0.0536)	(0.0112)
Package	-0.0030	0.0011	-0.0092	-0.0070
	(0.0118)	(0.0095)	(0.0071)	(0.0068)
IO	0.0517***	0.0076		
	(0.0158)	(0.0134)		
Constant	0.1189***	0.1188***	0.1296***	0.0828***
	(0.0157)	(0.0131)	(0.0278)	(0.0125)
Data source FEs	No	No	Yes	Yes
Observations	586	586	586	586
R^2	0.3332	0.4419	0.3924	0.6705

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Heteroskedasticity robust standard errors in parentheses. Column (3) applies a within-data source transformation. The reported R^2 is column (3) is its within measure. Column (4) adds data source dummy variables as additional controls.

The estimates presented in Column (4) suggest that methodological choices do not significantly alter the effect size of the EU membership trade effect. Specifically, there is no statistically significant difference between estimates obtained using Stata's standard high-dimensional fixed effects PPML package and those derived from the bias-correcting procedure developed by Weidner and Zylkin (2021). The used data sources appear to cover a sufficiently large time dimension, ensuring that the incidental parameter bias has little impact on the quantitative size of the estimates.

Restricting the sample period to the years 2000–2014, the shortest period available across all data sources, has, by contrast, a statistically significant negative effect on the estimated EU membership trade effect. On average, and across all data sources, the estimate is approximately 7.2 percentage points smaller when the analysis is limited to these 15 years. In our heterogeneity analysis summarized in Section 4.3.4 below, we examine this issue in greater depth.

The choice of country sample size does not systematically affect the magnitude of the estimated EU membership trade effect. Using the fixed effects estimator in Column (3), we find a small, negative, and statistically significant effect, suggesting a reduction in the estimated average effect size by approximately 1.5 percentage points when restricting the sample to the main economies covered by WIOD. In the other three columns, however, the meta-regression estimate is even smaller and not statistically different from zero. We, however, return to this issue with a more detailed analysis in Section 4.3.4.

Other data source restrictions, such as using aggregated versus detailed industry-level data, three-year interval data versus consecutive annual data, or replacing exports with imports have some impact on the estimated effect size of EU membership for trade, though not very large ones. In our preferred WLS specification with dataset dummy variables, reported in Column (4), using disaggregated industry-level data increases the estimated average EU membership trade effect by approximately 2 percentage points. Similarly, the effects for imports and interval data are positive, statistically significant, and of comparable magnitude ranging between 2.2 and 2.5 percentage points increases in the estimated effect size.

One might question the ad-hoc choice of using a three-year interval length. To address this concern, Table D1 presents a robustness check in which we extend the interval length to four and five years, respectively. As a result, the number of available estimates decreases to 569 for the two extended interval lengths. The modification to five year interval data has virtually no impact on the meta-regression analysis, as indicated by the parameter estimates reported in Column (8) of Table D1. The estimates based on five year interval data continue to yield average EU membership trade effect estimates that are approximately 2 percentage points larger than those derived from consecutive annual data. When using the four-year interval, the effect reverses, indicating a 2.7 percentage point smaller average trade effect estimate of EU membership (Column (4) of Table D1). Consequently, the choice of interval length appears to influence the estimated effect size of trade policy measures in structural gravity model estimations.

The estimated impact of using input-output table-based data sources versus official statistics on the size of the EU membership trade effect estimates remains ambiguous, as shown in the first two columns of Table 6. As noted above, the parameter associated with a data source's compiling method cannot be identified when applying the data source fixed-effects estimators. The OLS estimator in Column (1) suggests that data sources based on input-output tables lead to estimated EU membership trade effects that are approximately 5.2 percentage points larger. However, when accounting for the uncertainty of effect size estimates, the WLS estimator yields a much smaller and statistically insignificant impact of the data compilation method.

The most significant differences in the effect size estimates for the EU membership trade effect arise from two dataset choices: restricting the data sources to manufacturing industries only and incorporating domestic trade flows in the estimation. When limiting the sectoral coverage to manufacturing industries in data sources that also cover additional sectors, the estimated average EU membership trade effect decreases by 5.75 percentage points in our preferred specification reported in Column (4) of Table 6. The other three specifications provide similar quantitative effects, although the estimate in Column (3) is slightly smaller and not statistically significant. In our heterogeneity analysis in Section 4.3.4, we will revisit this issue and examine the heterogeneity of the EU membership trade effect across industries in greater detail.

The inclusion of domestic trade in gravity model specifications has the most significant impact on the estimated effect size of the EU membership trade effect and is robust across all four estimators. In our preferred specification in Column (4), the average estimated EU membership trade effect is approximately 21.4 percentage points larger in datasets that account for domestic trade. The smallest estimate, reported in Column (1), amounts to 18.6 percentage points. All estimates are statistically significant at the 1% level. This finding indicates that EU membership not only increases trade with other EU members at the expense of non-participating third countries but also induces a substitution of domestic sales with cross-border trade within the European common market. Moreover, the results highlight that data sources excluding domestic trade tend to significantly underestimate the overall trade effect of EU membership, as they fail to capture this substitution, a channel that is well accounted for in international trade theory.

The results of the meta-regression analysis in this section illustrate the impact of dataset restrictions on the estimated effect size of the EU membership trade effect. Our estimates indicate that the largest trade effect arises when a researcher utilizes a data source that includes domestic trade flows, spans the longest possible time period segmented into three or five year intervals, uses total trade data across all available industries, or employs highly disaggregated industry-level data. The next section presents some sub-sample analyses and robustness checks for these findings.

4.3.3 Subsample analysis and robustness checks

Table 7 provides an overview of the subsample analysis and additional robustness checks. The first six columns present results for different sub-samples of the EU membership trade effect estimates, while the last two columns exclude outliers from the ITPD-E (see Figure 3 and replace heteroskedasticity-robust standard errors with data source-clustered standard errors. All specifications employ the preferred WLS estimator with data source fixed effects.

Column (1) applies the meta-regression to 421 estimates from datasets restricted to cross-border trade flows, while Column (2) studies estimates from datasets that also incorporate domestic trade. The two columns highlight heterogeneous patterns that contribute to the overall findings presented in Section 4.3.2. The smaller EU membership trade effect estimates for the restricted time period appears to be driven by both types of datasets, those with and without domestic trade flows. In contrast, the larger estimates from datasets with disaggregated industries and three-year interval data originate exclusively from datasets limited to cross-border trade flows. Additionally, the smaller EU membership trade effect estimates observed in datasets covering only manufacturing industries is also driven by datasets that exclusively include cross-border trade flows. When domestic trade flows are added, restricting the sample to manufacturing industries instead results in an average EU membership trade effect estimate that is 16.2 percentage points larger.

Columns (3) and (4) examine a potential heterogeneity in the effects of dataset restrictions between datasets with disaggregated industry-level data for manufacturing industries and country-level aggregated trade flows. Column (3) includes 436 estimates derived from datasets with aggregated data, while Column (4) is based on 150 estimates of the EU membership trade effect from disaggregated datasets. The effect of restricting the analysis to manufacturing industries cannot be identified in Column (4), as disaggregated trade data are only available for manufacturing. The estimates reported in both columns indicate that, unlike the decision to include or exclude domestic trade flows, dataset restrictions do not lead to heterogeneous effect sizes of the EU membership trade effect across datasets with different levels of industrial disaggregation. The estimated parameters for different restrictions exhibit comparable levels of statistical significance and quantitative magnitudes across both columns.

Table 7: Meta-regression analysis: Subsample estimates and robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Subsamples						Drop outliers	Clustered SE
	Dom=0	Dom=1	Ind=0	Ind=1	Manu=0	Manu=1		
Domestic			0.2095***	0.2198***	0.1513***	0.2363***	0.2336***	0.2137***
			(0.0124)	(0.0186)	(0.0211)	(0.0112)	(0.0092)	(0.0525)
WIODcountries	-0.0136**	0.0072	-0.0050	-0.0200	0.0007	-0.0156*	-0.0087	-0.0083*
	(0.0055)	(0.0128)	(0.0078)	(0.0139)	(0.0118)	(0.0080)	(0.0062)	(0.0045)
Industry	0.0328***	-0.0212				0.0222***	0.0225***	0.0198**
	(0.0064)	(0.0145)				(0.0085)	(0.0083)	(0.0087)
Year0014	-0.0733***	-0.0708***	-0.0688***	-0.0775***	-0.0587***	-0.0764***	-0.0750***	-0.0716***
	(0.0065)	(0.0149)	(0.0096)	(0.0145)	(0.0147)	(0.0087)	(0.0073)	(0.0161)
3yearintervals	0.0325***	-0.0026	0.0276***	0.0189	0.0254*	0.0247**	0.0256***	0.0250
	(0.0066)	(0.0150)	(0.0093)	(0.0168)	(0.0142)	(0.0103)	(0.0076)	(0.0169)
Imports	0.0226**		0.0313**	0.0030	0.0493***	-0.0012	0.0225**	0.0223
	(0.0093)		(0.0136)	(0.0132)	(0.0164)	(0.0097)	(0.0093)	(0.0203)
Manu	-0.0974***	0.1616***	-0.0617***				-0.0901***	-0.0575
	(0.0062)	(0.0432)	(0.0117)				(0.0076)	(0.0341)
Package	-0.0118**	0.0051	-0.0023	-0.0038	-0.0047	-0.0061	-0.0038	-0.0070
	(0.0056)	(0.0119)	(0.0078)	(0.0147)	(0.0121)	(0.0079)	(0.0062)	(0.0048)
Constant	0.1081***	0.4385***	0.0820***	0.0549***	0.1077***	0.0209*	0.1043***	0.0828**
	(0.0094)	(0.0144)	(0.0136)	(0.0164)	(0.0159)	(0.0115)	(0.0100)	(0.0302)
Data source FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	421	165	436	150	204	382	562	586
R^2	0.6877	0.7711	0.6695	0.7154	0.6199	0.7632	0.7782	0.6705

Notes: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. All columns apply the WLS estimator with data source fixed effects. Columns (1) to (7) report heteroskedasticity robust standard errors in parentheses. Column (8) reports data source-clustered standard errors in parentheses. Column (1) only accounts for estimate bases solely on cross-border trade flows. Column (2) includes estimates from datasets with domestic trade added to the datasets. Column (3) restrictes the sample of estimates to the ones from aggregated trade flows, column (4) considers estimates based on industry-level disaggregated trade flow data. Column (5) focuses on estimates based on total trade, Column (6) is restricted to estimates from datasets that only include manufacturing trade. Column (7) drops the negative estimates from aggregated ITPD-E data as likely outliers.

We further split our sample of EU membership trade effect estimates into those based exclusively on trade in manufacturing industries and those based on total goods trade. The estimation results for these two robustness checks are presented in Columns (5) and (6) of Table 7. By construction, disaggregated industry-level data are only available in data sources that focus on manufacturing industries. Consequently, we cannot identify an effect for disaggregated data in Column (5). In data sources covering total trade, the impact of including domestic trade flows on the estimated EU membership trade effect is quantitatively smaller, amounting to a 15.1 percentage points increase. Replacing exports with import data is only statistically significant when using datasets that use total trade flows, leading to an EU membership trade effect estimate that is 4.9 percentage points larger. For datasets that contain only manufacturing trade data, restricting the country coverage to that of WIOD results in a 1.5 percentage point reduction in the effect size of the average estimated EU membership trade effect. The effect of other dataset restrictions is homogeneous across datasets with total goods trade versus manufacturing trade flows.

In Column (7), 24 estimates based on aggregated trade data from ITPD-E are excluded. Due to missing data on domestic trade, the simple aggregation of industry-level data to overall manufacturing trade flows introduces measurement errors, generally leading to an under-reporting of domestic trade. As a result, EU membership trade effect estimates from these datasets are negative and statistically highly significant (see Figure 3). However, excluding these outlying estimates does not systematically affect the main findings of our baseline meta-regression analysis. The signs of the parameter estimates and their statistical significance remain unchanged. The quantitative effects of including domestic trade flows and restricting data to manufacturing industries alone both increase slightly in absolute terms, reaching 23.4 and -9 percentage points changes in the average effect size of the EU membership trade effect estimate, respectively.

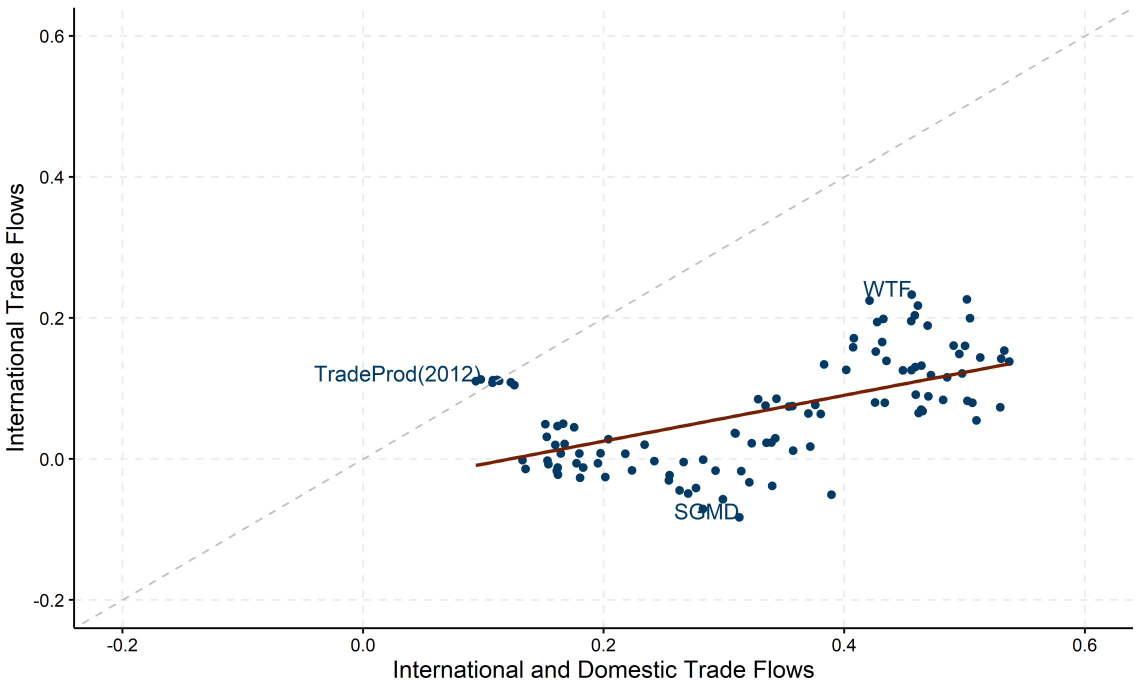
The final column of Table 7 replaces heteroskedasticity-robust standard errors with data source-clustered standard errors while also accounting for data source fixed effects. This adjustment affects the statistical significance of some dataset restriction variables in our meta-regression specification. The impact of using only manufacturing trade data becomes statistically insignificant, whereas restricting the country coverage to WIOD is now statistically significant. The effects from other sample restrictions remain statistically significant, highlighting the robustness of our main findings even under a highly restrictive specification that explicitly accounts for correlation in effect size estimates for the EU membership trade effect within the twelve alternative data sources.

4.3.4 Additional heterogeneity analyses

This section presents additional results from the various data sources used to estimate the trade effect of EU membership, with a focus on potential heterogeneity in effect size estimates across data sources and sample restrictions. In a first step, we graphically examine the role of domestic trade flows in shaping the estimated effect size. For data sources that include domestic trade flows, we are able to estimate two EU membership trade effects for each of the sample restrictions discussed in Section 3.2. For the first estimate we exclude domestic trade, while for the second we include it. Figure 4 plots these EU membership trade effect estimates against each other. The vertical axis represents the effect size based solely on international trade flows, while the horizontal axis shows the estimate size when domestic trade flows are included under the same sample restriction. If domestic trade flows had no impact on the effect size, both estimates would be of equal magnitude and all plotted points would align along the 45-degree line.

The scatter plot in Figure 4 is positioned to the right of the 45-degree line. In almost all cases, the inclusion of domestic trade increases the magnitude of the EU membership trade effect estimate for a given sample restriction. The further a point is located to the right, the larger is the effect once domestic trade added to the data, whereas the higher a point is positioned, the larger the effect based solely on international trade flows is. The only exception is observed for estimates derived from the 2012 version of TradeProd, where the points are closely aligned with the 45-degree line. In this data source, adding domestic trade to cross-border trade flows in the used data does not systematically alter the estimated effect size of EU membership on trade. Additionally, the fitted regression line suggests a slope parameter smaller than one, indicating that the inclusion of domestic trade flows does not simply add a strictly proportional value to the effect size estimate. Instead, it increases the estimated EU membership trade effect in a disproportionate manner.

Figure 4: Estimates without domestic trade versus with domestic trade flows included.



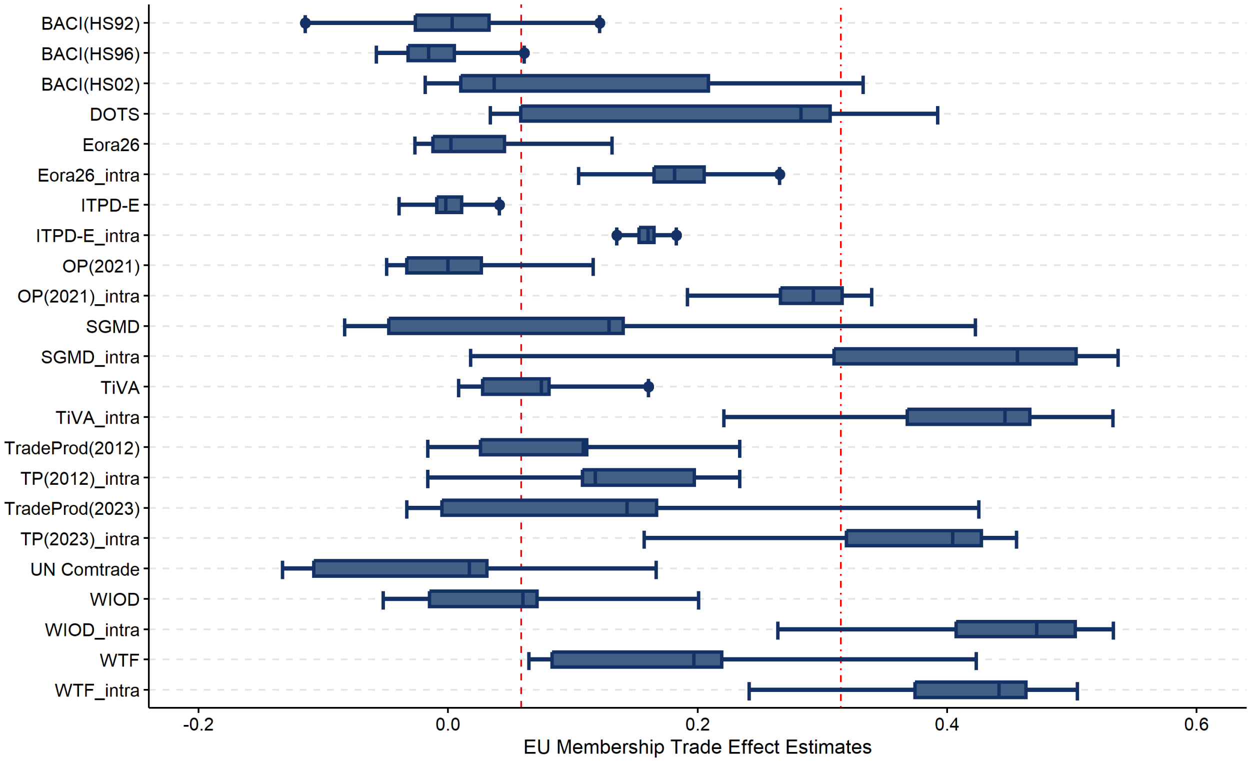
Notes: Each dot represents an estimate for the EU membership trade effect for data without domestic trade flows versus the same dataset restriction but additionally including domestic trade.

Figure 5 provides another comparison of the estimated EU membership trade effects across all data sources using box plots. These estimates are derived from sample restrictions I to VI from Table 2. For data sources that include domestic trade flows, we provide separate box plots for estimates from datasets with and without domestic trade. Consistent with our previous findings, Figure 5 reveals a clear pattern across all databases: incorporating domestic trade flows systematically leads to larger EU membership trade effect estimates. When domestic trade is included, the average estimated effect size amounts to 31.5 percent (dashed line). Estimates based solely on international trade flows take on an average of 5.9 percent (dot-dashed line) which is not statistically distinguishable from zero.

In line with Figure 4, the inclusion of domestic trade flows consistently amplifies the estimated effect size across all databases, albeit to varying degrees. Data from TradeProd(2012) produce the most consistent estimates when comparing results based on international trade flows alone with those that also incorporate domestic trade. Estimates derived from databases utilizing input-output tables (TiVA and WIOD) are smaller then the average of 5.9 percent when only international trade flows are considered, but become larger than the average estimate of 31.5 percent once domestic trade is added to the estimation sample.

Additionally, the results not only vary depending on the inclusion of domestic trade but also differ substantially across databases. Estimates from the DOTS and WTF databases, both of which capture total trade in all goods rather than manufacturing goods alone, tend to be larger than those from other sources. This finding underscores the importance of sectoral heterogeneity in estimating the trade effects of EU membership.

Figure 5: Distribution of EU membership trade effect estimates across data sources and sample restrictions.



Notes: Databases are ordered on the vertical axis. Data sources ending in _intra indicate samples that include domestic trade. The horizontal axis reports box plots for the EU membership trade effect estimates based on dataset restrictions I to VI.. The dashed (red) line marks the average effect size of the EU membership trade estimates using only international trade flows, the (red) dot-dashed line denotes the average estimate once domestic trade flows are included.

Furthermore, estimates derived from older BACI revisions are generally smaller in magnitude, whereas the most recent BACI (HS02) revision yields larger estimates with greater variability across different sample restrictions.

Table 6 illustrates that estimates based on the restricted time period from 2000 to 2014 are smaller than those derived from using all available years in each of the twelve data sources. According to our preferred specification, limiting the observation period to these 15 years reduces the estimated EU membership trade effect by more than seven percentage points. Since the period from 2000 to 2014 encompasses the three most recent eastward enlargement rounds, our findings in Table 6 suggest that the EU membership trade effect might be weaker for the accession of new Eastern European member states. This smaller trade effect is consistent with the results reported in Table 4 of Mika and Zymek (2018), which show that the estimated EU membership trade effect is twice as large when using DOTS data for the period from 1992 to 2002 compared to an extended observational period until 2013.

The twelve alternative data sources cover up to five major EU enlargement rounds, which occurred in 1986, 1995, 2004, 2007, and 2013. The 15 countries that joined the EU by 1995 are commonly referred to as "Old European" member states. Subsequent enlargements primarily involved Central and Eastern European countries, with ten of them joining in 2004, followed by Bulgaria and Romania in 2007 and Croatia in 2013.

To further investigate differences in EU trade effects before and after 1995, we utilize three data sources that cover the longest time periods. The results, reported in Table 8, are based on the country-level specifications of the structural gravity model outlined in Equations (5) and (6). Panel A presents estimates based solely on international trade flows, while for Panel B domestic trade is added to the data. Table 8 suggests that the observed time trend in EU membership trade effect estimates depends on whether domestic trade is included or not. When relying exclusively on international trade flows, the EU trade effect estimates exhibit a clear downward trend after 1995, indicating a decline in its quantitative magnitude. However, this trend reverses when domestic trade is accounted for. Our estimates suggest that trade diversion from domestic trade has been more pronounced for members that joined the EU after 1995. This finding aligns with Spornberger (2022), who emphasizes that excluding domestic trade flows introduces a significant negative bias when analyzing the trade effects of EU eastward enlargement rounds.

Table 8: EU trade effect estimates before and after 1995

	Panel A: International trade flows						*Panel B: Including domestic trade*			
	SGMD		WTF		DOTS		SGMD		WTF	
EU*Border$_{ijt}$ (until 1995)	0.1941***		0.2083***		0.4572***		0.1682***		0.1313***	
	(0.0364)		(0.0365)		(0.0482)		(0.0458)		(0.0297)	
EU*Border$_{ijt}$ (after 1995)		0.0206		0.1087***		0.1463***		0.3967***		0.5011***
		(0.0369)		(0.0359)		(0.0336)		(0.0427)		(0.0452)
FTAs$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Inter$_{ijt}$	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Three-way FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.9934	0.9937	0.9928	0.9922	0.9889	0.9923	0.9988	0.9989	0.9992	0.9983
Obs.	224,831	681,600	191,065	578,698	224,443	487,457	226,344	683,719	192,727	582,347

Notes: *** denotes statistical significance at the 1% level. Estimates reported in Panel A are based on the gravity model Specification (5). Panel B reports estimates from Specification (6). Standard errors are clustered at the exporter-importer pair level and are reported in parentheses.

Our preferred specification in Table 6 suggests that restricting datasets to manufacturing industries reduces the estimated EU membership trade effect by approximately 5.75 percentage points. This finding indicates a heterogeneous EU membership trade effect across industries, which we explore further in Table 9. ITPD-E, TiVA, and WIOD cover agricultural, manufacturing, and services industries, allowing us to separately estimate our structural gravity model specifications for these three broad sectors. These data sources additionally include domestic trade flows, enabling us to compare results when these flows are included or excluded from the estimation. Panel A presents the estimates from Equation (5) and Panel B reports those with domestic trade included from (6).

The findings from Table 9 can be summarized as follows: The estimated EU membership trade effect is largest for agricultural industries, followed by services industries, while it is quantitatively the smallest for manufacturing industries. This result helps to explain why data sources that account for total trade flows, such as DOTS and WTF, yield larger estimates of the EU membership trade effect compared to sources that focus exclusively on manufacturing industries (e.g., SGMD).

The results presented in Panel B of Table 9 further indicate that the inclusion of domestic trade flows has the most significant impact on the estimated effect size of the EU membership trade effect for manufacturing industries. This suggests a substantial substitution of domestic sales by cross-border trade in this sector. When considering only international trade flows, the estimated EU membership trade effect is either statistically not different from zero or of a modest effect size of approximately 9%. Once domestic trade flows are added to the data, these estimates increase to a range of 18.2% to 58.7%. EU membership appears to enhance trade in agricultural products primarily through additional trade creation with other EU members, whereas the EU membership effect in manufactured goods is driven mainly by a substitution of domestic trade with imports. In the case of services, both effects appear to be relevant, as suggested by the third columns for each of the three data sources.

Table 9: EU membership trade effect estimates across three broad sectors

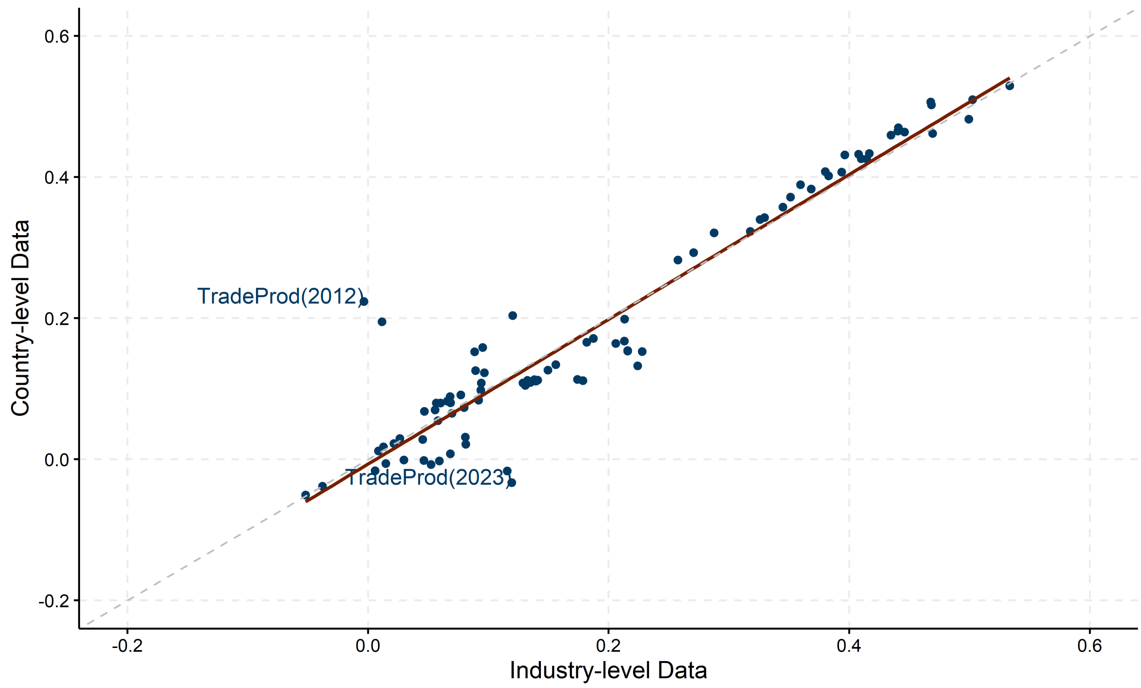
	ITPD-E			TiVA			WIOD		
	Agriculture	Manu	Service	Agriculture	Manu	Service	Agriculture	Manu	Service
Panel A: International trade flows									
EU*Border$_{ijt}$	0.8763***	0.0214	0.1051	0.6735***	0.0870**	0.2360***	0.7267***	0.0652	0.5605***
	(0.0734)	(0.0383)	(0.0991)	(0.0906)	(0.0406)	(0.0315)	(0.1021)	(0.0498)	(0.0769)
FTAs$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Three-way FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.9690	0.9931	0.9759	0.9833	0.9930	0.9932	0.9804	0.9949	0.9717
Obs.	394,056	713,524	73,275	106,128	106,128	106,128	34,890	34,890	34,890
Panel B: Including domestic trade flows									
EU*Border	0.8650***	0.1675***	0.6672***	0.7894***	0.4596***	0.2649***	0.6280***	0.4617***	0.5069***
	(0.0564)	(0.0363)	(0.0717)	(0.0667)	(0.0351)	(0.0242)	(0.0666)	(0.0382)	(0.0425)
FTAs$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Inter$_{ijt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Three-way FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.9965	0.9972	0.9984	0.9970	0.9986	0.9997	0.9978	0.9990	0.9996
Obs.	396,581	714,880	74,909	107,736	107,736	107,736	35,735	35,735	35,735

Notes: *** denotes statistical significance at the 1% level. ** denotes statistical significance at the 5% level. Estimates reported in Panel A are based on the gravity model Specification (5). Panel B reports estimates from Specification (6). Standard errors are clustered at the exporter-importer pair level and are reported in parentheses.

Estimates of the EU membership trade effect based on industry-level data are approximately 2 percentage points larger than those derived from aggregated country-level data (Column (4) of Table 6). Figure 6 provides further evidence on the impact of aggregating trade flow data for estimates of the EU membership trade effect by plotting effect size estimates from disaggregated sources against those obtained after aggregating the same data sources to the bilateral country level.

Figure 6 illustrates that the impact of trade flow data aggregation varies across different data sources and sample restrictions. In samples where the estimated EU membership trade effects are relatively small or close to zero, most estimates tend to be larger for disaggregated datasets, as indicated by their location to the right of the 45-degree line. However, in the case of the 2012 version of TradeProd, the estimates for aggregated data are larger at the lower end of the effect size distribution. For EU membership trade effect estimates exceeding 30%, aggregated trade flow data generally yield slightly larger estimates, as reflected by the observations being located just above the 45-degree line in the upper-right corner of Figure 6. In the vast majority of cases, the differences in effect size estimates for these observations are, however, small and not statistically significant.

Figure 6: Estimates from aggregated country-level data versus disaggregated industry-level data.



Notes: Each dot represent an estimate for the EU membership trade effect obtained using industry-level data versus an estimate from the the same data sources that are aggregated to the country-level.

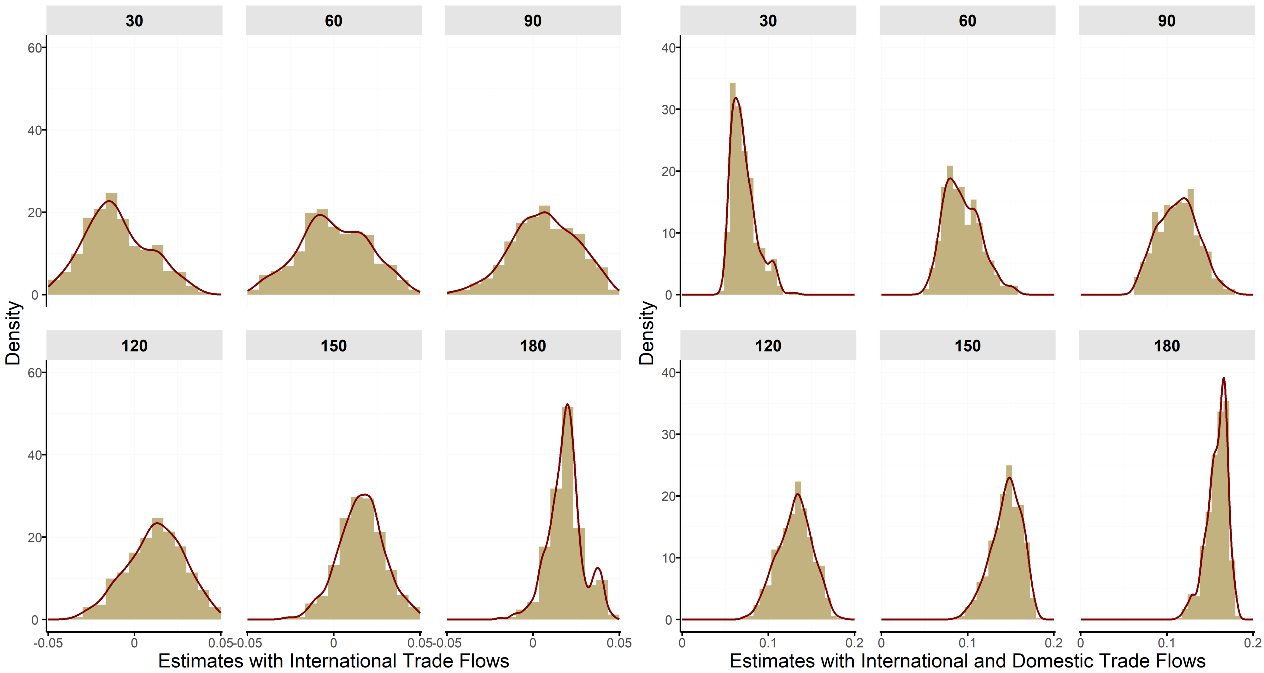
Our main findings indicate that restricting the sample to countries covered by WIOD has only a minor impact on the estimated effect size of the EU membership trade effect. In the meta-regression analysis, the estimates are predominantly negative but generally too small to be statistically distinguishable from zero. Four of our data sources, ITPD-E, TradeProd (2012), Eora26 and WTF, cover substantially larger country samples, allowing us to assess the generalizability of the main finding to other potential country restrictions in the data. We define up to six samples that include all EU member states, adding 30 randomly selected non-EU countries in each step. The smallest sample thus includes 30 countries in the control group, while the largest comprises 180. For each sample size, we conduct 500 iterations of the random selection process to generate a distribution of EU membership trade effect estimates across different control group sizes. Figure 7 visually represents the distribution of these estimates across the six sample sizes and distinguishes between gravity model estimates based on data with and without domestic trade flows.

Figure 7 presents the estimation results for ITPD-E and suggests that limited country coverage tends to underestimate the EU trade effect when domestic trade flows are excluded.[18] This finding is consistent with the meta-regression results discussed above (see Column (1) of Table 7) and the findings reported in Hou (2020). However, when domestic trade flows are included, the EU membership trade effect estimates increase.

The distribution of EU membership trade effect estimates based solely on international trade flows is right-skewed for small sample sizes but gradually shifts to a left-skewed distribution as the sample size increases.

[18]Figures D1 to D3 in Appendix D report the results for TradeProd (2012), Eora26, and WTF.

Figure 7: Distribution of estimates using different sample sizes of ITPD-E.



Notes: The vertical axis is the kernel density. The horizontal axis reports the EU membership trade effect estimates. The number above each plot denotes the number of randomly selected control group countries.

Smaller samples tend to yield lower estimates, while larger samples with a broader country coverage produce larger EU membership trade effect estimates when domestic trade flows are excluded. This finding aligns with Rose (2017)'s study on the effects of the European Monetary Union. Moreover, this result remains robust even when domestic trade flows are included. Additionally, as the number of countries increases, the variation in the estimates decreases. In samples with 180 countries in the control group, the distribution of the estimated EU membership trade effect converges to its average value.

This section provides further evidence on underlying heterogeneity in how data source selection choices and sample restrictions influence the effect size of EU membership trade estimates. Similar sample restrictions lead to varying quantitative impacts for the obtained estimates depending on the data source. In line with our main findings, the decision to include or exclude domestic trade flows in the data has a dominating impact on the estimated effect size, almost always increasing its quantitative magnitude. Moreover, the trade effect of EU membership varies across different time periods. Whether the effects are larger or smaller after 1995 crucially depends on whether domestic trade is considered or not. When domestic trade is included (excluded), the Eastern European enlargement rounds result in larger (smaller) trade effect estimates compared to the early EU enlargements until 1995.

A similar pattern emerges across broad economic sectors. When considering only cross-border trade flows, the estimated EU membership trade effect for manufacturing goods is small and often statistically indistinguishable from zero. In contrast, the effects in the agricultural and services industries are quantitatively larger and less sensitive to the exclusion of domestic trade. Significant positive effect estimates for manufacturing industries only appear once domestic trade is included in the estimation sample. Furthermore, replacing detailed industry-level data with aggregated country-level trade flow data from the same sources leads to only minor differences in the effect size of the estimated EU membership trade effect. Lastly, we find that the number of countries in the control group also influences the effect size. In small samples, the estimates are smaller and skewed to the left, while increasing the number of countries shifts the distribution to the right. Including domestic trade further amplifies this rightward shift.

5 Conclusions

Theoretical modeling and econometric estimation of trade and welfare effects of trade policies have made remarkable progress over the past two decades. The availability of larger and higher-quality panel datasets for trade flows enables a closer and more direct connection of international trade theory with the empirical estimation of structural gravity models. The application of PPML estimators with a comprehensive set of exporter-time, importer-time, and exporter-importer fixed effects effectively accounts for multilateral resistance in international trade relationships. This approach allows the identification of causal partial equilibrium trade effect estimates of trade policies based on empirical specifications of structural gravity model equations. These causal trade effect estimates typically serve as inputs for new quantitative general equilibrium trade models, which are used to assess the welfare implications of trade policies, taking general equilibrium effects into account.

The increasing complexity of available data sources on international and domestic trade flows has sparked a series of contributions to the literature discussing how to estimate and quantify trade policy effects in structural gravity models. Head and Mayer (2014), Bergstrand et al. (2015), Yotov et al. (2016), Yotov (2022), Larch and Yotov (2024), and Larch et al. (2025) provide in-depth discussions of alternative approaches and offer various recommendations for best practices. These recommendations include, among others, the incorporation of domestic trade, the use of consecutive annual panel data for disaggregated industry-specific trade flows covering as many countries as possible, and the application of the PPML estimator with a full set of fixed effects and clustered standard errors.

Most of these recommendations follow from theoretical and econometric reasoning and/or are derived from empirical analyses of effect size heterogeneity in estimates based on a specific gravity model specification for a given dataset. However, the impacts of data source selection choices and imposed sample restrictions has received little attention in the international economics literature. This paper contributes to our understanding of the implications of such choices for the estimated effect size of trade policies by empirically revisiting the trade effects of EU membership, which remains the most comprehensive trade agreement to date.

We have used twelve different and commonly applied data sources and imposed nine different sample restrictions on the data. The resulting samples served as basis for estimating 586 EU membership trade effects using standard gravity model specifications that only vary over the treatment of domestic trade flows and accurately account for the level of disaggregation in each of the samples. In a next stage, we employed a meta-regression and additional heterogeneity analyses to systematically study the role of data source selection and dataset restriction choices for the effect size of the EU membership trade effect estimate.

Our meta-regression analysis reveals significant differences in the effect size of the estimated EU membership trade effect across alternative data sources. The choice of a data source can substantially influence the effect size of the estimate from a structural gravity model specification, with consequences for the welfare implications derived from new quantitative trade models.

In addition to the selection of data sources, dataset restrictions also play a crucial role in shaping the estimated effect size of EU membership for trade. The inclusion of domestic trade flows induces the largest change in the effect size, indicating a notable substitution of domestic trade with trade between EU member states after a country's accession. Furthermore, the country coverage, the level of industrial disaggregation, and the time period considered (i.e., different EU enlargement phases) all have important implications for the magnitude of the EU membership trade effect estimate.

Our findings further show that datasets with a larger number of control group economies tend to produce larger effect size estimates. Moreover, when data sources allow for a detailed breakdown of trade flows at the industry level, the quantitative magnitude of the estimates are also larger, provided that the available level of disaggregation is fully exploited in the analysis. Finally, we find that the EU membership trade effect is significantly different between the early accession rounds and the post-1995 EU eastward enlargement. The estimates are larger for the eastward enlargement rounds when domestic trade flows are included in the data and smaller when they are excluded. Notably, we also identify substantial differences in effect size estimates across the broad sectors of agriculture, manufacturing, and services, further indicating a heterogeneous nature of the EU trade effect.

The findings of this paper may have broader implications for the estimation of trade policies beyond the example of the EU membership trade effect. Our analysis highlights the importance of data source selection and dataset restriction choices for the quantitative magnitude of the estimates obtained from structural gravity model specifications. In some cases, the choice of a particular data source inherently imposes relevant dataset restrictions. Traditional and widely used sources, such as IMF DOTS and UN Comtrade, do not include domestic trade flows. Furthermore, IMF DOTS only provides total bilateral trade flows at the country-pair level without industrial disaggregation. While BACI offers the most detailed product-level bilateral trade flow data, it also lacks information on domestic sales.

This is particularly relevant, as domestic trade flows appear to be crucial for the effect size of estimates for trade policies. While theoretical trade models highlight the substitution of domestic sales with cross-border trade among countries in a trade agreement as a central mechanism, our meta-regression analysis provides empirical support for its quantitative significance. Efforts by data source compilers to further improve the availability and quality of domestic production and trade data would be highly beneficial for enhancing the accuracy of trade policy estimates in structural gravity models.

The international economics literature predominantly studies the impact of trade policies on trade in manufacturing goods. WTO's SGMD, for example, covers only manufacturing industries but augments the Comtrade data with domestic trade. International trade in agricultural goods is considered to be more strongly restricted by national regulation and larger trade barriers. Services trade data are typically of lower quality and are available at only a higher level of aggregation. Furthermore, the services trade modes differ systematically from those for goods trade, which requires alternative economic reasoning and modeling. Our findings, however, suggest that the trade effect of EU membership varies quantitatively across these sectors and focusing solely on manufacturing goods may provide an incomplete picture of the overall gains from trade policies. Initiatives to incorporate more detailed and higher-quality data on trade in agricultural goods and services could thus prove beneficial for empirically studying the quantitative trade effects of trade policies.

We do not find significant differences in the effect size of the estimates when using industry-level data versus aggregated country-pair trade flows, provided that we estimate homogeneous effects across industries. This changes when we allow for heterogeneous effects across the three broad sectors: agriculture, manufacturing and services. The latter finding aligns with recent research on an aggregation bias in gravity model trade flow estimates (Breinlich et al., 2024; French and Zylkin, 2024; Larch and Yotov, 2024). Furthermore, detailed industry-level trade data include larger shares of zero trade flows, which also allows for the study of the impact of a trade policy on the extensive margin of trade (Helpman et al., 2008; Baier et al., 2014; Ad*ão* et al., 2024; French and Zylkin, 2024). The number of included control group countries also affects the estimated effect size of EU membership for trade. Based on our findings, we recommend using detailed industry-level data rather than country-pair aggregates and to estimate heterogeneous trade policy effects across industries for datasets covering as many countries as possible. A promising avenue for data compilers, therefore, may be to increase efforts to provide further disaggregated trade flow data for more countries worldwide.

The findings from this paper suggest that the quantitative magnitude trade effect estimates from trade policies may crucially depend on data source selection and sample restriction choices. To enhance comparability across contributions to the literature, harmonizing data sources might be more promising than further increasing the number of alternative data sources available to researchers. This would enable the community to apply common standards for data source selection and sample restrictions. In our view, the ITPD-E database combines the desired data source characteristics in the most compelling way and could serve as a starting point for efforts to harmonize dataset standards for estimating trade effects of trade policies using econometric specifications of the structural gravity model.

References

Ad*ão*, R., C. Arkolakis, and S. Ganapati (2024). From heterogeneous firms to heterogeneous trade elasticities: The aggregate implications of firm export decisions. Working Paper. `https://arkolakis.com/wp-content/uploads/draft_20241130.pdf`.

Allen, T., C. Arkolakis, and Y. Takahashi (2020). Universal gravity. *Journal of Political Economy 128*(2), 393–433.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review 69*(1), 106–116.

Anderson, J. E. and E. Van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170–192.

Anderson, J. E. and E. Van Wincoop (2004). Trade costs. *Journal of Economic Literature 42*(3), 691–751.

Anderson, J. E. and Y. V. Yotov (2010). The changing incidence of geography. *American Economic Review 100*(5), 2157–86.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Baier, S. L. and J. H. Bergstrand (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics 71*(1), 72–95.

Baier, S. L., J. H. Bergstrand, and M. Feng (2014). Economic integration agreements and the margins of international trade. *Journal of International Economics 93*(2), 339–350.

Bergstrand, J. H., M. W. Clance, and J. M. C. Santos Silva (2025). The tails of gravity: Using expectiles to quantify the trade-margins effects of economic integration agreements. Discussion Papers in Economics DP 01/25, University of Surrey.

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.

Borchert, I. and Y. V. Yotov (2017). Distance, globalization, and international trade. *Economics Eetters 153*, 32–38.

Breinlich, H., D. Novy, and J. M. C. Santos Silva (2024). Trade, gravity, and aggregation. *The Review of Economics and Statistics 106*(5), 1418–1426.

Caliendo, L. and F. Parro (2015). Estimates of the trade and welfare effects of NAFTA. *The Review of Economic Studies 82*(1), 1–44.

Campos, R. G., J. Timini, and E. Vidal (2021). Structural gravity and trade agreements: Does the measurement of domestic trade matter? *Economics Letters 208*, 110080.

Cheng, I.-H. and H. J. Wall (2005). Controlling for heterogeneity in gravity models of trade and integration. *Federal Reserve Bank of St. Louis Review 87*(1), 49–63.

Dai, M., Y. V. Yotov, and T. Zylkin (2014). On the trade-diversion effects of free trade agreements. *Economics Letters 122*(2), 321–325.

De Sousa, J., T. Mayer, and S. Zignago (2012). Market access in global and regional trade. *Regional Science and Urban Economics 42*(6), 1037–1052.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Egger, P. and M. Larch (2008). Interdependent preferential trade agreement memberships: An empirical analysis. *Journal of International Economics 76*(2), 384–399.

Egger, P., M. Larch, and Y. V. Yotov (2022). Gravity estimations with interval data: Revisiting the impact of free trade agreements. *Economica 89*(353), 44–61.

Egger, P. and Y. Wolfmayr (2018). International trade data and empirical patterns. In *Handbook of International Trade and Transportation*, pp. 79–140.

Egger, P. H. and S. Nigai (2015). Structural gravity with dummies only: Constrained ANOVA-type estimation of gravity models. *Journal of International Economics 97*(1), 86–99.

Egger, P. H. and K. E. Staub (2016). GLM estimation of trade gravity models with fixed effects. *Empirical Economics 50*(1), 137–175.

Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics 97*(1), 76–85.

Feenstra, R. C. and J. Romalis (2014). International prices and endogenous quality. *The Quarterly Journal of Economics 129*(2), 477–527.

Felbermayr, G., J. Gröschl, and I. Heiland (2022). Complex europe: Quantifying the cost of disintegration. *Journal of International Economics 138*, 103647.

French, S. and T. Zylkin (2024). The effects of free trade agreements on product-level trade. *European Economic Review 162*, 104673.

Gaulier, G. and S. Zignago (2010). BACI: International trade database at the product-level. the 1994-2007 version. Working Papers 2010-23, CEPII.

Glick, G. and A. K. Rose (2016). Currency unions and trade: A post-EMU reassessment. *European Economic Review 87*, 78–91.

Guilhoto, J. M., C. Webb, and N. Yamano (2022). Guide to OECD TiVA indicators, 2021 edition. OECD Science, Technology and Industry Working Papers No. 2022/02.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of International Economics*, Volume 4, pp. 131–195.

Heid, B., M. Larch, and Y. V. Yotov (2021). Estimating the effects of non-discriminatory trade policies within structural gravity models. *Canadian Journal of Economics 54*(1), 376–409.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics 123*(2), 441–487.

Hou, J. (2020). Revisiting the trade effects of the euro: Data sources and various samples. *Empirical Economics 59*(6), 2731–2777.

Kasy, M. (2021). Of forking paths and tied hands: Selective publication of findings, and what economists should do about it. *Journal of Economic Perspectives 35*(3), 175–92.

Larch, M., J. A. Monteiro, R. Piermartini, and Y. V. Yotov (2025). On the effects of GATT/WTO membership on trade: They are positive and large after all. *Canadian Journal of Economics 58*(1), 281–328.

Larch, M., S. Shikher, and Y. V. Yotov (2025). Estimating gravity equations: Theory implications, econometric developments, and practical recommendations. CGPA Working Paper 2025-01, Drexel University.

Larch, M., J. Wanner, and Y. V. Yotov (2018). Bi- and unilateral trade effects of joining the euro. *Economics Letters 171*, 230–234.

Larch, M., J. Wanner, Y. V. Yotov, and T. Zylkin (2019). Currency unions and trade: A PPML re-assessment with high-dimensional fixed effects. *Oxford Bulletin of Economics and Statistics 81*(3), 487–510.

Larch, M. and Y. V. Yotov (2024). Estimating the effects of trade agreements: Lessons from 60 years of methods and data. *The World Economy 47*(5), 1771–1799.

Lenzen, M., D. Moran, K. Kanemoto, and A. Geschke (2013). Building Eora: A global multi-region input-output database at high country and sector resolution. *Economic Systems Research 25*(1), 20–49.

Marini, M. M., M. R. Dippelsman, and M. M. Stanger (2018). New estimates for direction of trade statistics. IMF Working Papers 2018/016, International Monetary Fund.

Mayer, T., G. Santoni, and V. Vicard (2023). The CEPII trade and production database. CEPII Working Paper 2023-01, CEPII Research Center.

Mika, A. and R. Zymek (2018). Friends without benefits? New EMU members and the "Euro Effect" on trade. *Journal of International Money and Finance 83*, 75–92.

Monteiro, J.-A. (2020). Structural gravity manufacturing sector dataset: 1980-2016. World Trade Organization, Geneva.

Nicita, A. and M. Olarreaga (2007). Trade, production, and protection database, 1976-2004. *World Bank Economic Review 21*(1), 165–171.

Oberhofer, H. and M. Pfaffermayr (2021). Estimating the trade and welfare effects of Brexit: A panel data structural gravity model. *Canadian Journal of Economics 54*(1), 338–375.

Oberhofer, H., M. Pfaffermayr, and R. Sellner (2021). Revisiting time as a trade barrier: Evidence from a panel structural gravity model. *Review of International Economics 29*(5), 1382–1417.

Olivero, M. P. and Y. V. Yotov (2012). Dynamic gravity: Endogenous country size and asset accumulation. *Canadian Journal of Economics 45*(1), 64–92.

Pfaffermayr, M. (2020). Constrained poisson pseudo maximum likelihood estimation of structural gravity models. *International Economics, 161*, 188–198.

Rose, A. K. (2004). Do we really know that the WTO increases trade? *American Economic Review 94*(1), 98–114.

Rose, A. K. (2017). Why do estimates of the EMU effect on trade vary so much? *Open Economies Review 28*(1), 1–18.

Santos Silva, J. M. C. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Santos Silva, J. M. C. and S. Tenreyro (2011). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters 112*(2), 220–222.

Spornberger, J. (2022). EU integration and structural gravity: A comprehensive quantification of the border effect on trade. *Review of International Economics 30*(4), 915–938.

Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives 15*(3), 131–150.

Stanley, T. D. and S. Jarrell (2005). Meta-regression analysis: A quantitative method of literature surveys. *Journal of Economic Surveys 19*(3), 299–308.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. De Vries (2015). An illustrated user guide to the World Input-Output Database: The case of global automotive production. *Review of International Economics 23*(3), 575–605.

Timmer, M. P., B. Los, R. Stehrer, and G. J. de Vries (2016). An anatomy of the global trade slowdown based on the WIOD 2016 release. GGDC research memorandum number 162, University of Groningen.

Weidner, M. and T. Zylkin (2021). Bias and consistency in three-way gravity models. *Journal of International Economics 132*, 103513.

Wolfmayr, Y., K. Friesenbichler, H. Oberhofer, M. Pfaffermayr, I. Siedschlag, M. Di Ubaldo, M. Tong Koeckling, and W. Yang (2019). *The Performance of the Single Market for Goods after 25 Years*. Study for the European Commission, DG Internal Market, Industry, Entrepreneurship and SMEs. http://dx.doi.org/10.2139/ssrn.3462650.

Yotov, Y. V. (2012). A simple solution to the distance puzzle in international trade. *Economics Letters 117*(3), 794–798.

Yotov, Y. V. (2022). On the role of domestic trade flows for estimating the gravity model of trade. *Contemporary Economic Policy 40*(3), 526–540.

Yotov, Y. V., R. Piermartini, and M. Larch (2016). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. WTO iLibrary.

Appendix A Country and industry coverage by data source

Table A1 lists all countries included in each data source. Table A2 provides information on the industries covered and the level of disaggregation available. The seven data sources that include industry-level data use different classification systems: ITPD-E and TiVA follow ISIC Rev.4, WIOD applies ISIC Rev.3, while BACI, TradeProd (2012, 2023), and UN Comtrade use the SITC Rev.3 classification.

Table A1: Country coverage by data source

Databases	Country coverage
BACI	ABW, AFG, AGO, AIA, ALB, AND, ANT, ARE, ARG, ARM, ASM, ATF, ATG, AUS, AUT, AZE, BDI, BEL, BEN, BES, BFA, BGD, BGR, BHR, BHS, BIH, BLM, BLR, BLZ, BMU, BOL, BRA, BRB, BRN, BTN, BWA, CAF, CAN, CCK, CHE, CHL, CHN, CIV, CMR, COD, COG, COK, COL, COM, CPV, CRI, CUB, CUW, CXR, CYM, CYP, CZE, DEU, DJI, DMA, DNK, DOM, DZA, ECU, EGY, ERI, ESP, EST, ETH, FIN, FJI, FLK, FRA, FSM, GAB, GBR, GEO, GHA, GIB, GIN, GMB, GNB, GNQ, GRC, GRD, GRL, GTM, GUM, GUY, HKG, HND, HRV, HTI, HUN, IDN, IND, IOT, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KIR, KNA, KOR, KWT, LAO, LBN, LBR, LBY, LCA, LKA, LSO, LTU, LUX, LVA, MAC, MAR, MDA, MDG, MDV, MEX, MHL, MKD, MLI, MLT, MMR, MNE, MNG, MNP, MOZ, MRT, MSR, MUS, MWI, MYS, MYT, NAM, NCL, NER, NFK, NGA, NIC, NIU, NLD, NOR, NPL, NRU, NZL, OMN, PAK, PAN, PCN, PER, PHL, PLW, PNG, POL, PRK, PRT, PRY, PSE, PYF, QAT, ROU, RUS, RWA, SAU, SCG, SDN, SEN, SGP, SHN, SLB, SLE, SLV, SMR, SOM, SPM, SRB, SSD, STP, SUR, SVK, SVN, SWE, SWZ, SXM, SYC, SYR, TCA, TCD, TGO, THA, TJK, TKL, TKM, TLS, TON, TTO, TUN, TUR, TUV, TZA, UGA, UKR, URY, USA, UZB, VCT, VEN, VGB, VNM, VUT, WLF, WSM, YEM, ZAF, ZMB, ZWE.
DOTS	ABW, AGO, ALB, ARE, ARG, ARM, ATG, AUS, AUT, AZE, BDI, BEL, BEN, BFA, BGD, BGR, BHR, BHS, BIH, BLR, BLZ, BOL, BRA, BRB, BRN, BTN, BWA, CAF, CAN, CHE, CHL, CHN, CIV, CMR, COD, COG, COL, COM, CPV, CRI, CYP, CZE, DEU, DJI, DMA, DNK, DOM, DZA, ECU, EGY, ERI, ESP, EST, ETH, FIN, FJI, FRA, FSM, GAB, GBR, GEO, GHA, GIN, GMB, GNB, GNQ, GRC, GRD, GTM, GUY, HKG, HND, HRV, HTI, HUN, IDN, IND, IRL, IRN, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KIR, KNA, KOR, KWT, LAO, LBN, LBR, LBY, LCA, LKA, LSO, LTU, LUX, LVA, MAC, MAR, MDA, MDG, MDV, MEX, MHL, MKD, MLI, MLT, MMR, MNG, MOZ, MRT, MUS, MWI, MYS, NAM, NER, NGA, NIC, NLD, NOR, NPL, NZL, OMN, PAK, PAN, PER, PHL, PNG, POL, PRT, PRY, PYF, QAT, ROU, RUS, RWA, SAU, SDN, SEN, SGP, SLB, SLE, SLV, SRB, SSD, STP, SUR, SVK, SVN, SWE, SWZ, SYC, SYR, TCD, TGO, THA, TJK, TKM, TLS, TON, TTO, TUN, TUR, TZA, UGA, UKR, URY, USA, UZB, VCT, VEN, VNM, VUT, WSM, YEM, ZAF, ZMB.
Eora26	ABW, AFG, AGO, ALB, AND, ANT, ARE, ARG, ARM, ATG, AUS, AUT, AZE, BDI, BEL, BEN, BFA, BGD, BGR, BHR, BHS, BIH, BLR, BLZ, BMU, BOL, BRA, BRB, BRN, BTN, BWA, CAF, CAN, CHE, CHL, CHN, CIV, CMR, COD, COG, COL, CPV, CRI, CUB, CYM, CYP, CZE, DEU, DJI, DNK, DOM, DZA, ECU, EGY, ERI, ESP, EST, ETH, FIN, FJI, FRA, GAB, GBR, GEO, GHA, GIN, GMB, GRC, GRL, GTM, GUY, HKG, HND, HRV, HTI, HUN, IDN, IND, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KOR, KWT, LAO, LBN, LBR, LBY, LIE, LKA, LSO, LTU, LUX, LVA, MAC, MAR, MCO, MDA, MDG, MDV, MEX, MKD, MLI, MLT, MMR, MNE, MNG, MOZ, MRT, MUS, MWI, MYS, NAM, NCL, NER, NGA, NIC, NLD, NOR, NPL, NZL, OMN, PAK, PAN, PER, PHL, PNG, POL, PRK, PRT, PRY, PSE, PYF, QAT, ROU, ROW, RUS, RWA, SAU, SDS, SEN, SGP, SLE, SLV, SMR, SOM, SRB, STP, SUD, SUR, SVK, SVN, SWE, SWZ, SYC, SYR, TCD, TGO, THA, TJK, TKM, TTO, TUN, TUR, TWN, TZA, UGA, UKR, URY, USA, USR, UZB, VEN, VGB, VNM, VUT, WSM, YEM, ZAF, ZMB, ZWE.
ITPD	ABW, AFG, AGO, AIA, ALB, AND, ANT, ARE, ARG, ARM, ASM, ATA, ATF, ATG, AUS, AUT, AZE, BDI, BEL, BEN, BES, BFA, BGD, BGR, BHR, BHS, BIH, BLM, BLR, BLZ, BMU, BOL, BRA, BRB, BRN, BTN, BVT, BWA, CAF, CAN, CCK, CHE, CHL, CHN, CIV, CMR, COD, COG, COK, COL, COM, CPV, CRI, CUB, CUW, CXR, CYM, CYP, CZE, DEU, DJI, DMA, DNK, DOM, DZA, ECU, EGY, ERI, ESH, ESP, EST, ETH, FIN, FJI, FLK, FRA, FRO, FSM, GAB, GBR, GEO, GHA, GIB, GIN, GMB, GNB, GNQ, GRC, GRD, GRL, GTM, GUM, GUY, HKG, HMD, HND, HRV, HTI, HUN, IDN, IND, IOT, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KIR, KNA, KOR, KWT, LAO, LBN, LBR, LBY, LCA, LKA, LSO, LTU, LUX, LVA, MAC, MAR, MDA, MDG, MDV, MEX, MHL, MKD, MLI, MLT, MMR, MNE, MNG, MNP, MOZ, MRT, MSR, MUS, MWI, MYS, MYT, NAM, NCL, NER, NFK, NGA, NIC, NIU, NLD, NOR, NPL, NRU, NZL, OMN, PAK, PAN, PCN, PER, PHL, PLW, PNG, POL, PRK, PRT, PRY, PSE, PYF, QAT, ROU, RUS, RWA, SAU, SCG, SDN, SEN, SGP, SGS, SHN, SLB, SLE, SLV, SMR, SOM, SPM, SRB, SSD, STP, SUR, SVK, SVN, SWE, SWZ, SXM, SYC, SYR, TCA, TCD, TGO, THA, TJK, TKL, TKM, TLS, TON, TTO, TUN, TUR, TUV, TWN, TZA, UGA, UKR, UMI, URY, USA, UZB, VAT, VCT, VEN, VGB, VNM, VUT, WLF, WSM, YEM, ZAF, ZMB, ZWE.
OP (2021)	ALB, AUS, AUT, BGR, BLX, BRA, CAN, CHE, CHN, COL, CRI, CYP, CZE, DEU, DNK, EGY, ESP, EST, ETH, FIN, FRA, GBR, GRC, HUN, IDN, IND, IRL, ISL, ISR, ITA, JOR, JPN, KAZ, KEN, KGZ, KOR, LKA, LTU, LVA, MAR, MDA, MEX, MKD, MUS, MWI, MYS, NLD, NOR, NZL, PAN, PER, PHL, POL, PRT, ROU, RUS, SVK, SVN, SWE, TUR, TZA, UKR, URY, USA, ZAF.
SGMD	ABW, AFG, AGO, AIA, ALB, AND, ANT, ARE, ARG, ARM, ATG, AUS, AUT, AZE, BDI, BEN, BFA, BGD, BGR, BHR, BHS, BIH, BLR, BLX, BLZ, BMU, BOL, BRA, BRB, BRN, BTN, BWA, CAF, CAN, CCK, CHE, CHL, CHN, CIV, CMR, COG, COK, COL, COM, CPV, CRI, CSK, CUB, CXR, CYM, CYP, CZE, DEU, DJI, DMA, DNK, DOM, DZA, ECU, EGY, ERI, ESH, ESP, EST, ETF, ETH, FIN, FJI, FLK, FRA, FRO, FSM, GAB, GBR, GEO, GHA, GIB, GIN, GLP, GMB, GNB, GNQ, GRC, GRD, GRL, GTM, GUF, GUY, HKG, HND, HRV, HTI, HUN, IDN, IND, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KIR, KNA, KOR, KWT, LAO, LBN, LBR, LBY, LCA, LKA, LSO, LTU, LVA, MAC, MAR, MDA, MDG, MDV, MEX, MHL, MKD, MLI, MLT, MMR, MNE, MNG, MNP, MOZ, MRT, MSR, MTQ, MUS, MWI, MYS, NAM, NCL, NER, NFK, NGA, NIC, NIU, NLD, NOR, NPL, NRU, NZL, OMN, PAK, PAN, PCN, PER, PHL, PLW, PNG, POL, PRK, PRT, PRY, PSE, PYF, QAT, REU, ROU, RUS, RWA, SAU, SCG, SDN, SEN, SGP, SHN, SLB, SLE, SLV, SMR, SOM, SPM, SRB, STP, SUR, SVK, SVN, SVU, SWE, SWZ, SYC, SYR, TCA, TCD, TGO, THA, TJK, TKL, TKM, TMP, TON, TTO, TUN, TUR, TUV, TWN, TZA, UGA, UKR, URY, USA, UZB, VCT, VEN, VGB, VNM, VUT, WLF, WSM, YDR, YEM, YUG, ZAF, ZAR, ZMB, ZWE.
TiVA	ARG, AUS, AUT, BEL, BGR, BRA, BRN, CAN, CHE, CHL, CHN, COL, CRI, CYP, CZE, DEU, DNK, ESP, EST, FIN, FRA, GBR, GRC, HKG, HRV, HUN, IDN, IND, IRL, ISL, ISR, ITA, JPN, KAZ, KHM, KOR, LAO, LTU, LUX, LVA, MAR, MEX, MLT, MMR, MYS, NLD, NOR, NZL, PER, PHL, POL, PRT, ROU, ROW, RUS, SAU, SGP, SVK, SVN, SWE, THA, TUN, TUR, TWN, USA, VNM, ZAF.
TradeProd	Both version: AFG, AGO, ALB, ARE, ARG, ARM, AUS, AUT, AZE, BDI, BEN, BFA, BGD, BGR, BHR, BHS, BIH, BLR, BLZ, BMU, BOL, BRA, BRB, BRN, CAF, CAN, CHE, CHL, CHN, CIV, CMR, COG, COL, CPV, CRI, CUB, CYP, CZE, DEU, DNK, DOM, DZA, ECU, EGY, ERI, ESP, EST, ETH, FIN, FJI, FRA, GAB, GBR, GEO, GHA, GMB, GRC, GTM, HKG, HND, HRV, HTI, HUN, IDN, IND, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KOR, KWT, LAO, LBN, LBR, LBY, LCA, LKA, LTU, LVA, MAC, MAR, MDA, MDG, MDV, MEX, MKD, MLT, MMR, MNE, MNG, MOZ, MUS, MWI, MYS, NER, NFK, NGA, NIC, NLD, NOR, NPL, NZL, OMN, PAK, PER, PHL, PNG, POL, PRT, PRY, QAT, ROU, RUS, RWA, SAU, SCG, SDN, SEN, SGP, SLV, SOM, SRB, SUN, SUR, SVK, SVN, SWE, SWZ, SYR, THA, TJK, TON, TTO, TUN, TUR, TZA, UGA, UKR, URY, USA, UZB, VEN, VNM, YEM, ZAF, ZMB, ZWE. 2012 version only: ABW, AIA, AND, ANT, ASM, ATA, ATF, ATG, BA1, BLX, BTN, BVT, CCK, COK, COM, CSH, CXR, CYM, DJI, DMA, ESH, ET1, FLK, FSM, GIB, GIN, GNB, GNQ, GRD, GRL, GUM, GUY, HMD, IOT, KIR, KNA, MCO, MHL, MLI, MNP, MRT, MSR, NCL, NIU, NRU, PAN, PCN, PLW, PRK, PYF, SGS, SHN, SLB, SLE, SMR, SPM, STP, SYC, TCA, TCD, TGO, TKL, TKM, TMP, TUV, TWN, UMI, VAT, VCT, VGB, VIR, VUT, WLF, WSM, YU1, ZAR. 2023 version only: BEL, BWA, CSK, DDR, LSO, LUX, NAM, PSE.
UN Comtrade	ABW, AFG, AGO, AIA, ALB, AND, ANT, ARE, ARG, ARM, ATG, AUS, AUT, AZE, BDI, BEL, BEN, BFA, BGD, BGR, BHR, BHS, BIH, BLR, BLZ, BMU, BOL, BRA, BRB, BRN, BTN, BWA, CAF, CAN, CHE, CHL, CHN, CIV, CMR, COD, COG, COK, COL, COM, CPV, CRI, CUB, CYP, CZE, DEU, DJI, DMA, DNK, DOM, DZA, ECU, EGY, ERI, ESP, EST, ETH, FIN, FJI, FRA, FRO, FSM, GAB, GBR, GEO, GHA, GIN, GLP, GMB, GNB, GRC, GRD, GRL, GTM, GUF, GUY, HKG, HND, HRV, HTI, HUN, IDN, IND, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KIR, KNA, KOR, KWT, LAO, LBN, LBY, LCA, LKA, LSO, LTU, LUX, LVA, MAC, MAR, MDA, MDG, MDV, MEX, MKD, MLI, MLT, MMR, MNE, MNG, MOZ, MRT, MSR, MTQ, MUS, MWI, MYS, MYT, NAM, NCL, NER, NGA, NIC, NLD, NOR, NPL, NZL, OMN, PAK, PAN, PER, PHL, PLW, PNG, POL, PRT, PRY, PSE, PYF, QAT, REU, ROU, RUS, RWA, SAU, SCG, SDN, SEN, SGP, SLB, SLE, SLV, SRB, STP, SUR, SVK, SVN, SWE, SWZ, SYC, SYR, TCA, TGO, THA, TJK, TKM, TLS, TON, TTO, TUN, TUR, TUV, TZA, UGA, UKR, URY, USA, UZB, VCT, VEN, VNM, VUT, WSM, YEM, ZAF, ZMB, ZWE. (the floowing countries are only included in importers) ASM, ATA, ATF, BES, BLM, BVT, CCK, CSK, CUW, CXR, CYM, DDR, ESH, FLK, GIB, GNQ, GUM, HMD, IOT, LBR, MHL, MNP, NFK, NIU, NRU, PCI, PCN, PRK, SGS, SHN, SMR, SOM, SPM, SSD, SUN, SXM, TCD, TKL, UMI , VAT, VGB, WLF, YMD, YUG.
WIOD	AUS, AUT, BEL, BGR, BRA, CAN, CHE, CHN, CYP, CZE, DEU, DNK, ESP, EST, FIN, FRA, GBR, GRC, HRV, HUN, IDN, IND, IRL, ITA, JPN, KOR, LTU, LUX, LVA, MEX, MLT, NLD, NOR, POL, PRT, ROU, RUS, SVK, SVN, SWE, TUR, TWN, USA.
WTF	ABW, AFG, AGO, AIA, ALB, AND, ANT, ARE, ARG, ARM, ASM, ATA, ATF, ATG, AUS, AUT, AZE, BDI, BEL, BEN, BES, BFA, BGD, BGR, BHR, BHS, BIH, BLM, BLR, BLX, BLZ, BMU, BOL, BRA, BRB, BRN, BRX, BTN, BUN, BVT, BWA, CAF, CAN, CCK, CHE, CHL, CHN, CIV, CMR, COD, COG, COK, COL, COM, CPV, CRI, CSK, CUB, CUW, CXR, CYM, CYP, CZE, DDR, DEU, DJI, DMA, DNK, DOM, DZA, ECU, EGY, ERI, ESH, ESP, EST, ETH, FIN, FJI, FLK, FRA, FRE, FRO, FSM, GAB, GBR, GEO, GHA, GIB, GIN, GLP, GMB, GNB, GNQ, GRC, GRD, GRL, GTM, GUF, GUM, GUY, HKG, HMD, HND, HRV, HTI, HUN, IDN, IND, IOT, IRL, IRN, IRQ, ISL, ISR, ITA, JAM, JOR, JPN, KAZ, KEN, KGZ, KHM, KIR, KNA, KOR, KWT, LAO, LBN, LBR, LBY, LCA, LKA, LSO, LTU, LUX, LVA, MAC, MAR, MDA, MDG, MDV, MEX, MHL, MKD, MLI, MLT, MMR, MNE, MNG, MNP, MOZ, MRT, MSR, MTQ, MUS, MWI, MYS, MYT, NAM, NCL, NER, NFK, NGA, NIC, NIU, NLD, NOR, NPL, NRU, NZE, NZL, OMN, PAK, PAN, PCI, PCN, PER, PHL, PLW, PNG, POL, PRK, PRT, PRY, PSE, PYF, QAT, REU, ROU, RUS, RWA, SAU, SCG, SDN, SEN, SGP, SGS, SHN, SLB, SLE, SLV, SMR, SOM, SPE, SPM, SRB, SSD, STP, SUR, SVK, SVN, SVU, SWE, SWZ, SXM, SYC, SYR, TCA, TCD, TGO, THA, TJK, TKL, TKM, TLS, TON, TTO, TUN, TUR, TUV, TWN, TZA, UGA, UKR, UMI, URY, USA, UZB, VAT, VCT, VEN, VGB, VNM, VUT, WLF, WSM, YDR, YEM, YUG, ZAF, ZMB, ZWE.

Table A2: Industry breakdown by data source

Databases	Industries
BACi & UN Comtrade	Organic chemicals, Inorganic chemicals, Dyeing, tanning and colouring materials, Medicinal and pharmaceutical products, Essential oils for perfume materials and cleaning preparations, Fertilizers other than group 272, Plastics in primary forms, Plastics in non-primary forms, Chemical materials and products, n.e.s., Leather, leather manufactures and dressed furskins, Rubber manufactures, n.e.s., Cork and wood manufactures (excluding furniture), Paper and paper manufactures, Textile yarn and related products, Non metallic mineral manufactures, n.e.s., Iron and steel, Non-ferrous metals, Manufactures of metal, n.e.s., Power generating machinery and equipment, Specialised machinery, Metal working machinery, Other industrial machinery and parts, Office machines and automatic data processing machines, Telecommunication and sound recording apparatus, Electrical machinery, apparatus and appliances, n.e.s., Road vehicles, Other transport equipment, Prefabricated buildings, sanitary, heating and lighting fixtures, n.e.s, Furniture and parts thereof, Travel goods, handbags, etc., Articles of apparel & clothing accessories, Footwear, Professional and scientific instruments, n.e.s., Photo apparatus, optical goods, watches and clocks, Miscellaneous manufactured articles, n.e.s.
ITPD	Accumulators primary cells and batteries, Agricultural and forestry machinery, Aircraft and spacecraft, Articles of concrete cement and plaster, Automobile bodies trailers & semi-trailers, Bakery products, Basic chemicals except fertilizers, Basic iron and steel, Basic precious and non-ferrous metals, Bearings gears gearing & driving elements, Bicycles and invalid carriages, Builders' carpentry and joinery, Building and repairing of ships, Building/repairing of pleasure/sport. boats, Carpets and rugs, Casting of iron and steel, Cement lime and plaster, Cocoa chocolate and sugar confectionery, Coke oven products, Cordage rope twine and netting, Corrugated paper and paperboard, Cutlery hand tools and general hardware, Cutting shaping & finishing of stone, Dairy products, Distilling rectifying & blending of spirits, Domestic appliances n.e.c., Dressing & dyeing of fur; processing of fur, Electric motors generators and transformers, Electricity distribution & control apparatus, Electronic valves tubes etc., Engines & turbines (not for transport equipment), Fertilizers and nitrogen compounds, Food/beverage/tobacco processing machinery, Footwear, Furniture, Games and toys, Glass and glass products, Grain mill products, Insulated wire and cable, Jewellery and related articles, Knitted and crocheted fabrics and articles, Lifting and handling equipment, Lighting equipment and electric lamps, Luggage handbags etc.; saddlery & harness, Macaroni noodles & similar products, Machine tools, Machinery for metallurgy, Machinery for mining & construction, Machinery for textile apparel and leather, Made-up textile articles except apparel, Malt liquors and malt, Man-made fibres, Measuring/testing/navigating appliances etc., Medical surgical and orthopaedic equipment, Motor vehicles, Motorcycles, Musical instruments, Office accounting and computing machinery, Optical instruments & photographic equipment, Other articles of paper and paperboard, Other chemical products n.e.c., Other electrical equipment n.e.c., Other fabricated metal products n.e.c., Other food products n.e.c., Other general purpose machinery, Other manufacturing n.e.c., Other non-metallic mineral products n.e.c., Other publishing, Other rubber products, Other special purpose machinery, Other textiles n.e.c., Other transport equipment n.e.c., Other wood products; articles of cork/straw, Ovens furnaces and furnace burners, Paints varnishes printing ink and mastics, Parts/accessories for automobiles, Pesticides and other agro-chemical products, Pharmaceuticals medicinal chemicals etc., Plastic products, Plastics in primary forms; synthetic rubber, Pottery china and earthenware, Prepared animal feeds, Printing, Processing of nuclear fuel, Processing/preserving of fish, Processing/preserving of fruit & vegetables, Processing/preserving of meat, Publishing of books and other publications, Publishing of newspapers journals etc., Publishing of recorded media, Pulp paper and paperboard, Pumps compressors taps and valves, Railway/tramway locomotives & rolling stock, Refined petroleum products, Refractory ceramic products, Reproduction of recorded media, Rubber tyres and tubes, Sawmilling and planing of wood, Service activities related to printing, Soap cleaning & cosmetic preparations, Soft drinks; mineral waters, Sports goods, Starches and starch products, Steam generators, Struct.non-refractory clay; ceramic products, Structural metal products, Sugar, TV and radio receivers and associated goods, TV/radio transmitters; line comm. apparatus, Tanks reservoirs and containers of metal, Tanning and dressing of leather, Textile fibre preparation; textile weaving, Tobacco products, Vegetable and animal oils and fats, Veneer sheets plywood particle board etc., Watches and clocks, Weapons and ammunition, Wearing apparel except fur apparel, Wines, Wooden containers
TiVA	Basic metals, Chemical and chemical products, Coke and refined petroleum products, Computer, electronic and optical equipment, Electrical equipment, Fabricated metal products, Food products, beverages and tobacco, Machinery and equipment, nec, Manufacturing nec; repair and installation of machinery and equipment, Motor vehicles, trailers and semi-trailers, Other non-metallic mineral products, Other transport equipment, Paper products and printing, Pharmaceuticals medicinal chemical and botanical products, Rubber and plastics products, Textiles, textile products, leather and footwear, Wood and products of wood and cork
TradeProd (2012)	Food, Beverage, Tobacco, Textiles, Wearing apparel, except footwear, Leather and products of leather, leather substitutes and fur, Footwear, except vulcanized or moulded rubber or plastic footwear, Wood and cork products, except furniture, Furniture and fixtures, except primarily of metal, Paper and paper products, Printing, publishing and allied industries, Industrial chemicals, Other chemical products, Petroleum refineries, Rubber products, Plastic products not elsewhere classified, Pottery, china and earthenware, Glass and glass products, Other non-metallic mineral products, Iron and steel basic industries, Non-ferrous metal basic industries, Fabricated metal products, except machinery and equipment, Machinery except electrical, Electrical machinery apparatus, appliances and supplies, Transport equipment, Professional and scientific, and measuring and controlling equipment not elsewhere classified, and of photographic and optical goods
TradeProd (2023)	Food, Textiles, Wood-Paper, Chemicals, Minerals, Metals, Machines, Vehicles, Other
WIOD	Basic and fabricated metals, Chemicals and pharmaceuticals, Coke and refined petroleum products , Computer, electronic, electrical and optical prod., Food, beverages, tobacco, Furniture, other manufacturing, Machinery, Non-metallic mineral products, Pulp, paper, printing, publishing, Rubber and plastic products, Textiles, wearing apparel, leather prod., Transport equipment, Wood and products of wood and cork

Appendix B Number of estimates by data source and sample restrictions

Table B1: Number of estimates from different data sources and sample restrictions

Samples/Databases		I	II	III	IV	V	VI	VII	VIII	IX	Total
Panel A: International trade flows											
Country-level databases	BACI (HS92)	2	2	2	2	2	2	-	8	20	314
	BACI (HS96)	2	2	2	2	2	2	-	8	20	
	BACI (HS02)	2	2	2	2	2	2	-	8	20	
	DOTS	2	2	2	2	2	2	12	-	-	
	Eora26	2	2	2	2	2	2	-	-	-	
	ITPD	2	2	2	2	2	2	-	-	12	
	OP (2021)	2	2	2	2	-	-	-	-	-	
	SGMD	2	2	2	2	2	2	-	-	-	
	TiVA	2	2	2	2	2	2	-	-	12	
	TradeProd (2012)	2	2	2	2	2	2	-	-	-	
	TradeProd (2023)	2	2	2	2	2	2	-	-	-	
	UNComtrade	2	2	2	2	2	2	12	-	24	
	WIOD	2	2	-	-	2	-	-	-		
	WTF	2	2	2	2	2	2	-	-	-	
Industry-level databases	BACI (HS92)	1	1	2	2	1	2	-	6	-	107
	BACI (HS96)	1	1	2	2	1	2	-	6	-	
	BACI (HS02)	1	1	2	2	1	2	-	6	-	
	ITPD	1	1	1	1	1	1	-	-	-	
	TiVA	2	2	2	2	2	2	-	-	-	
	TradeProd (2012)	1	1	2	2	2	2	-	-	-	
	TradeProd (2023)	1	1	2	2	2	2	-	-	-	
	UNComtrade	1	1	2	2	1	2	9	-	-	
	WIOD	2	2	-	-	2	-	-	-	-	
Panel B: Including domestic trade flows											
Country-level databases	Eora26	2	2	2	2	2	2	-	-	-	122
	ITPD	2	2	2	2	2	2	-	-	12	
	OP (2021)	2	2	2	2	-	-	-	-	-	
	SGMD	2	2	2	2	2	2	-	-	-	
	TiVA	2	2	2	2	2	2	-	-	12	
	TradeProd (2012)	2	2	2	2	2	2	-	-	-	
	TradeProd (2023)	2	2	2	2	2	2	-	-	-	
	WIOD	2	2	-	-	2	-	-	-		
	WTF	2	2	2	2	2	2	-	-	-	
Industry-level databases	ITPD	1	1	1	1	1	1	-	-	-	43
	TiVA	2	2	2	2	2	2	-	-	-	
	TradeProd (2012)	1	1	2	2	1	2	-	-	-	
	TradeProd (2023)	1	1	2	2	2	2	-	-	-	
	WIOD	2	2	-	-	2	-	-	-	-	
Total											586

Notes: Sample I includes the original datasets, Sample II is the sub-sample from 2000 to 2014, Sample III only includes WIOD countries, sample IV includes WIOD countries from 2000 to 2014. Sample V uses three-year intervals and Sample VI additionally restricts the country sample to WIOD countries. Sample VII uses imports as dependent variables. Sample VIII uses the three different revisions of the HC classifications in the BACA dataset. Sample IX uses total trade from BACI and UN Comtrade databases instead of total manufacturing trade. Please see Table 2 for more details.

Appendix C Simply heterogeneity tests for effect size heterogeneity

Table C2: Test statistics for effect size heterogeneity across data sources and sample restrictions

	Number of estimates	Cochran Q test	I^2 statistic	H^2 statistic
All estimates	586	10365.58***	95.55%	22.48
BACI (HS92)	55	183.88***	71.61%	3.52
BACI (HS96)	55	246.74***	83.30%	5.99
BACI (HS02)	55	281.87***	84.45%	6.43
DOTS	24	158.85***	87.67%	8.11
Eora26	24	768.49***	97.72%	43.93
ITPD	60	492.56***	93.46%	15.29
OP (2021)	16	168.49***	91.63%	11.95
SGMD	24	376.23***	95.41%	21.77
TiVA	72	2129.83***	96.69%	30.19
TradeProd (2012)	43	155.03***	74.62%	3.94
TradeProd (2023)	44	628.97***	94.42%	17.93
UN Comtrade	66	312.53***	82.60%	5.75
WIOD	24	798.89***	97.22%	35.94
WTF	24	209.11***	91.04%	11.16
Test of group differences		259.16***		
International trade flows	421	3036.21***	88.37%	8.60
Include domestic trade flows	165	3020.01***	97.05%	33.91
Test of group differences		139.16***		
Country-level data	436	6172.33***	94.77%	19.11
Industry-level data	150	4136.34***	96.80%	31.27
Test of group differences		0.04		
Exports as dependent variables	541	9682.48***	95.65%	22.98
Imports as dependent variables	45	634.18***	93.92%	16.44
Test of group differences		2.46		
All industries	204	3371.01***	95.99%	24.83
Manufacturing	382	6967.35***	95.22%	20.91
Test of group differences		2.94*		
HS 1992 Rev.	55	183.88***	71.61%	3.52
HS 1996 Rev.	55	246.74***	83.30%	5.99
HS 2002 Rev.	55	281.87***	84.45%	6.43
Test of group differences		11.64***		

Notes: ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. The null hypothesis of Cochran Q test is that the effect sizes are homogeneous within data sources and sample restrictions. The range of I^2 statistic is defined between 0% and 100%. Larger values indicate more dispersion in the effect size. The H^2 statistic takes on a value of zero in the case of effect size homogeneity.

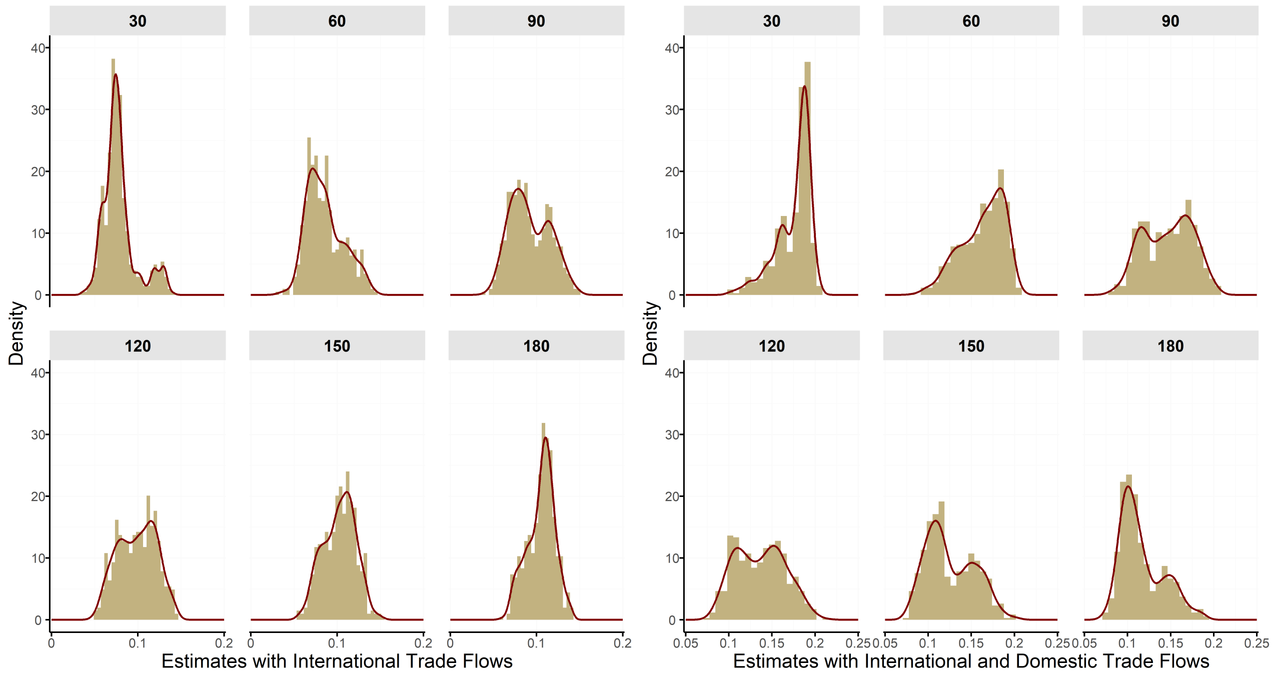
Appendix D Alternative time intervals and country coverage

Table D1: Meta-regression analysis: Four and five year interval data

Variables	(1) OLS	(2) WLS	(3) FE	(4) WLS	(5) OLS	(6) WLS	(7) FE	(8) WLS
Domestic	0.2100***	0.2295***	0.2070**	0.2141***	0.1864***	0.2108***	0.2037**	0.2133***
	(0.0197)	(0.0145)	(0.0679)	(0.0107)	(0.0206)	(0.0144)	(0.0719)	(0.0106)
WIODcountries	-0.0148	-0.0158*	-0.0144***	-0.0092	-0.0116	-0.0121	-0.0139**	-0.0077
	(0.0115)	(0.0095)	(0.0044)	(0.0066)	(0.0118)	(0.0094)	(0.0046)	(0.0070)
Industry	0.0289**	0.0224**	0.0344	0.0265***	0.0057	0.0170*	0.0087	0.0178**
	(0.0114)	(0.0100)	(0.0244)	(0.0083)	(0.0112)	(0.0099)	(0.0093)	(0.0084)
Year0014	-0.0818***	-0.0787***	-0.0830***	-0.0786***	-0.0794***	-0.0753***	-0.0765***	-0.0721***
	(0.0125)	(0.0108)	(0.0166)	(0.0083)	(0.0126)	(0.0106)	(0.0174)	(0.0082)
4yearintervals	-0.0218	-0.0232*	-0.0228	-0.0271***				
	(0.0146)	(0.0122)	(0.0149)	(0.0077)				
5yearintervals					0.0258*	0.0208*	0.0251	0.0195**
					(0.0150)	(0.0123)	(0.0233)	(0.0084)
Imports	0.0406*	0.0293	0.0264	0.0210*	0.0254	0.0123	0.0204	0.0116
	(0.0216)	(0.0199)	(0.0246)	(0.0117)	(0.0200)	(0.0185)	(0.0288)	(0.0119)
Manu	-0.0296*	-0.0369***	-0.0299	-0.0542***	-0.0297*	-0.0374***	-0.0316	-0.0521***
	(0.0155)	(0.0118)	(0.0500)	(0.0103)	(0.0163)	(0.0117)	(0.0516)	(0.0112)
Package	-0.0032	0.0006	-0.0095	-0.0081	-0.0030	0.0003	-0.0095	-0.0080
	(0.0115)	(0.0095)	(0.0061)	(0.0067)	(0.0119)	(0.0095)	(0.0068)	(0.0070)
IO	0.0633***	0.0241*			0.0382**	-0.0010		
	(0.0150)	(0.0127)			(0.0159)	(0.0132)		
Constant	0.1036***	0.1090***	0.1210***	0.0719***	0.1202***	0.1209***	0.1279***	0.0839***
	(0.0158)	(0.0136)	(0.0273)	(0.0125)	(0.0157)	(0.0131)	(0.0248)	(0.0130)
Datasource FEs	No	No	Yes	Yes	No	No	Yes	Yes
Obs.	569	569	569	569	569	569	569	569
R^2	0.3890	0.4770	0.4090	0.7098	0.3299	0.4420	0.3907	0.6627

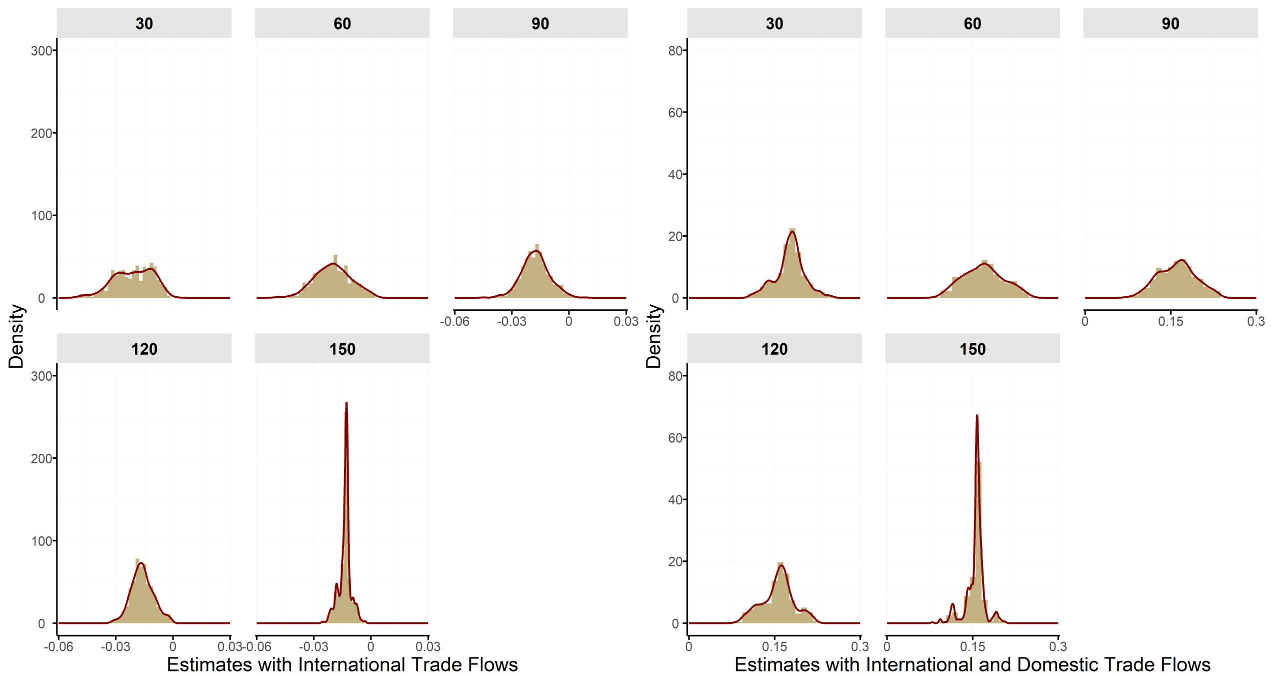
Notes: ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. Heteroskedasticity robust standard errors in parentheses. Columns (1) to (4) are based on four year interval data, column (5) to (8) use five year intervals.OP (2021) is excluded as it is already constructed as three-year interval dataset in the raw data.

Figure D1: Distribution of estimates using different sample sizes of TradeProd (2012)



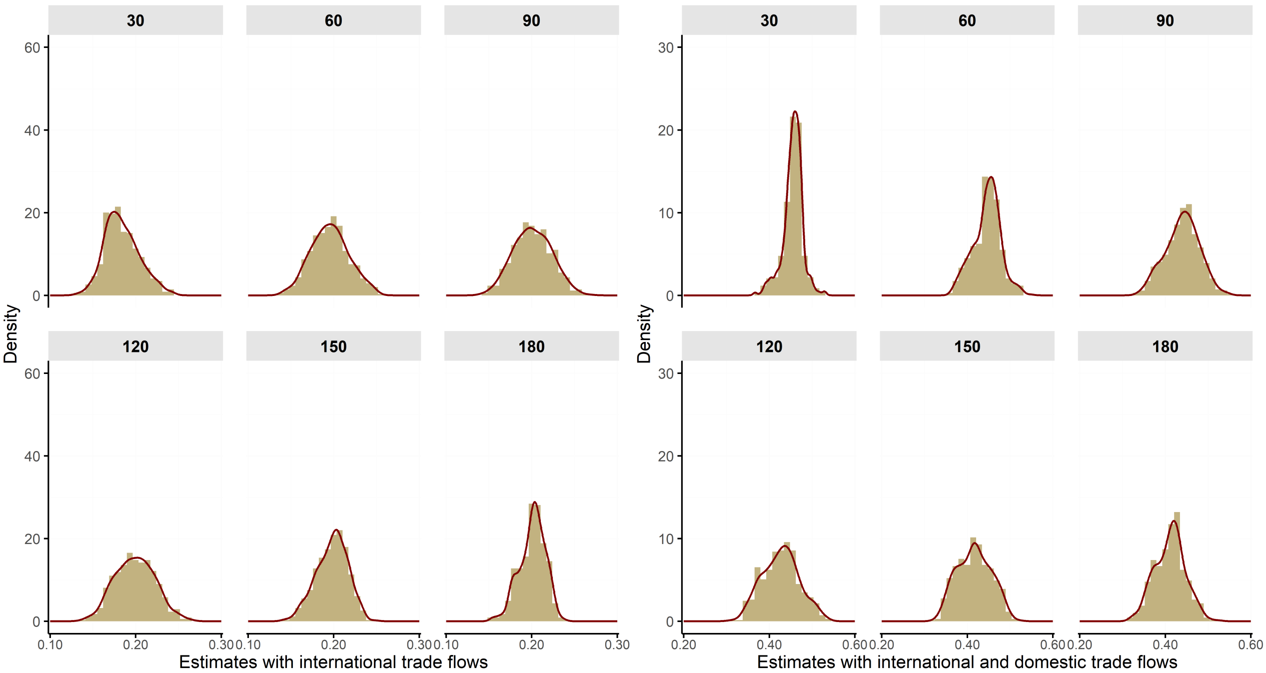
Notes: The vertical axis is the kernel density. The horizontal axis reports the EU membership trade effect estimates. The number above each plot denotes the number of randomly selected control group countries.

Figure D2: Distribution of estimates using different sample sizes of Eora26.



Notes: The vertical axis is the kernel density. The horizontal axis reports the EU membership trade effect estimates. The number above each plot denotes the number of randomly selected control group countries.

Figure D3: Distribution of estimates using different sample sizes of WTF.



Notes: The vertical axis is the kernel density. The horizontal axis reports the EU membership trade effect estimates. The number above each plot denotes the number of randomly selected control group countries.