Missing Tariffs

Feodora A. Teti*

December 19, 2024

Abstract

This paper reveals significant errors in a key variable in international trade: tariff rates. The issues arise from incomplete reporting, leading to measurement error from false interpolation by the data provider, the World Integrated Trade Solution (WITS), and selection bias from dropping tariffs when no trade is recorded. I develop a novel interpolation algorithm to correct these issues and construct a global tariff dataset. Reestimating recent studies relying on WITS data highlights the importance of correcting these errors. Studies using cross-country tariff variation, such as estimates of the trade elasticity, are particularly sensitive to the mishandling of the tariff data.

Keywords: Tariffs, MFN, Preferences, Trade Elasticity, WITS

JEL-Classification: F13, F14

*ifo Institute, Princeton University, University of Munich, CEPR, and CESifo (e-mail: teti@ifo.de).

I am very thankful to Treb Allen, Pol Antràs, Costas Arkolakis, Andrew Bernard, Matilde Bombardini, Arnaud Costinot, Carsten Eckel, Gabriel Felbermayr, Lisandra Flach, Alejandro Graziano, Matthew Grant, Gene Grossman, Elhanan Helpman, Dennis Novy, Gianluca Orefice, Monika Schnitzer, Claudia Steinwender, Felix Tintelnot, Michael Waugh, and Joachim Winter, as well as participants of numerous seminars and conferences for their valuable comments and suggestions. I gratefully acknowledge financial support received from Senatsausschuss Wettbewerb (SAW) under grant no. SAW-2016-ifo-4, the German Research Foundation through CRC TRR 190, and the scholarship from the Bernt Rohrer Foundation. All errors are my own.

1 Introduction

Tariffs are nearly as old as trade itself. Jean-Baptiste Colbert's reforms in 1667 are an early example of how governments used tariffs to protect domestic industries (Smith 1776). Throughout much of the 20th century, multilateral and bilateral trade agreements drove tariffs to historically low levels. But the tide has turned. The US–China trade war and the European Union's recent tariffs on Chinese electric vehicles starkly highlight the return of protectionism. This policy reversal raises important questions. To what extent do tariffs drive inflation? Can they revive domestic jobs? And do tariffs help build industries critical to national security? To answer these questions, one can turn to past episodes of tariff changes, which provide insights into their effects on prices, labor market outcomes, and industrial policy. Yet, despite their importance, comprehensive and accurate data on import tariffs remain surprisingly limited.

As Anderson and Van Wincoop (2004, p. 693) starkly observed, *"the grossly incomplete and inaccurate information on policy barriers available to researchers is a scandal and a puzzle."* This problem arises because tariff rates are reported only when countries choose to do so, resulting in substantial gaps in the time series data due to inconsistent reporting. This issue is further exacerbated by the mishandling in the World Integrated Trade Solution (WITS) database—one of the main providers and a standard source for cross-country tariff rates—of missing observations in the raw data, which introduces nonclassical measurement error and selection bias.[1] These errors pose significant challenges for empirical analysis, raising a fundamental question: How can we accurately estimate the effects of tariffs without correct data?

This paper proposes a new methodology to construct a granular global tariff database that imputes the missing tariff rates and substantially reduces the errors introduced by WITS. By incorporating trade policy features such as gradual tariff phase-outs and sequentially deepening trade agreements, the new algorithm used to compile the database improves the accuracy and reliability of tariff data. To determine how important the mistakes in the existing data are in practice, I focus on estimating the trade elasticity using tariff variation, a fundamental parameter in international economics research that determines the gains from trade in many workhorse trade models. I reestimate the main results of prominent studies using the improved database and show that the errors in the WITS data bias existing estimates of the trade elasticity. The direction of the bias varies across studies because of the nonclassical character of the

[1] As of June 2024, WITS had been cited as source for tariff rates in 126 articles published in the leading economics journals since 2010 (including the *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, *Review of Economic Studies*, *Review of Economics and Statistics*, *AEJ: Economic Policy*, *AEJ: Microeconomics*, *Macroeconomics*, *AEJ: Applied Economics*, *Journal of the European Economic Association*, and *Journal of International Economics*). The complete list of papers is available upon request. Not all papers identified through this exercise necessarily rely on biased tariff data. The extent of potential errors depends on how the authors downloaded the tariffs (e.g., via WITS' download tool or bulk download option) and the cleaning steps taken to address missing tariff rates.

measurement error and selection bias, which can push the bias in either direction depending on the empirical approach, sample composition, and data-cleaning steps taken by the respective authors.

There are multiple sources of error in the tariff rates reported in WITS. Many countries report their tariff rates inconsistently, with preferential tariff rates being especially underreported relative to most favored nation (MFN) tariff rates. WITS addresses these gaps by interpolating missing preferential tariff rates with the MFN rates, creating artificial spikes in bilateral time series data.[2] In 2001, for instance, tariff rates were misreported for 30% of the worldwide imports within regional trade agreements, with interpolation errors leading to an overstatement of the actual rates by an average of 6.9 percentage points. This gives rise to nonclassical measurement error that is always positive and occurs only among country pairs with trade agreements, typically involving large trade flows.

Moreover, the WITS data rely on importers' reports of positive trade flows to UN Comtrade, such that observations for product–country pair–years with zero trade flows are excluded and low-income countries are disproportionately underrepresented in the data.[3] For 2001, we observe reported tariff rates for only 3% of all product–country pairs and cover just 67% of global imports. Hence, the WITS data introduce positive selection bias, which is particularly problematic for researchers seeking to analyze the effects of tariffs on trade across countries with varying income levels. Furthermore, the lack of a full set of tariff rates for all country pair–product combinations makes it impossible to explore important questions, such as the effects of tariffs on the extensive margin of trade or the determinants of tariff protectionism.

A key question is the extent to which the WITS data are used in academic work and how many errors remain unaddressed. A survey of papers published in the top-five general interest economics journals since 2010 shows that 10 out of 21 papers either made no corrections or, in an effort to handle the missing data, fill in the tariff rates with flawed data, such as the incorrectly interpolated preferential tariff rates provided by WITS. These methods, while practical, may inadvertently amplify the bias in the tariff data rather than reduce it. Notably, all but one of these studies rely on cross-country variation in both types of tariffs, preferential and MFN. In such empirical settings. detecting errors is particularly challenging due to the complexity of the data, which spans the importer, exporter, product, and time dimensions. By contrast, studies with narrower scopes—such as those focusing on a single country or a small set of products—tend to address the mistakes in the tariff data effectively (e.g., Conconi et al. 2018; Head and Mayer 2019). Similarly, studies that rely solely on changes in MFN tariffs

[2] The errors described here correspond to the version of the WITS data available through the WITS download tool available as of November 20, 2024, for the "effectively applied tariff". The download procedure is outlined in the appendix.

[3] Gaulier and Zignago (2010) show that imports from low-income countries are underrepresented in UN Comtrade.

generally face fewer errors due to the higher quality of these data (e.g., Bagwell and Staiger 2011; Handley and Limão 2017).

To address the identified issues in the WITS data, I fill in all the missing tariff rates—both preferential and MFN—and then retrieve the lowest available statutory rate.[4] To achieve this, I first minimize the number of missing observations by combining five different sources of tariff data, including unprocessed data from UNCTAD, the WTO, the International Trade Center (ITC), national authorities, and detailed phase-in schedules for 149 free trade agreements (FTAs). To fill in the remaining missing observations, I develop a novel algorithm that infers the missing tariff rates based on reported rates. This algorithm accounts for trade policy features such as gradual tariff phase-outs and the dynamics of multiple, sequentially deepening agreements. Cross-validation with external sources enhances confidence in the reliability of the imputed observations, particularly for preferential tariffs. A validation exercise confirms that my novel algorithm outperforms alternative methods for filling in the missing tariff rates. The final dataset includes 6.7 billion tariff rate observations across 200 countries and their partners over 34 years at the HS6 level, offering a more precise and granular view of global tariff rates than previously available, with wide-ranging applications beyond research in the field of international economics.[5]

To demonstrate the practical importance of correcting tariff data, I examine how the errors in WITS impact existing estimates of the trade elasticity that use tariff variation.[6] This parameter is key for trade economists, as it measures how import quantities respond to changes in trade costs and is crucial for calculating the gains from international trade in many workhorse models (Costinot and Rodríguez-Clare 2014).[7]

Perhaps surprisingly, the direction of the bias arising from the missing tariff rates that we should expect when estimating the effects of tariffs on trade flows is a priori ambiguous because of three competing forces: attenuation bias, differential measurement error, and sample selection. First, the falsely interpolated tariff rates always overstate the true rates, which are often zero or low, as is typical for preferential tariffs. This systematic upward bias in the tariff rates introduces nonclassical measurement error, which leads to attenuation bias that exceeds the magnitude expected under classical random measurement error assumptions. Second, because the measurement error affects only country pairs with trade agreements, which are systematically different from other trading partners, it is correlated with the regression error

[4] This paper focuses solely on applied MFN and preferential tariffs, abstracting from other types of tariffs such as antidumping duties, safeguard measures, or the tariffs introduced during the US–China trade war.

[5] The data are available to researchers here: https://feodora-teti.weebly.com/.

[6] There are also other approaches that do not rely on variation in tariffs (e.g., Alessandria et al. 2024; Broda and Weinstein 2006; Feenstra 1994; Hummels 2001; Shapiro 2021; Simonovska and Waugh 2014; Soderbery 2015; Soderbery 2018).

[7] E.g., applying different elasticities to the ACR gains-from-trade formula (Arkolakis et al. 2012) shows that the estimated gains for the US in 2000 are 7% with a trade elasticity of -1 but only 0.72% with an elasticity of -10.

term. The unobservables contained in the error term, such as bilateral distance, are likely correlated with both trade volumes and tariff rates in opposing directions; thus, the differential character of the measurement error can counteract or even fully offset the attenuation bias. Last, sample selection introduces an additional bias that can go in either direction. Drawing from the literature on measurement error, I formally decompose the bias from the missing tariffs into these three components (Bound et al. 1994; Bound et al. 2001; Bound and Krueger 1991; Imai and Yamamoto 2010; Schennach 2022).

The replication study focuses on three prominent papers estimating the trade elasticity: Arkolakis et al. (2018), Boehm et al. (2023), and Caliendo and Parro (2015). For all three of the studies, reestimating their main results with the corrected tariff data significantly changes the estimated coefficient of the tariff elasticity. In Arkolakis et al. (2018), the attenuation bias dominates, and the tariff elasticity increases in magnitude from -4.28 to -9.81. Conversely, the findings for Boehm et al. (2023) suggest that their reported estimates overstate the absolute value of the tariff elasticity, primarily because of sample selection. With the corrected data, the long-run elasticity coefficient adjusts from -2.14 to -0.80, and the short-run coefficient from -0.62 to -0.46. In the case of Caliendo and Parro (2015), the correction for measurement error approximately cancels out the impact of the correction for sample selection, with the elasticity, overall, moving only slightly from -4.55 to -4.65. These differences in the sensitivity of the trade elasticity estimates to the flawed tariff data provided by WITS can be attributed to nonclassical measurement error and selection biasing the estimates in different directions. The extent to which one factor or the other dominates depends on the identification strategy, the methods used to handle the missing tariffs issue, and the countries in the sample in the respective studies.

The results of the replication study relate to the extensive literature seeking to estimate the trade elasticity with variation in tariff rates, as surveyed, e.g., in Costinot and Rodríguez-Clare (2014), Head and Mayer (2014), and Hillberry and Hummels (2013). Despite the centrality of the trade elasticity in the international economics literature, so far, there is no consensus on the magnitude of this parameter. The estimates range from -10 to -1 and depend heavily on the method, setting, and sample used (e.g., Boehm et al. 2023; Caliendo and Parro 2015; Fajgelbaum et al. 2020; Fink et al. 2005; Fontangé et al. 2022; Imbs and Mejean 2015; Romalis 2007). While one might attribute the discrepancies in these estimates to poor data quality, this paper demonstrates that other factors are at play. Correcting errors in tariff rates increases the variance in the trade elasticity estimates across the three replicated studies, emphasizing that differences stem from empirical methods and sample choices. This finding underscores that the challenges researchers face in estimating the trade elasticity extend beyond data quality to potential issues such as treatment heterogeneity across countries or sectors, difficulties in aggregation, and, more broadly, the need for more careful identification strategies in future research.

While previous efforts have been made to compile datasets on global tariff rates (e.g., Barari and Kim 2020; Bown and Crowley 2016; Caliendo et al. 2023; Guimbard et al. 2012), my work advances these by offering unique coverage in terms of countries (200), time periods (yearly data for 1988–2021), and level of disaggregation (HS6-digit) and improved accuracy due to my novel filling algorithm. Among existing datasets, the most comparable are Caliendo et al. (2023) and CEPII's MAcMap-HS6 (Guimbard et al. 2012), though they differ significantly from the one presented here. Both rely on manually inferring missing preferential tariff rates based on case-by-case review of the legal texts of regional trade agreements (RTAs)—a process prone to errors because of the complexity of such agreements. For instance, CEPII's MAcMap-HS6 incorrectly reports tariffs of 0% for Eastern European countries and the EU in 2001.[8] By contrast, my algorithm leverages expert judgment from the Design of Trade Agreements (DESTA) team (Dür et al. 2014) and the WTO to classify whether agreements allow for phasing in. It uses preferential tariff rates reported in any year within the RTA implementation period to effectively infer phase-out schedules at the product level, fully utilizing the available data. A validation exercise using simulated patterns of missing tariff schedules shows that my algorithm produces more accurate results than previous contributions in the literature. Further, the coverage of Caliendo et al. (2023) and MAcMap-HS6 is substantially more limited, rendering them unsuitable for replicating the three papers analyzed in this study (see Section 4.4 for a more detailed discussion).

The remainder of this paper is organized as follows. Section 2 discusses how WITS mishandles and exacerbates the missing data issues. Section 3 surveys how the WITS data are used in existing research. Section 4 introduces the new methodology to fill in the gaps. Section 5 examines the direction of the biases from nonclassical measurement error and sample selection. Section 6 replicates three prominent studies on the tariff elasticity. Section 7 concludes.

2 Missing Tariffs in Standard Sources

This section explains how WITS's data processing introduces errors that compromise the suitability of its tariff rates for empirical research. The tariff rates provided by WITS are affected by two issues: false interpolation and positive selection.

[8] Official Journal of the European Communities No L114/27.

2.1 False Interpolation

WITS compiles tariff data from UN Trade and Development's (UNCTAD's) Trade Analysis and Information System (TRAINS), the World Trade Organization (WTO), and the ITC, offering a tool to download data at the HS6 level for nearly 200 countries, with some observations dating back to 1988.[9] It offers comprehensive coverage of MFN (bound and applied) and bilateral preferential tariff rates but does not cover other types of tariffs, such as retaliatory tariffs or antidumping duties. Additionally, WITS provides information on trade flows and some nontariff measures.[10]

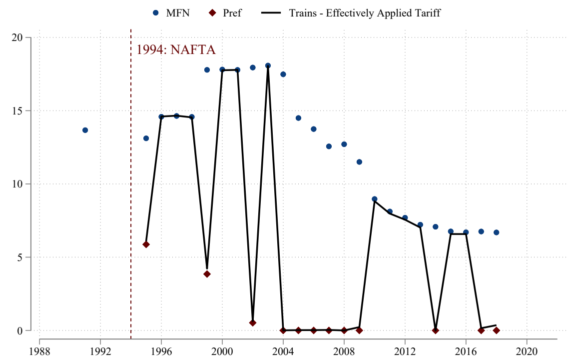
To determine the tariffs that countries impose on each other, WITS uses the concept of the min-tariff, defined as the lowest available tariff rate. If the preferential tariff t_{ijkt} that country i imposes on country j in year t for product k is available, it is used; otherwise, the tariff equals the MFN tariff t_{ikt}. Formally, this can be expressed as $t_{ijkt} = min\,\{t_{ikt}; t_{ijkt}\}$.[11]

The min-tariff rule relies on a crucial assumption: When a preferential tariff rate is not reported for a given year, it is assumed that no trade agreement exists between the countries. However, countries do not consistently report their complete tariff schedules—including both their MFN and preferential rates under all trade agreements—for every year. When observations for preferential tariffs are missing, the min-tariff rule leads to false interpolation, causing tariffs to jump to the (higher) MFN level. These artificial spikes in the time series data do not reflect genuine changes in trade policy but are instead artifacts of WITS's flawed interpolation method.

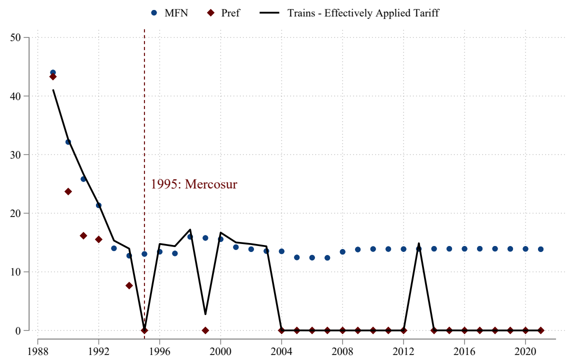
Figure 1 shows the average tariff across more than 5,000 HS6-level products for selected country pairs that signed an RTA during the period of observation, highlighting how the missing observations in the reported preferential tariff rates and WITS's false imputation lead to substantial measurement error. The red dots represent the simple average of the preferential tariffs across all products k for a given year, while the blue dots give the simple average of the MFN tariffs. The solid line represents the average tariff that WITS reports for the respective country pair.

Panel (a) shows Mexico's tariffs on US imports. Before NAFTA's implementation on January 1, 1994, Mexico imposed its MFN tariffs on US goods; afterward, the preferential tariffs negotiated under NAFTA were applied.[12] As Figure 1 indicates, Mexico did not consistently report its MFN and preferential tariff rates for all years: The MFN tariff rates are complete from 1995 onward but only sporadically reported for the years before then, while the preferential tariffs were reported even less frequently, as illustrated by the sparse red dots. The min-tariff rule leads to

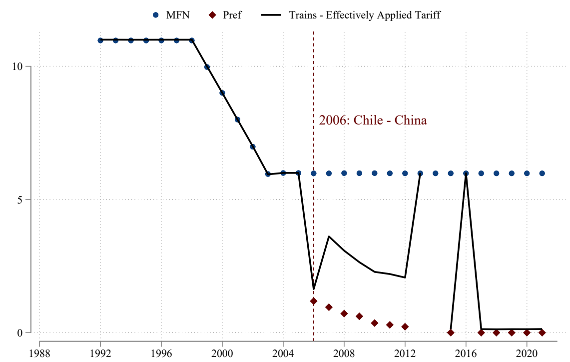
[9] WITS can be accessed here: https://wits.worldbank.org/. Most of the raw data since 2010 are sourced from the ITC. For more details on data providers, see the WITS homepage: https://wits.worldbank.org/dataproviders.html.
[10] See Ederington and Ruta (2016) for a summary of the nontariff measures available through WITS.
[11] WITS refers to this variable as the "the effectively applied tariff" UNCTAD (2011, p. 95).
[12] See Besedes et al. (2020) for more details.



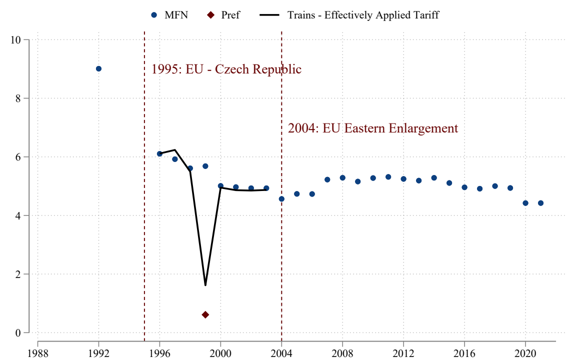
(a) NAFTA: Mexico–United States



(b) Mercosur: Brazil–Argentina



(c) Chile–China



(d) EU East. Enlargement: Czech Rep.–EU

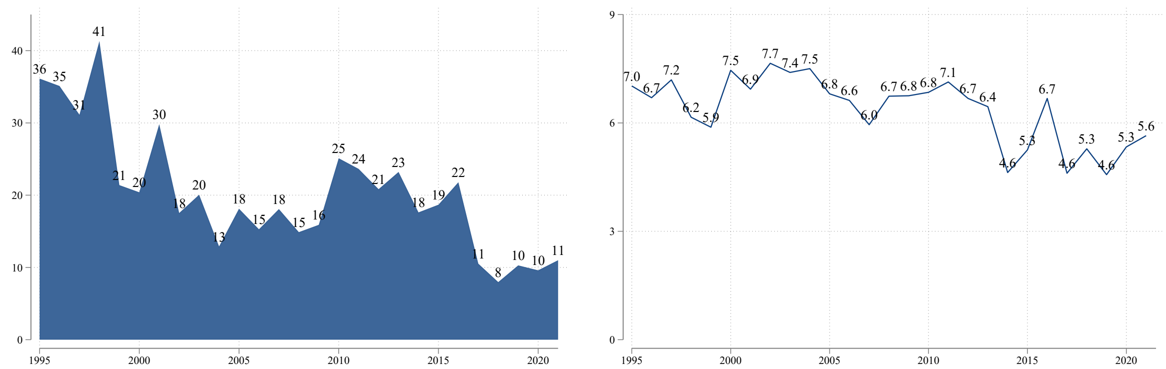
Figure 1. Missing Tariffs and False Imputation in WITS: Four Examples

Note: The figure shows the unweighted average over all HS6-level products k of the preferential, MFN, and "effectively applied tariff" rates reported by WITS for selected country pairs. The spikes indicate missing preferential tariff rates in the presence of reported MFN tariffs.

massive variation over time in the tariffs, with spikes in each year for which the preferential rates are missing but the MFN rates are available making the data unsuitable for any empirical analysis.

The problem of false imputation due to missing preferential tariff rates is not unique to Mexico—rather, it is prevalent for countries around the world involved in other important RTAs. For example, trade within Mercosur, Latin America's largest free trade area, has been largely tariff-free since the customs union came into effect in 1995 following a four-year phase-in period. Again, the misreporting leads to measurement error in the WITS data (see Figure 1, Panel (b)). Similar issues appear with Chile's tariffs on Chinese imports (Panel (c)). Panel (d) shows the time series of average tariffs between the Czech Republic and the EU. Although the Czech Republic joined the EU in 2004, most tariff cuts occurred earlier because of the 1995 Europe Agreement, which eliminated tariffs on EU exports—something difficult to detect in WITS, where preferential tariff rates are reported only for 1999.

For the observed spikes to be interpretable as actual changes in trade policy rather than artifacts of the incorrect WITS interpolation, they would need to correspond to some event that temporarily raised the tariff rate to exactly the level of the MFN tariff. One potential explanation is the suspension of an RTA, but this has occurred for only 2.23% of country pairs.[13] Moreover, there is no evidence of suspensions for any of the four country pairs in Figure 1. Other potential explanations such as the entry into force of antidumping tariffs are insufficient since WITS does not include them in the first place. In general, it is difficult to find alternative explanations for the tariff jumps that align perfectly with MFN levels across more than 5,000 products.[14]



(a) Share of Imports in RTAs with False MFN Tariffs

(b) Average Interpolation Error (in pp.)

Figure 2. Extent of Falsely Interpolated Tariff Rates in RTAs

Note: Panel (a) displays the share of imports within RTAs impacted by falsely interpolated tariffs. Panel (b) shows the average tariff interpolation error in percentage points. Trade flows within the European Union and those subject to zero MFN tariffs are excluded.

As Figure 2 shows, using WITS data results in MFN tariffs being mistakenly applied to a significant share of imports within RTAs. These numbers exclude trade within the European Union, as WITS does not report those tariffs. Including intra-EU trade would artificially inflate the share of flawed tariffs, as missing tariffs within the bloc can be easily corrected by replacing them with zero, reflecting the European Single Market's policy of eliminating trade barriers among member states. The share was particularly high in the late 1990s, peaking at 41%, before declining to 25% by 2010 and stabilizing around 10–11% by 2019. The average interpolation error ranges from 4.6 to 7.7 percentage points, and imports between low-income countries are

[13]This number is based on all RTAs in force in 2021 according to Mario Larch's RTA database (Egger and Larch 2008).

[14]The differences in unweighted averages in Figure 1 are due solely to aggregation bias.

particularly affected by false interpolation, especially in earlier years (see appendix). These numbers do not reflect a fixed sample of countries over time. Instead, they likely result from different countries failing to report tariffs in different years, which complicates panel data analysis and can lead to additional loss of observations.

2.2 Positive Selection

WITS typically reports far fewer observations than the approximately 199 million that should be available each year, considering the 200 importing countries that impose tariffs on 199 exporters for approximately 5,000 products. For example, for 2010, WITS reports only 5.5 million observations. This discrepancy occurs because WITS includes only importer–exporter–product combinations with positive trade flows. This is problematic for empirical research because of the endogenous relationship between tariffs and trade: Higher tariffs often correlate with lower trade volumes, and in some cases, tariffs can be high enough to prevent trade entirely. Consequently, the WITS data are subject to positive selection bias.[15] Additionally, this selected sample limits researchers' ability to explore important questions, such as the effects of tariffs on the extensive margin of trade or the determinants of tariff protectionism.

Even more troublesome is that WITS underreports trade flows because of the incompleteness of the data from its primary source, UN Comtrade. Brilliant work by CEPII shows that many countries either fail to fully report their trade to UN Comtrade or provide data only at an aggregated level (Gaulier and Zignago 2010). CEPII addresses these gaps by using mirror-data, where missing data on import flows for one country are substituted with the corresponding export data reported by its trade partner. Unsurprisingly, low-income countries are the most affected by poor reporting compliance. Hence, any analysis using the WITS tariff data will suffer from selection bias, even when the focus is on observations with positive trade flows.

Figure 3 shows the share of observations that are missing from WITS because of underreporting of trade flows in UN Comtrade, calculated as the number of observations available in WITS relative to the larger dataset from CEPII's Baci trade data. The share of missing observations is notably higher for earlier years, reflecting improvements in UN Comtrade's coverage over time.[16] This share also negatively correlates with income: While relatively low for high-income countries, it has fluctuated around 50% for least-developed countries (LDCs) since 2001. Additionally, low- and middle-income countries in the Americas are notably underrepresented in the WITS sample. Furthermore, the trade flows for the country pair–products that are not reported in WITS but are included in CEPII's Baci database are systematically

[15]This issue relates to the literature on tariff aggregation, which highlights that the use of import weights often leads to underestimation of a country's level of tariff protectionism (Anderson and Neary 1994; Kee et al. 2008).

[16]The Baci data are available for years starting in 1995.

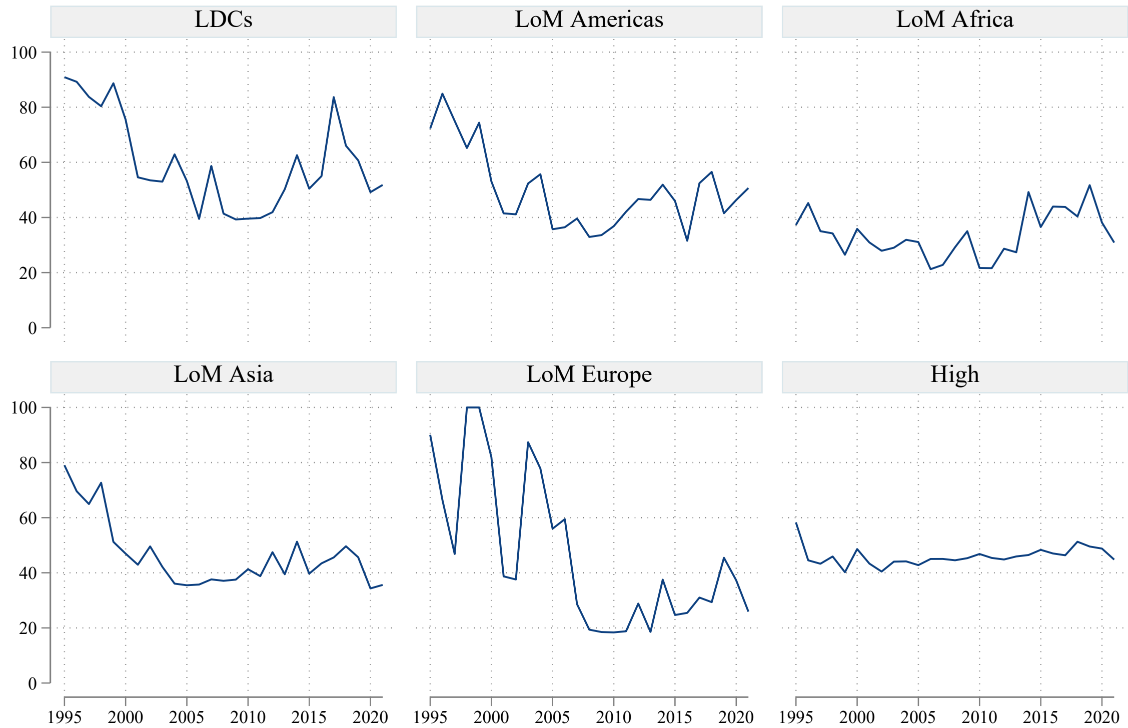


FIGURE 3. SHARE OF MISSING WITS OBSERVATIONS DUE TO UNDERREPORTING IN UN COMTRADE

Note: The figure shows the share of missing WITS observations attributable to underreporting of trade flows in UN Comtrade, calculated as $1 - \frac{N_{WITS}}{N_{Baci}}$, where N is the number of observations. Income groups follow the World Bank classification: least developed countries (LDCs), regional low- or middle-income (LoM) countries, and high-income countries.

lower than those, that are included in both with a ratio of reported trade flow observations to missing ones ranging from 1.23 to 2.79 for 1995—2021 (see Online Appendix B.3).

3 Role of WITS Tariff Data in Economic Research

Given the issues with the tariff data in WITS, it is crucial to understand which empirical results might be impacted by these errors. Using Google Scholar, I compiled a list of papers using the WITS data and published in the top-five general interest economics journals (*American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*) since 2010. I identified 21 such papers, summarized in Table 1 and described in more detail in Appendix A. These papers can be grouped into five main categories by how tariffs are used: studies identifying the trade elasticity (four papers), studies of product quality (one paper), studies of bilateral or multilateral trade policy effects (six papers), analyses of firm-level outcomes (five papers), and studies using tariffs for constructing counterfactuals or model moments (five papers).

Table 1. Biases and Corrections in Tariff Rates in WITS-Based Papers in Top-Five Journals

			# Importer		Tariff Type	
			Many	1 or 2	both	MFN
Bias due to False Interpolation & Selection						
✗	2	do not address data issues, relying on raw data as provided.	2	–	2	–
⚠	8	handle selection issues with flawed tariffs, carrying errors across years.	7	1	8	–
Potential Bias due to Selection						
?	5	may be biased due to selection bias but unclear.	4	1	1	4
Correct Data						
ℹ	3	rely on legal texts for information about tariff cuts.	1	2	2	1
★	1	use accurate raw data.	–	1	–	1
✓	2	correctly fill in missing data at the HS6 level for missing tariffs, effectively addressing data issues.	1	1	2	–
21 Papers using WITS Data						

Note: This table categorizes the 21 papers published in top-five economics journals (*American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*) since 2010 that use WITS data, based on how they address the issues inherent in the dataset. The classification is derived from a review of the data-cleaning steps described in the original papers and, where available, their online appendices hosted on the journals' websites. Additional details are provided in Appendix A.

I reviewed how each paper utilized the WITS data, focusing on the type of tariff (applied MFN, bound MFN, or preferential), country coverage, level of disaggregation, and time frame, as well as whether any corrections were applied to the data. Three key insights emerged: First, of the 21 papers reviewed, 10 either made no corrections or relied on imputing missing MFN and preferential tariff rates with flawed values, propagating errors across year. All but one of these studies use cross-country data, and all rely on variation in both MFN and preferential tariffs. Consequently, it is the studies using tariff variation to estimate the trade elasticity or analyzing firm-level outcomes, as well as studies using tariffs to construct counterfactuals that are particularly impacted by the errors in the tariff rates. The notable exception is the study by Feenstra and Romalis (2014), who undertook a data cleaning process related to what I will propose later in this paper.[17]

The studies working within well-defined institutional contexts and using disaggregated data were likelier to correct data issues successfully, as in such settings errors are much easier to spot. For example, Head and Mayer (2019) manually coded the tariffs for fewer than 40 HS6 products, while Conconi et al. (2018) filled the gaps in Mexico's tariff schedule under NAFTA using subsequent-year data. This suggests that addressing the WITS data problems is more feasible when the research focus is narrower.

[17]Feenstra and Romalis (2014) used an earlier version of the data introduced in Caliendo et al. (2023).

Furthermore, the WITS data were often used to retrieve information about MFN tariff changes (in six out of the 21 papers). This type of research is subject to significantly fewer errors because of the better data quality for MFN rates (e.g., Bagwell and Staiger 2011; Handley and Limão 2017). However, these kind of studies may still be biased due to selection.

4 Filling the Gaps: A New Methodology for Constructing Global Tariff Data

In this section, I present a new methodology for addressing missing tariff rates. To fill the gaps in the data, I fill in all missing tariff rates—both preferential and MFN—before retrieving the lowest available statutory tariff rate. To do so, I proceed in three steps: First, I combine data from WITS with other databases containing tariff information, to reduce the number of missing observations. Second, I develop a new algorithm to impute the remaining gaps in the data. These first two steps are performed separately for MFN and preferential tariffs to ensure complete coverage for both. Finally, I combine the MFN and preferential tariff datasets—now without missing observations—and use them to retrieve the tariffs t_{ijkt} for all importer–exporter–product–time combinations, thereby addressing the positive selection bias in the WITS data.

4.1 Primary Data Sources

I obtain information from five primary sources: detailed phase-out schedules available through the WTO's RTA Database; country-reported raw tariff schedules from TRAINS and the WTO's Integrated Data Base (IDB), both accessed via WITS; data from ITC's Market Access Map; and US and EU tariff schedules from national authorities, kindly provided by researchers at the World Bank (Forero-Rojas et al. 2018). Table 2 summarizes the coverage of the respective sources and shows the available years and number of importing countries that they cover.

For 149 RTAs, I collected detailed phase-out schedules for all tariff lines and participating countries, which eliminate the need for imputation when they are available. These schedules, found in the WTO's RTA database, have particularly good coverage of more recent FTAs.[18] These digitized schedules—totaling 384 for 149 RTAs—detail the timing of tariff cuts for each tariff line and participating party. Because of formatting inconsistencies, the raw files were manually standardized, and I aggregated the data to the HS6 product level by averaging over the tariff lines. The main advantage of the phase-out schedules is their completeness: They cover

[18]The data can be accessed at https://rtais.wto.org/.

Table 2. Primary Sources for Tariffs

Source	type of tariffs	covered years	# importing countries
1) Phase-out schedules	preferential	1994–2021	91 (149 RTAs)
2) UNCTAD's TRAINS	MFN & preferential	1988–2021	197
3) WTO's IDB	MFN	1996–2021	163
4) ITC's Market Access Map	MFN & preferential	2007–2021	197
5) National authorities	MFN & preferential	1997–2017	EU & US

Note: This table reports primary tariff data sources, their coverage by year, and the number of importing countries. The sources are listed in the order used to fill in the missing observations: The first row serves as the primary source, with subsequent sources used sequentially as needed.

all tariffs for all national tariff lines and all years. Hence, whenever this type of information is available, no imputation is necessary. Therefore, for preferential tariffs, the schedules serve as my primary source of information—only when they are not available do I use data from the other sources.

In addition to the error-prone processed data, WITS provides a lightly processed version of the preferential and MFN tariff rates based on raw tariff line data from UNCTAD's TRAINS. WITS aggregates these data to the HS6 level (with unweighted means) and organizes them by tariff schedule, including MFN and preferential tariff rates from various trade agreements.[19] While this version does not suffer from selection bias, as the inclusion of observations is not conditional on country pairs having positive trade flows, the issue of missing information in both preferential and MFN tariff rates remains unresolved, leading to gaps for nonreporting years that still need to be addressed. TRAINS provides broad coverage, encompassing 200 importers and their partners for years from 1988 onward, making it the core primary source of tariff data for many country pairs.

Unprocessed versions of the WTO's IDB data are not available for bulk download. The only accessible version is provided by WITS through its download tool and is affected by false interpolation and selection bias. While reliable information on preferential tariffs cannot be recovered, the primary issue with MFN tariffs is selection bias. Because I use this information only to complement other sources, its limitations are not a concern.

The ITC's Market Access Map is another established source for tariff data,[20] covering bound, applied MFN, and preferential tariffs from 2007 onward for 197 countries at the tariff line level for various tariff schedules, and relying on voluntary country reporting. Consequently, this source can be expected to be affected by issues similar to those in WITS, i.e., missing

[19]This version of the data is available through the bulk download option on WITS at https://wits.worldbank.org/.
[20]Available at https://www.macmap.org/. This is the raw data that CEPII's MAcMap-HS6 is based on.

observations for both preferential and MFN tariffs.[21] Similarly to the unprocessed TRAINS data, the ITC data do not condition the inclusion of observations on country pairs having positive trade flows, so selection bias is not present. Since 2010, TRAINS has supplied WITS with ITC-collected tariff data, so the raw data for recent years are identical across both sources. However, the country overlap varies, as some countries report to the ITC only in certain years and others report only to TRAINS. The data, initially at the tariff line level, are aggregated to the HS6 level.

Last, I use data provided by US and EU national authorities. The US International Trade Commission gathers and publishes tariff rates for the United States and the European Commission those for the EU.[22] All data cleaning and preparation has been done by researchers from the World Bank (Forero-Rojas et al. 2018), who have kindly provided me with access. Similarly to the phase-out schedules, the original data give tariffs at the lowest level of disaggregation, i.e., the national tariff line, and have to be aggregated to HS6 product level to be matched to the tariff data from the remaining sources.

For preferential tariffs, the phase-out schedules from the WTO's RTA database serve as the primary information source. Missing observations of preferential tariffs are sequentially filled in with information provided by TRAINS, then with the ITC data and, last, for the US and EU, with data from the national sources. For MFN tariffs, TRAINS serves as the primary source; the missing data are sequentially filled in with information provided by the WTO's IDB, ITC, and national authorities. Ad valorem equivalents for non ad valorem tariffs are taken only from TRAINS; for all other sources, only ad valorem tariffs are included.

4.2 Novel Imputation Algorithm

Having combined all available raw data on tariffs, I carefully correct for the remaining missing tariff rates. Although the scale of the missing data problem is much more pronounced for preferential tariffs, reporting compliance is also imperfect for MFN tariffs t_{ikt}. Rather than replacing missing MFN tariffs by linearly interpolating between available observations, I set each missing observation equal to the nearest preceding observation. If there is no preceding observation, the missing MFN rate is set equal to the nearest subsequent observation. Figure 4 illustrates the procedure. This procedure accounts for anecdotal evidence that countries are likelier to update schedules after a significant tariff change and follows Caliendo et al. (2023).

[21] The ITC also provides an imputation method for the effectively applied tariff. Setting the criterion "minimum rate" instead of "by each trade agreement" in the associated download tool yields data series affected by false interpolation due to the min-rule.

[22] The raw data are available at https://dataweb.usitc.gov/tariff/annual and https://circabc.europa.eu/faces/jsp/extension/wai/navigation/container.jsp.

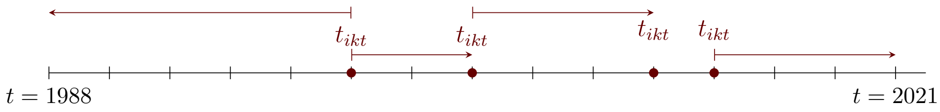


Figure 4. Imputation Algorithm for MFN Tariffs

The imputation algorithm for preferential tariffs follows the same principle but accounts for the presence of RTAs: Preferential tariffs are not imputed for years when no RTA is in effect. This requires linking every reported preferential tariff to an RTA, including its year of entry into force and, if applicable, the year it became inactive. As such data are not readily available, I construct it by combining information from five different databases. Details on the data construction are provided in the appendix.

The algorithm for imputing missing preferential tariffs is illustrated in Figure 5. Panel (a) shows the imputation process for cases with no phase-in period. When my newly created RTA database confirms that an RTA is in force for a given country pair ij, I use the preceding preferential tariff if available; otherwise, I use the first subsequent one. Notably, even a single reported observation following the RTA's entry into force is sufficient to correctly interpolate the preferential tariff for all missing years, as the tariff remains constant.[23]

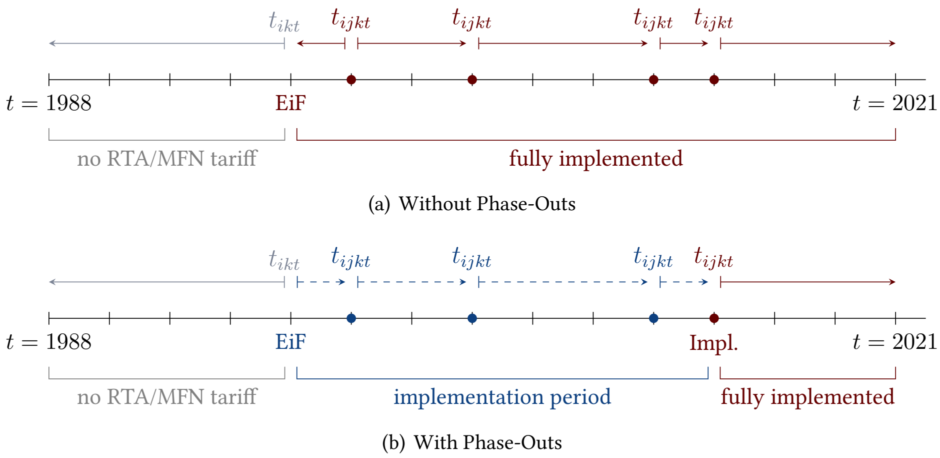


Figure 5. Imputation Algorithm for Preferential Tariffs

Note: The solid line represents missing observations filled by forward and backward filling, while the dashed line represents linear interpolation. "EiF" stands for entry into force, "Impl." for year or full implementation. More details can be found in the appendix.

Trade agreements can be divided into two groups: RTAs under which tariffs are gradually phased out over time and RTAs under which all tariff reductions occur when the RTA enters into force. In agreements with phased-out tariffs, there is often significant heterogeneity across products in the timing of the phasing-out. For instance, in NAFTA 51% of all tariff lines were

[23] For agreements notified under the enabling clause, I only interpolate forward, not backward. More details on this can be found in the appendix.

cut immediately, while the MFN tariffs on the remaining lines were phased out over five, ten, or fifteen years (Besedes et al. 2020). Imputing the preferential tariff rates applied during the phase-in period is challenging: Using preceding observations risks overstating the true rate, while using subsequent observations may understate it.

I address this issue in two ways. First, as described above, I supplement the existing preferential tariff rates with tariff schedules from the WTO's RTA Database, which provide phase-out details at the tariff line level. Hence, when these tariff schedules are available, there is no need for imputation. Second, I complement the RTA data with information on whether tariffs were phased out and, if so, the year by which the phasing-out was to be fully implemented.[24] For cases where RTA phase-in occurs, I propose linearly interpolating between the MFN tariff rate in the year prior to the RTA's entry into force and all reported preferential tariff rates from the year of RTA entry into force to the implementation year. After the implementation year, I use forward and backward filling to complete the preferential tariff series. Figure 5 (b) illustrates this approach.

Unfortunately, the RTA databases indicate only the final year when the full implementation must be completed, hence, there is no product dimension. However, when countries report their preferential tariff rates during the implementation period, linear interpolation can capture differences in phase-out schedules. For products with immediate tariff elimination, the linear interpolation yields a steeper time series, closely approximating the step function that would characterize the true tariff rates when tariffs drop from the MFN level to zero as the RTA enters into force. Conversely, for products with gradual tariff reductions, the decline in the linearly interpolated time series appears more gradual, accurately reflecting the phase-out period. The sooner importing countries report their preferential tariff rates after the RTA's year of entry into force, the more effectively the algorithm can capture these product-level differences in phase-ins.

It is worth emphasizing that accurate information on the existence and type of an RTA is vital for accurate imputation of preferential tariffs. Combining multiple RTA data sources is essential for cross-checking and validating the presence of RTAs, thereby ensuring the reliability of the imputation process. The algorithm also incorporates various additional trade policy features. Ignoring these factors could result in biased preferential tariff rates. I will now outline these features, with more details described in the appendix.

Sequential deepening of RTAs Not all country pairs have signed just one RTA; bilateral trade relations often evolve and deepen over time. For instance, Mexico initially received preferential access to the US for certain goods through the General System of Preferences

[24] These data are available through the WTO's RTA Database and DESTA.

(GSP). When NAFTA came into effect, the covered product scope expanded, and existing preferential tariffs that were not yet zero were eventually eliminated completely under NAFTA. For agreements that deepen over time, failing to account for potentially differing scopes of the products covered by the RTA could lead to falsely impute preferential tariff rates. For example, if agricultural products were excluded from the initial agreement, RTA^1_{ij}, between countries i and j but included in a subsequent agreement, RTA^2_{ij}, the imputation algorithm might incorrectly extend the preferences under RTA^2_{ij} to years when only RTA^1_{ij} was in force. This could result in the imputed preferential tariffs being too low if countries do not report them for RTA^1_{ij} but do for RTA^2_{ij}. Conversely, if the reporting pattern is reversed, the imputed tariffs could be too high for the period when only RTA^2_{ij} was in force.

The algorithm addresses this issue by assigning each reported preferential tariff for product k to its corresponding trade agreement and hence an entry into force year. It does so by identifying the first year in which a preferential tariff for product k is observed. In that year, the algorithm determines which RTA is in force and assigns its entry into force year to the preferential tariff rate for product k, allowing the entry into force years to vary by product k. To implement this, I construct a coherent list of trade agreements that includes the full sequence of agreements. Using this list, I perform the interpolation only within each trade agreement. When the agreement name changes, no backward interpolation is performed; instead, only forward interpolation is applied.

To do so, I need to make sure that the newly compiled RTA database includes the full sequence of agreements and that the names of the RTA only changes when there is a potential change in the scope of products covered by tariff reductions. For example, the data should include the GSP for the US–Mexico relationship until 1993 and replace it with NAFTA from 1994 onward to account for the significant expansion in product coverage. By contrast, the USMCA, which entered into force in 2020, should not be included, as it introduced nontariff changes, such as stricter rules of origin, without altering preferential tariffs.

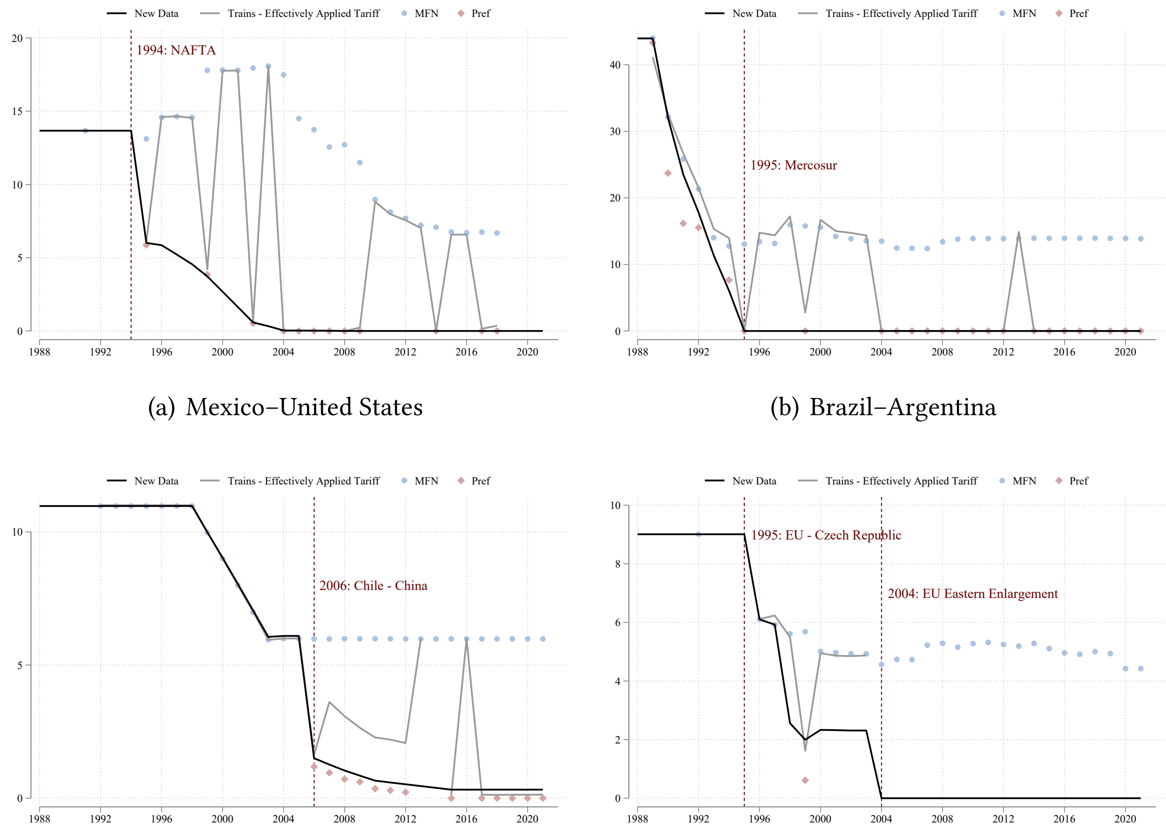
To achieve this, I leverage the fact that some trade agreements are more likely to have a broader product scope than others. Specifically, I distinguish between agreements notified under Article XXIV of the GATT/WTO and those notified under the Enabling Clause. The key difference is that Article XXIV agreements are required to liberalize substantially all trade, whereas Enabling Clause agreements are not. Therefore, when trade relations between a country pair deepen from an Enabling Clause agreement to an Article XXIV agreement, the scope of covered products is expected to expand. Accordingly, I add information on whether an agreement covers substantially all trade. In addition, I incorporate information on substantial changes in the coverage of trade agreements to capture potential changes in product scope when country pairs sequentially deepen their existing Article XXIV agreements. Further details on the data construction process are provided in the appendix.

With this dataset, I interpolate missing tariff observations separately for each RTA. For example, from 1988 to 1993, the United States granted Mexico preferential market access under the GSP, and I interpolate only the missing tariffs during this period using the reported preferential tariffs under GSP. From 1994 onward, I interpolate the remaining missing observations to reflect the preferential tariffs under NAFTA. If a product k was already preferential under GSP, I assume it remains preferential under NAFTA, but not vice versa. This approach ensures that preferences from one agreement are not incorrectly extended to periods covered by another RTA, thereby increasing the accuracy of the imputed tariff rates.

Multiple RTAs Countries may sometimes qualify for multiple preferential tariff schemes, such as the Everything but Arms arrangement and the EU's GSP, in which cases I assume that the lowest available tariff applies. However, challenges arise when not all tariff preferences for all schemes are reported. For example, before the EU's eastern enlargement, the Baltic countries had access to both GSP and Europe Agreement preferences, but for 2000, only GSP tariffs are reported, and the (lower) Europe Agreement preferential tariffs are missing. Simply assuming that the minimum tariff applies without recognizing that data are missing results in overestimation of the imputed tariff, as the algorithm would mistakenly use the (higher) GSP rate for the interpolation. Conceptually, this issue is analogous to the false interpolation of missing preferential tariff rates with MFN rates, leading to similar spikes in the time series. To correct for this, the algorithm takes the minimum of all reported preferential tariffs but sets them as missing if they increase over time. In the case of the EU, this procedure ensures that the interpolation algorithm does not use the (higher) GSP tariffs, which would yield too high tariffs, but instead correctly identifies the preferential tariff observation for the year 2000 as missing for both the Europe Agreement and the GSP. This allows the algorithm to accurately interpolate using information from the Europe Agreement for preceding and subsequent years.

Graduation from Nonreciprocal Trade Arrangements Low-income countries receiving preferences through nonreciprocal trade arrangements can lose preferential access when they "graduate," that is, when they reach a certain income level. Graduation can be observed in the RTA data and is accounted for by the algorithm. However, the algorithm cannot fully account for cases where countries graduate from specific products, which occurs when there is no longer a "competitive need " for the preferential rate (Ornelas 2016), as to the best of my knowledge, there is no reliable primary data source documenting product-level graduation.

Changing Nomenclatures To make use of all available information over time, it is crucial to convert all 6-digit product codes into their counterparts under the first available nomenclature, HS88/92. Without this conversion, only missing observations for product codes in the same nomenclature or that remained unchanged over the 30-year period would be filled in.[25]



(a) Mexico–United States

(b) Brazil–Argentina

(c) Chile–China

(d) Czech Republic–European Union

FIGURE 6. IMPUTED TARIFF RATES FOR SELECTED COUNTRY PAIRS

Note: The graph shows the unweighted average of the imputed tariff rates for selected country pairs.

Having addressed these issues in the interpolation algorithm, I next retrieve the statutory tariffs that countries impose on their partners. To do this, I follow WITS's approach by applying the min-tariff rule, but I use the fully interpolated MFN and preferential tariff rates for all observations. This involves constructing a matrix that includes every country pair–product–year combination $ijkt$, thereby addressing the selection bias present in the WITS data. Figure 6 illustrates how the imputation method changes the time series for the examples discussed in Section 2, showing that the troublesome spikes are eliminated for all the country pairs.

[25]Concordance tables are available through WITS.

4.3 Validation of the Algorithm

To validate the proposed algorithm, I require a subset of tariff observations free from misreporting to evaluate how effectively the algorithm corrects the false interpolation performed by WITS. This involves simulating missing data patterns and applying the algorithm to fill in the artificially created gaps. An ideal reporting country for this validation exercise would have near-perfect reporting compliance and RTAs, its trade policy would incorporate some of the features discussed earlier, and its MFN tariff rates would be nonzero for at least some products; otherwise, the impact of WITS's false interpolation—which replaces missing preferential tariff rates with the MFN rates—would be minimal. Japan meets these criteria, with its 21 RTAs that deepen sequentially and offer multiple preferences to some exporting countries, as well as its nonzero (specifically, 5.2%) average MFN tariff rate.[26]

I simulate different levels of data availability, randomly keeping only 10%, 20%, 30%, and so on up to 90% of all tariff schedule–years to mimic real-world scenarios with varying shares of missing preferential tariff data. Then, I apply the novel algorithm and alternative methods to fill in the missing preferential tariff rates. Each simulation is repeated 250 times. The validation exercise focuses on the filling algorithm for the preferential tariffs, it does not simulate missing MFN tariff rates or the bias due to sample selection. The results of the new algorithm are compared to those of alternative methods, offering insights into the algorithm's relative effectiveness. Additionally, the analysis highlights the importance of each step in the algorithm's design.

Table 3 displays the average and 25th/75th percentile (in parentheses) of the mean absolute error (MAE, in percentage points) of the 250 replications, comparing the proposed algorithm to alternative methods for interpolating missing tariff rates across varying levels of data availability. Column (2) reports the results of the WITS interpolation method, which serves as a benchmark reflecting the outcomes obtained with the WITS interpolation method, i.e., replacing missing preferential tariffs with MFN tariffs. Column (3) replaces all missing preferential tariff observations with zero. Column (4) fills missing observations using preceding preferential tariff rates, or subsequent ones if preceding values are unavailable. Column (5) uses linear interpolation to fill in the gaps between reported observations. For missing values at the start or end of the sample, forward and backward filling is applied: The first reported tariff rate is used for earlier missing values and the last reported tariff rate for later ones.

Several key findings emerge. First, the algorithm demonstrates substantial improvements in data quality relative to that of the WITS data, regardless of the share of reported tariffs,

[26] Japan's reporting to WITS is not entirely complete; a few tariff schedule–years are missing. However, all but one (the tariff schedule for the Japan–Singapore RTA) were successfully retrieved from the WTO's RTA Database, ensuring completeness.

Table 3. Mean Absolute Error Across Interpolation Methods

Reported	Algorithm	WITS	Zero	Fill-In	Linear	CEPII-RTA
	(1)	(2)	(3)	(4)	(5)	(6)
10%	1.00	1.94	0.92	0.89	0.89	1.88
	[0.80; 1.17]	[1.91; 1.99]	[0.85; 0.90]	[0.81; 0.86]	[0.82; 0.87]	[1.86; 1.94]
20%	0.67	1.71	0.88	0.82	0.82	1.70
	[0.55; 0.76]	[1.65; 1.78]	[0.86; 0.90]	[0.79; 0.85]	[0.79; 0.85]	[1.65; 1.76]
30%	0.54	1.49	0.89	0.81	0.81	1.53
	[0.46; 0.58]	[1.43; 1.57]	[0.87; 0.91]	[0.79; 0.84]	[0.79; 0.84]	[1.48; 1.60]
40%	0.42	1.28	0.90	0.80	0.80	1.36
	[0.36; 0.46]	[1.22; 1.36]	[0.90; 0.91]	[0.79; 0.80]	[0.79; 0.79]	[1.31; 1.42]
50%	0.34	1.06	0.90	0.80	0.80	1.19
	[0.30; 0.38]	[0.99; 1.13]	[0.90; 0.91]	[0.79; 0.79]	[0.79; 0.79]	[1.13; 1.25]
60%	0.28	0.83	0.90	0.79	0.79	1.02
	[0.24; 0.29]	[0.75; 0.89]	[0.90; 0.90]	[0.79; 0.79]	[0.79; 0.79]	[0.96; 1.07]
70%	0.21	0.62	0.90	0.79	0.79	0.86
	[0.19; 0.24]	[0.56; 0.69]	[0.90; 0.90]	[0.79; 0.79]	[0.79; 0.79]	[0.81; 0.91]
80%	0.17	0.41	0.90	0.79	0.79	0.70
	[0.14; 0.19]	[0.36; 0.46]	[0.89; 0.90]	[0.79; 0.79]	[0.79; 0.79]	[0.65; 0.74]
90%	0.13	0.25	0.89	0.79	0.79	0.57
	[0.10; 0.13]	[0.16; 0.26]	[0.89; 0.89]	[0.79; 0.79]	[0.79; 0.79]	[0.51; 0.58]

Note: The mean absolute error (MAE) is reported in percentage points. Values in parentheses represent the 25th and 75th percentiles. The table compares the performance of various interpolation methods across different shares of reported tariff rates. Japan's average MFN tariff level, calculated from 1988 to 2021, is 5.22%. See main text for details about the alternative interpolation algorithms.

reducing the MAE by a factor of 1.85–3.13. Second, to outperform the alternative methods, the algorithm requires at least 20% of the tariff rates to be reported and achieves notable gains when reporting reaches 30%, with MAEs 1.2–1.3 times smaller than those of the alternative methods. Performance continues to improve as data availability increases, with the largest gains observed at 90% reporting, where the algorithm's MAE is 5.9–6.7 times smaller than that of the alternative approaches. Third, the alternative methods perform poorly at higher reporting shares, often yielding worse results than WITS when 80% or more of tariff schedules are reported. The comparison stresses the importance of explicitly accounting for trade policy features, such as gradual phase-ins vs. full implementation of agreements, sequential deepening of agreements, and overlap in agreements. Conceptually, the failure to incorporate these aspects is what primarily distinguishes the alternative interpolation methods from the algorithm.

Column (6) underscores the importance of the accuracy of the trade agreement data for the algorithm's performance. Here, instead of using the entry-into-force dates from my newly compiled dataset of RTAs, I replace them with information from CEPII's gravity database (Conte et al. 2022) and use linear interpolation to fill in the missing preferential tariff rates. Linear interpolation is used instead of the algorithm to isolate the effect of inaccuracies in the

RTA data: Using the algorithm with the CEPII RTA data confounds the algorithm's ability to account for trade policy with differences in the RTA data. The MAEs when the algorithm is run on the CEPII RTA data are substantially larger across most patterns of missingness, illustrating the critical role of the accuracy of trade agreement data in achieving reliable interpolation results.

4.4 Comparison with Previous Work

While many databases provide information on tariffs for individual countries (e.g., US ITC, TARIC for the EU, and ALADI for Latin American countries) or selected product categories (e.g., the Agricultural Market Access Database), few offer comprehensive coverage of cross-country tariff rates spanning a global set of countries, all products, and extended time periods. The Global Tariff Database (GTD) built with the new methodology introduced here provides unique coverage, including tariff rates for 200 countries, yearly data from 1988 to 2021, HS6-level disaggregation, and improved accuracy from the use of my novel algorithm to infill missing data. Among existing datasets, the most comparable are the data from Caliendo et al. (2023, CFRT) and CEPII's MAcMap-HS6 (Guimbard et al. 2012). Both make valuable contributions but differ significantly from the GTD in coverage and methodology, particularly in their handling of missing preferential tariff observations and data interpolation.

The CFRT dataset provides global tariff rates at the SITC4 level for 1984–2011, offering particular improvements on historical tariff data and especially detailed coverage for the US.[27] These data are much more aggregated, with the SITC4 classification containing only approximately 1,000 products, compared to over 5,000 in the HS6 classification. To fill in missing preferential tariff rates, CFRT manually review approximately 100 RTAs and GSP programs to identify their start dates and how the typical tariff preference was phased in. In contrast, the GTD includes 149 detailed phase-out schedules and relies on an algorithmic solution, which is significantly less error prone than CFRT's manual process. For example, the CFRT observations for Mexico's tariff on the US align with their GTD counterparts for years until 2007 but incorrectly spike to MFN levels for the years from 2008 to 2010. Additionally, as of December 2024, CFRT provide publicly available data only for country pair–industry combinations with positive trade flows.[28]

CEPII's MAcMap-HS6 offers tariff data for all importer–exporter–product combinations, effectively avoiding selection bias. The data are available for 2001–2019 but only at three-year

[27] Some US tariff rates are inferred from tariff revenues, reflecting de facto rather than statutory rates. This mix makes cross-country comparisons problematic since revenue-based tariff data are not broadly available and diverge from statutory rates because of factors such as costly rules of origin or antidumping duties.

[28] This description is based on the version available as of December 2024. The data are available at https://rcfeenstra.github.io/CFRT/.

intervals.[29] It is particularly valuable for non ad valorem tariffs and the calculation of ad valorem equivalents (AVEs) for tariff rate quotas, making it especially useful for analyzing sectors where AVEs are prevalent, such as agriculture. Furthermore, CEPII incorporates additional information through individual data-sharing agreements with countries, potentially broadening the scope of the reported tariff schedules; however, evaluating the extent of this enhancement is challenging because of the limited availability of documentation.

Despite these strengths, MAcMap-HS6 has notable limitations. Similarly to the CFRT dataset, it relies on manual review of agreement texts to determine whether tariffs are phased out, an approach prone to errors given the complexity of legal agreements. For instance, it incorrectly reports tariffs of 0% for new EU member states in 2001 despite documented exceptions.[30] Its interpolation algorithm also underperforms relative to the GTD's: By relying on data from $t-1$ and $t-2$, or $t+1$ if earlier data are unavailable, it struggles with extended reporting gaps, particularly if data are missing for many years at the beginning of the sample or right after an agreement enters into force. This issue is evident in cases such as Chile, which failed to report preferential tariffs for its EU agreement (which entered into force in 2003) until 2006. For 2004, MAcMap-HS6 incorrectly indicates that Chile continued to impose the MFN tariff on EU countries, despite tariff reductions already being in effect. Moreover, MAcMap-HS6's interpolation method does not fully capture tariff reductions in the year an agreement enters into force, as it does not perform backwards interpolation until the entry into force year. This limitation can result in missing the most significant drop in tariffs, i.e., specifically when the RTA takes effect.

The primary improvement of the GTD over existing datasets lies in its more precise imputation algorithm. Table 3 highlights the importance of incorporating trade policy features such as sequential deepening on trade agreements and the coexistence of multiple agreements, which the algorithms of MAcMap-HS6 and CFRT do not explicitly address. Consequently, unless manual corrections address these challenges comprehensively—a difficult task given the complexity of global trade policy—the filling algorithms used by CFRT and MAcMap-HS6 may fail to fully account for them.

Beyond its superior algorithm, the GTD offers much broader coverage, spanning 200 countries and HS6-level disaggregation over a 34-year period. This comprehensive scope enables the GTD to fully replicate the three papers analyzed in the replication study later in this paper—a task that neither MAcMap-HS6 nor CFRT could achieve because of their limited coverage. Moreover, the GTD's algorithmic approach, in contrast to the labor-intensive manual methods employed for other datasets, is highly adaptable, making it ideal for future updates and the

[29] This description is based on the version available as of December 2024. The data are available at http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=12.

[30] Official Journal of the European Communities No L114/27.

integration of new trade agreements. The GTD enhances the quality and reliability of tariff data, effectively addressing long-standing gaps in data availability that have hindered cross-country tariff analysis and policy evaluation in international economics, as emphasized by Bown and Crowley (2016).

5 Sources of Bias: Measurement Error and Selection

This section contains a general discussion of the bias in estimates of the effect of tariffs on trade based on the flawed WITS tariff data. The total bias is the net result of three competing forces: attenuation bias, differential measurement error, and sample selection.

5.1 Nonclassical Measurement Error

Suppose that the true relation between the exports x_{ij} and ad valorem tariffs $\tau_{ij} = ln(1 + t_{ij})$, where t_{ij} is the tariff imposed by importing country i on exports from country j, is given by:

$$x_{ij} = -\sigma \, \tau_{ij} + \epsilon_{ij} \tag{1}$$

and ϵ_{ij} is uncorrelated with τ_{ij}. For ease of notation, the product dimension k and the time index t are omitted, but all the results can be extended easily. However, because some of the preferential tariff rates are falsely interpolated with MFN tariff rates t_i, the observed tariff is not τ_{ij} but instead:

$$\tilde{\tau}_{ij} = \tau_{ij} + u_{ij},$$

with $u_{ij} = ln(1 + t_i)$ for share ω_{ij} of country pairs with an RTA.

The resulting biased ordinary least squares (OLS) estimator of $-\tilde{\sigma}$ is

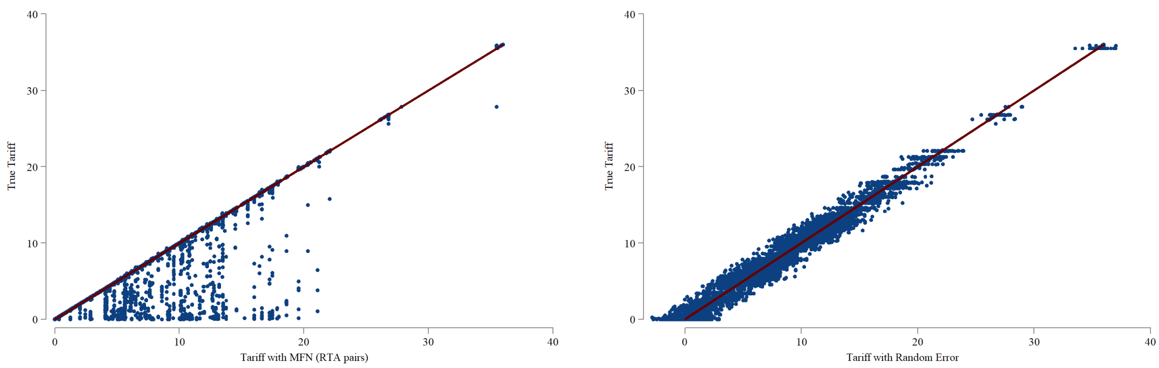
$$-\tilde{\sigma} = \frac{\text{Cov}(x_{ij}, \tilde{\tau}_{ij})}{\text{Var}(\tilde{\tau}_{ij})}$$

and can be rewritten as

$$-\tilde{\sigma} = -\sigma \underbrace{\left(\frac{\text{Var}(\tau_{ij}) + \text{Cov}(\tau_{ij}, u_{ij})}{\text{Var}(\tau_{ij}) + \text{Var}(u_{ij}) + 2 \cdot \text{Cov}(\tau_{ij}, u_{ij})} \right)}_{b_{\tau,\tilde{\tau}}} + \frac{\text{Cov}(\epsilon_{ij}, u_{ij})}{\text{Var}(\tau_{ij}) + \text{Var}(u_{ij}) + 2 \cdot \text{Cov}(\tau_{ij}, u_{ij})}$$

If the measurement error u_{ij} were uncorrelated with both τ_{ij} and ϵ_{ij}, i.e., $\mathrm{Cov}(\tau_{ij}, u_{ij}) = 0$ and $\mathrm{Cov}(\epsilon_{ij}, u_{ij}) = 0$, the biased estimate $-\tilde{\sigma}$ would be always attenuated, consistent with the result in the presence of classical measurement error. However, in the context of missing preferential tariff data and the resulting false imputations, these assumptions do not hold.

Under the WITS imputation method, missing preferential tariff observations are replaced with higher MFN tariff rates; therefore, the falsely interpolated tariffs will always be larger than the true ones. Figure 7 Panel (a) illustrates this relationship for the year 2001, plotting the true tariff τ_{ij} against their falsely imputed counterparts $\tilde{\tau}_{ij}$. If the error u_{ij} were indeed random, the data points would align around the 45-degree line, introducing solely noise to the explanatory variable but no systematic bias (Panel (b)). However, false interpolation occurs much more frequently for low tariff rates, as preferential rates are often zero or very low. Specifically, positive values of u_{ij} are associated with smaller true rates, making the measurement error nonclassical ($\mathrm{Cov}(\tau_{ij}, u_{ij}) < 0$).



(a) Nonclassical Measurement Error: $\tau \leq \tilde{\tau}$

(b) Random Measurement Error

Figure 7. Measurement Error Due to False Interpolation

Note: The figure shows the measurement error due to false interpolation. Panel (a) plots the true tariffs rates τ against the corrupted ones $\tilde{\tau}$ for the year 2001; Panel (b) shows how the two variables would be associated with each other if the measurement error were random with a mean of zero and variance of 1.

The term in parentheses can be compactly rewritten as the regression coefficient $b_{\tau,\tilde{\tau}}$. The size of $b_{\tau,\tilde{\tau}}$ depends on the identification strategy and its impact on the error structure but can always be recovered by a regression of the true on the falsely imputed tariffs. This is a general result that extends to other empirical approaches using OLS to estimate the tariff elasticity, including those that account for multilateral resistance terms, additional control variables or time-differencing—common strategies used in the literature to address endogeneity concerns. In the cross-sectional case, $b_{\tau,\tilde{\tau}}$ will be always smaller than 1 because $\tilde{\tau}_{ij}$ is always greater than τ_{ij} and can never become negative (see Panel (a)). Consequently, the negative covariance between τ_{ij} and u_{ij} ($\text{Cov}(\tau_{ij}, u_{ij}) < 0$) results in a stronger attenuation bias than would occur if the measurement error were classical.

Furthermore, $\text{Cov}(\epsilon_{ij}, u_{ij}) \neq 0$ because false interpolation occurs only for country pairs with an RTA; thus, tariffs with measurement error are not randomly distributed either across country pairs or over time. The literature on measurement error refers to this type of error as *differential measurement error* (e.g., Bound et al. 2001; Schennach 2022). Naturally, country pairs with a trade agreement differ systematically from those without; for instance, they are often geographically closer. Many variables included in the residual ϵ_{ij}, such as bilateral distance, income levels and consumer preferences, correlate with trade and tariffs in opposite directions. Consequently, using falsely interpolated tariff rates as an explanatory variable will most likely overstate the true effect of tariffs on trade.

5.2 Positive Selection

Even if false interpolation were not an issue (for example, because clean data are available), estimating $x_{ij} = -\sigma\tau_{ij} + \epsilon_{ij}$ from WITS data would still yield biased estimates because of positive selection. As previously discussed, WITS provides tariffs only for importing countries that report positive trade flows to UN Comtrade, and trade reporting patterns are systematically biased against low-income countries. This selection bias persists even when clean data are available, leading to overrepresentation of certain country pairs and underrepresentation of others.

This systematic pattern of missingness in tariff rates creates selection bias. Let $s_i = 1$ if all of importing country i's bilateral links are included in the WITS sample and $s_i = 0$ otherwise. s_i can be rewritten as a function of income gdp_i and other unobservables e_i, such that the selection equation is:

$$s_i = \gamma\; gdp_i + e_i$$

Estimating Equation (1) in the available sample yields a biased estimate of the elasticity $-\sigma$ whenever the sensitivity of trade flows to tariff changes is heterogeneous across countries.[31] In addition to the true value of $-\sigma$, the biased coefficient contains the selection bias π:

$$-\tilde{\sigma} = -\sigma + \underbrace{\frac{\text{Cov}(\epsilon_{ij}, \tau_{ij}|s_i = 1)}{\text{Var}(\tau_{ij}|s_i = 1)}}_{\text{selection}} = -\sigma + \pi$$

[31] Relatedly, Soderbery (2018) demonstrates that import demand elasticities vary across countries.

If the sample selection were indeed random, π would equal zero, and the elasticity estimate would be unbiased. However, this is not the case, as trade flows and tariffs are systematically different for those $ijkt$ observations that are not included in the WITS-sample.

While the conditional covariance and variance are not directly observable, we can identify the selection bias by comparing the estimated coefficient from the WITS sample (using the new data) to the true elasticity $-\sigma$ derived from the full dataset. The extent of the bias depends on the sample composition of the respective empirical specification. If the sample is limited to observations included in WITS, the bias will be minimal. By contrast, if the empirical strategy relies on tariff variation from low-income countries, incorporating the missing observations will have a greater impact. In the cross-sectional setting, the selection bias likely reduces the estimated tariff elasticity ($-\tilde{\sigma} < -\sigma$) because low-income countries, which are underrepresented in WITS, participate less in international trade and tend to impose higher tariffs on average.

5.3 Decomposition

Decomposing the total bias $\Delta\,\sigma$ into its three components requires comparing $-\tilde{\sigma}_s$ (the coefficient estimated from the WITS data), $-\sigma_s$ (the coefficient estimated from data corrected for false interpolation) and the true $-\sigma$ (the coefficient estimated from data corrected for false interpolation and selection), and the signal-to-noise ratio $b_{\tau,\tilde{\tau}}$. The total bias $\Delta\,\sigma$, defined as the difference between $-\tilde{\sigma}_s$ and $-\sigma$, consists of two parts: the bias attributable to false interpolation ($\Delta\,\sigma^M$) and that attributable to selection ($\Delta\,\sigma^S$).

$$\Delta\,\sigma = \Delta\,\sigma^M + \Delta\,\sigma^S = \underbrace{\overbrace{-\sigma_s(b_{\tau,\tilde{\tau}} - 1)}^{\textit{attenuation bias } \lambda \uparrow} + \overbrace{\Omega}^{\textit{differential ME} \downarrow}}_{\text{false interpolation} \uparrow \text{ or } \downarrow} + \underbrace{\pi}_{\text{selection} \uparrow \text{ or } \downarrow} \tag{2}$$

The bias attributable to false interpolation $\Delta\,\sigma^M$ is given by $-\tilde{\sigma}_s - (-\sigma_s)$ and can be further decomposed into its two components. The first isolates the attenuation bias λ, with its magnitude depending on the signal-to-noise ratio $b_{\tau,\tilde{\tau}}$ and the unbiased $-\sigma$, making it straightforward to determine. The second part of the bias comes from the differential character of the measurement error, denoted by Ω, i.e., $\mathrm{Cov}(\epsilon_{ij}, u_{ij}) \neq 0$. Its magnitude is calculated as the difference between the total bias due to interpolation ($\Delta\,\sigma^M$) and the attenuation bias λ. Finally, the selection bias is determined as $\Delta\,\sigma - \Delta\,\sigma^M$.

6 Revisiting the Trade Elasticity

In this section, I replicate three prominent studies estimating the trade elasticity—Arkolakis et al. (2018), Caliendo and Parro (2015), and Boehm et al. (2023)—to demonstrate the practical importance of correcting the flawed tariff data. I focus on estimates of the trade elasticity because it is a key parameter in international trade and, as highlighted in the systematic survey in Section 3, an empirical setting where the literature particularly struggles to address the issues in the WITS tariff data.[32] These three studies vary in their empirical strategies and approaches to handling missing tariff observations, leading to stark differences in the sensitivity of their results to the flaws in the tariff data provided by WITS.

6.1 Arkolakis et al. (2018)

Arkolakis et al. (2018, ARRY) develop a new model of trade and multinational production that they use to quantify the welfare implications of shocks driving increased specialization in innovation and production. The model delivers a novel testable implication—namely, that trade flows restricted to the parents and affiliates of firms from a given country are more sensitive to trade costs than overall, or unrestricted, trade flows. Using data from the Bureau of Economic Analysis (BEA) on the sales of US firms and their foreign affiliates, the authors take their theory to the data and estimate restricted and unrestricted trade elasticities using gravity equations. The two estimated elasticities then serve as key targets in the calibration of their model. ARRY rely on cross-sectional variation in tariff rates to estimate the restricted and unrestricted trade elasticities. The gravity regressions use unbalanced pair-level data for 62 countries in 1999 and account for the multilateral resistance terms with importer and exporter fixed effects (μ_i and μ_j). In addition, a vector of standard gravity controls $\boldsymbol{Z_{ij}}$ (bilateral distance, contiguity, same language, colonial history, and domestic trade) is included.

For the restricted gravity estimates, the data on firm-level trade flows are not available for reasons of confidentiality. Therefore, I analyze only how sensitive the results of the unrestricted gravity equation are to the flaws in the WITS data caused by the missing tariff observations. The regression equation for the unrestricted gravity model is

$$lnX_{ij} = -\sigma\tau_{ij} + \gamma\boldsymbol{Z_{ij}} + \mu_i + \mu_j + \epsilon_{ij} \tag{3}$$

To overcome the inherently difficult task of compiling bilateral tariff data, ARRY downloaded MFN tariffs from WITS and combined them with a dummy variable indicating whether coun-

[32] The analysis of Brancaccio et al. (2020) is excluded from the replication exercise as their dependent variable relies on confidential data unavailable to me.

tries i and j have a trade agreement. They assume that whenever a trade agreement exists, the applied tariff equals zero for all products k; otherwise, the MFN tariff applies. While interpolating missing preferential tariff rates as zero may appear reasonable, this approach introduces measurement error. For preferential tariffs that are greater than zero (whether because of gradual phase-outs or final rates lower than the MFN rate but higher than zero) or products that are exempted, this method risks underestimating the true tariff rates. Figure B.8 compares the proposed solution for constructing preferential tariffs with my algorithm, using the simulated missing patterns in the Japanese tariff data introduced in Section 4.3. The ARRY interpolation approach produces 3 to 33 times higher errors than the GTD, depending on the share of missing tariffs, indicating it is not ideal for addressing missing preferential tariffs.

The mismeasurement of preferential tariff rates is exacerbated by ARRY's reliance on incomplete trade agreement data, a common challenge given the limitations of the available datasets on RTAs. For instance, the association agreements between the EU and Eastern European countries that acceded in 2004 were already in effect by 1999, yet ARRY assume that the United Kingdom—an EU member state during their sample period—applied MFN rates to these countries. This yields upward bias in their measure of preferential tariffs, partially offsetting the downward bias introduced by the zero-filling method. This source of bias is of the same type as the one introduced in Table 3 Column (6) of the validation exercise, which demonstrates that relying on incomplete RTA data adds noise to the interpolated tariff rates, reducing its reliability.

Additionally, ARRY acknowledge the selection issue in the WITS data at the country pair–product level. As noted earlier, WITS reports MFN tariff rates only for $ijkt$ observations with positive trade flows recorded in UN Comtrade. To address the resulting missing observations, ARRY infer them using the average MFN tariff of importer i for product k at the 6-digit level, based on the MFN tariffs reported by WITS. However, due to their incomplete RTA data, for pair–products that actually have a preferential tariff rate, this interpolation step is effectively identical to the flawed interpolation performed by WITS. After these interpolation steps, the disaggregated tariff data are aggregated to the country pair level. These data manipulation steps result in an error that more closely resembles classical measurement error; the importer and exporter fixed effects amplify this (see Figure B.7).

Column (1) in Table 4 presents the results of rerunning Equation (3) with the data and code provided by the authors for replication. I successfully reproduce the paper's coefficient estimate of -4.28.[33] To assess the bias introduced by the flaws in the tariff data, I use the newly constructed HS6-level tariff data, aggregate them by averaging across all products within a

[33]The authors kindly provided additional information about the relevant sample, which is necessary to replicate their estimates but not included in the replication files. The standard errors differ slightly because the authors ran a seemingly unrelated regression using both restricted and unrestricted data, which are not available to me.

pair, and re-estimate Equation (3). This provides an estimate of $-\sigma$, the true tariff elasticity, unaffected by false interpolation or positive selection. To decompose the total bias into its distinct components, I first isolate the effect of false interpolation by restricting the sample at the HS6 level to pair–product observations ijk available in WITS, ensuring that only the bias from false interpolation is corrected while retaining the bias due to selection. This subsample of bilateral HS6 tariffs is then averaged to the pair level and used to estimate $-\sigma_s$.

Table 4. Reestimating the Tariff Elasticity: Results

Paper	Coefficients			Decomposition		
	$-\tilde{\sigma}_s$	$-\sigma_s$	$-\sigma$	$\Delta\ \sigma$	$\Delta\ \sigma^M$	$\Delta\ \sigma^S$
	(1)	(2)	(3)	(4)	(5)	(6)
Arkolakis et al. (2018)	-4.28*	-8.95***	-9.81***	5.53	4.67	0.86
	(2.36)	(3.03)	(3.24)		(85%)	(15%)
Obs.	295	295	301			
Caliendo and Parro (2015)	-4.55***	-6.15***	-4.48***	0.07	1.60	-1.66
	(0.35)	(0.62)	(1.00)		(-2333%)	(2433%)
Obs.	7,205	7,205	8,185			
Boehm et al. (2023)						
— *short run*	-0.62***	-0.47***	-0.46***	-0.17	-0.15	-0.02
	(0.12)	(0.13)	(0.06)		(88%)	(12%)
Obs. (in Mio.)	20.9	20.9	51.6			
— *long run*	-2.14***	-1.67***	-0.80***	-1.76	-0.47	-0.80
	(0.34)	(0.35)	(0.25)		(36%)	(64%)
Obs. (in Mio.)	7.5	7.4	19.8			

Note: The table shows the results when I replicate and reestimate Arkolakis et al. (2018, Table 1), Caliendo and Parro (2015, Table 1), and Boehm et al. (2023, Figure 2, horizons 1 and 10). Column (1) reports the replicated results using the original code and data provided by the authors. Column (2) reestimates the analysis, correcting only for false interpolation in the tariff rates, while Column (3) incorporates corrections for both false interpolation and selection bias. Columns (4) to (6) decompose the sources of bias.

The results suggest that ARRY's data underestimate the true $-\sigma$ by 5.53, yielding a coefficient of -9.81 after correction for both sources of bias. Interestingly, the corrected tariff data produce estimates consistent with those from ARRY's alternative empirical approach, which does not rely on tariff variation and yields a coefficient of -9.7. The interpolation steps employed by the authors eliminate the positive selection but introduce other forms of measurement error. A comparison of columns (1) to (3) reveals the contribution of each interpolation step to the total bias: the bias due to falsely assuming zero tariffs for all products of pairs with an RTA attenuates the coefficient by 4.67 while the bias due to incomplete RTA data, i.e., assuming MFN tariffs instead of preferential tariffs, contributes an additional attenuation of 0.86.

6.2 Caliendo and Parro (2015)

In their seminal work, Caliendo and Parro (2015, CP) include sectoral linkages, trade in intermediate goods, and sectoral heterogeneity in production in a Ricardian model to quantify the trade and welfare effects of tariff changes. In their quantification exercise, the trade elasticity is one of the central parameters. Therefore, CP propose a new method for estimating the trade elasticity that has minimal data requirements and is consistent with any trade model that delivers a gravity equation.

For the estimation of $-\sigma$, they consider three countries indexed by n, i, and h and take the cross-product of goods shipped in one direction between the three countries, from n to i, from i to h, and from h to n, and then the cross-product of the same goods shipped in the other direction, from n to h, from h to i, and from i to n. Employing the gravity equation and taking the ratio, it follows that

$$\frac{X_{ni}X_{ih}X_{hn}}{X_{nh}X_{hi}X_{in}} = \left(\frac{\kappa_{ni}\kappa_{ih}\kappa_{hn}}{\kappa_{nh}\kappa_{hi}\kappa_{in}}\right)^{-\sigma} \tag{4}$$

All the terms involving prices and other parameters cancel out, and bilateral trade is a function only of trade costs κ, which consist of tariffs $\tau_{ni} = ln(1 + t_{ni})$ and nontariff trade costs d_{ni}. The nontariff trade costs can be modeled quite generally as a linear function of bilateral symmetric trade costs $\mu_{ni} = \mu_{in}$ and importer and exporter characteristics ν_i and ν_n. All remaining asymmetric trade costs incurred in exporting from n to i are summarized in ϵ_{ni}. Hence, κ_{ni} can be rewritten as

$$ln(\kappa_{ni}) = ln(\tau_{ni}) + ln(d_{ni}) = ln(\tau_{ni}) + \mu_{ni} + \nu_n + \nu_i + \epsilon_{ni}$$

The main regression equation follows after we take logs and plug the nontariff trade costs into (4).

$$ln\left(\frac{X_{ni}X_{ih}X_{hn}}{X_{nh}X_{hi}X_{in}}\right) = -\sigma\ ln\left(\frac{\tau_{ni}\tau_{ih}\tau_{hn}}{\tau_{nh}\tau_{hi}\tau_{in}}\right) + \left(\frac{\epsilon_{ni}\epsilon_{ih}\epsilon_{hn}}{\epsilon_{nh}\epsilon_{hi}\epsilon_{in}}\right) \tag{5}$$

CP run their regressions at the sector level of the World Input–Output Database (WIOD). They download industry-level tariff rates for International Standard Industrial Classification (ISIC) industries from WITS, provided as unweighted averages of HS6-level product data, and aggregate them to WIOD sectors. However, any aggregated version of the bilateral tariff rates in WITS is inherently flawed because it relies on underlying product-level data affected by both false interpolation and selection biases. For the year 1993, even at the aggregated ISIC-level, the data suffer from missing observations. To address these gaps, CP propose filling in the missing tariff rates with values from preceding and subsequent years. However, because this interpolation is performed at the ISIC industry level, where tariff rates are substantially mismeasured because of WITS's mishandling of the missing data, their method carries the bias to other years.

To obtain tariff rates at the WIOD sector level that maintain the WITS sample consistency, I replicate the procedure employed by WITS when researchers download data. Specifically, for each year between 1989 and 1995, I keep only the HS6 products that are also included in WITS. Using this selected sample, I aggregate to the ISIC industry level and then apply the infilling algorithm proposed by CP. The resulting data are further aggregated to the WIOD sector level. Regressing Equation (5) with these tariffs yields $-\sigma_s$, which corrects for false interpolation but not for the positive selection inherent in WITS. To determine the true $-\sigma$, free from sample selection bias, I use the tariff rates from the GTD for 1993 and aggregate them to WIOD sectors.[34]

The replication study shows that, for CP, the biases due to false interpolation and positive selection nearly cancel each other out (-4.55 vs. -4.48 in columns (1) and (3)). When the sample selection bias inherent in the WITS data is left unaddressed but the false interpolation is corrected, the coefficient increases in absolute magnitude from -4.55 to -6.15. Including tariff observations for which the corresponding trade flows are zero when averaging to the industry-level pushes the tariff elasticity up to -4.48.

Excluding tariff observations for which the trade flows are zero pushes the tariff elasticity upward because high tariffs are often associated with low trade flows. Specifically, industries with prohibitively high tariffs at the HS6 level tend to have lower total trade flows than industries where all HS6 products are traded. Thus, the WITS sample selection bias negatively correlates with trade flows. Additionally, excluding HS6-level tariff observations with zero trade flows before aggregating to the industry level yields lower average industry-level tariffs than when include them. As a result, excluding HS6 products with zero trade flows before aggregation to the industry level leads to an overstatement of the tariff elasticity, such that $-\sigma_s < -\sigma$.

6.3 Boehm et al. (2023)

Boehm et al. (2023, BLP) introduce an innovative approach to estimate the trade elasticity across different time horizons by addressing the endogeneity of trade policy through a novel identification strategy. To account for omitted variables that vary by country pair–product, they time-difference the data. Changes in trade policy might still be endogenous, as, for

[34] As in CP (for the baseline results), all aggregation steps are done with unweighted averages.

example, policymakers might lower tariffs in response to a large increase in trade. To address these kinds of endogeneity issues, BLP propose a very innovative identification strategy: They compare imports from countries with MFN relations to imports from countries unaffected by changes in MFN tariffs, i.e., countries with a trade agreement. Crucially, the estimates are based only on the response of minor exporters, for which changes in MFN tariffs are plausibly exogenous.[35] Furthermore, BLP account for autocorrelation in tariff changes using local projections (Jordà 2005). The exogenous year-on-year changes in MFN tariffs are used as an instrument for future tariff changes to identify responses of trade at different horizons.

The instrument z_{ijkt} equals the change in MFN tariff rates between t and $t-1$ if country pair–products have MFN relations at t and $t-1$ and equals zero otherwise (see Equation (6)).

$$\begin{aligned} z_{ijkt} = & \left(ln\tau_{ikt}^{MFN} - ln\tau_{ikt-1}^{MFN}\right) \\ & \times \mathbf{1}\left(\tau_{ijkt} = \tau_{ikt}^{MFN}\right) \times \mathbf{1}\left(\tau_{ijkt-1} = \tau_{ikt-1}^{MFN}\right) \end{aligned} \tag{6}$$

The instrument is used to estimate the following first stage, with major partners excluded:

$$\Delta_h ln\tau_{ijkt} = \beta^h z_{ijkt} + \delta_{it}^{sh} + \delta_{jt}^{sh} + \delta_{ij}^{sh} + u_{ijkt}^h \tag{7}$$

$\Delta_h \ln \tau_{ijkt}$ denotes the time difference in $\ln \tau_{ijkt}$ between periods $t-1$ and $t+h$. δ_{it}^{sh} and δ_{jt}^{sh} are importer–HS4–year and exporter–HS4–year fixed effects, respectively, while δ_{ij}^{sh} represents importer–exporter–HS4 fixed effects.

The second stage is then

$$\Delta_h lnX_{ijkt} = -\sigma^h \Delta_h ln\tau_{ijkt} + \delta_{it}^{sh} + \delta_{jt}^{sh} + \delta_{ij}^{sh} + u_{ijkt}^h \tag{8}$$

The analysis is conducted at the HS6-level, and the only data correction step BLP take to address the missing tariff observations is to set the bilateral tariffs to zero for preferential trade within the European Union. Hence, the analysis is plagued by the false interpolation and sample selection done by WITS.

The false interpolation leads to a corrupted instrument $\tilde{z}_{ijkt}$: For importers i that fail to report their preferential tariff rates with partner j in two consecutive years, BLP falsely assume no trade agreement exists. Consequently, the instrument equals to the change in MFN rates instead of zero, i.e., $\tilde{z}_{ijkt} = z_{ijkt} + u_{ijkt}^z$. Put differently, the false interpolation misclassifies observations that should belong to the control group, i.e., those within an RTA, by incorrectly

[35] The exclusion of large trade partners is important for identification because changes in their trade policy might be endogenous—for example, because of lobbying activity by multinational enterprises (Blanchard and Matschke 2015).

assuming that they trade under MFN relations. The resulting error u^z_{ijkt} may be positive or negative, as it depends on whether MFN rates increased or decreased in the respective years in which the country fails to report its tariffs. However, in practice, most countries—particularly those with a preference for free trade, as indicated by their participation in RTAs—tend to lower their MFN tariffs over time on average. If the misclassified pairs had zero or negative growth in trade flows $\Delta_h lnX_{ijk}$, the measurement error in the instrument would attenuate the reduced form. This assumption is plausible for pairs that have been in an RTA long enough for trade growth to stabilize; the literature suggests that RTAs require 5–17 years for full adjustment (Alessandria et al. 2024; Anderson and Yotov 2022; Baier and Bergstrand 2007). However, for recently implemented RTAs, trade growth may still be unfolding, as the effects of tariff changes and associated trade adjustments take time to materialize. In such cases, leveraging $\tilde{z}_{ijkt}$ will overstate the true effect.

The first stage coefficient β^h decreases when false interpolation is corrected (see Table B.1): following an initial impulse, 74% of the change persists after one year and 64% after ten years with GTD tariffs, compared to 82% and 68%, respectively, with WITS data when keeping the sample constant. The discrepancy arises from the artificially high persistence in WITS data. For misclassified observations the WITS data predict a future tariff reduction when importer i reports the true (lower) preferential tariffs in $t + h$. Put differently, the dependent variable in the first stage suffers from false interpolation as well, i.e., $\Delta_h ln\tilde{\tau}_{ijkt} = \Delta_h ln\tau_{ijkt} + \Delta_h u_{ijkt}$. For misclassified pair–product observations, the base-year tariff $\ln \tau_{ijk,t-1}$ used to calculate future tariff changes $\Delta_h ln\tilde{\tau}_{ijkt}$ equals the false MFN rate. This results in a measurement error $\Delta_h u_{ijkt}$, which is negative when preferential rates are correctly reported in $t + h$, as tariffs mechanically drop from high MFN rates to low preferential rates in $t + h$. As a result, the h-horizon change in tariffs is negative and often large for these misclassified pairs, inflating the first-stage coefficient $\tilde{\beta}^h$. When also correcting for the selection bias, the long-run first stage coefficient is much lower (0.56) indicating even stronger mean reversion than in smaller sample.

Table 4, Columns (1) to (3) give the result of the replication analysis when I rerun the code made available by the authors and use the data provided by WITS and the GTD for the short-run (horizon $h = 1$) and long-run (horizon $h = 10$) elasticity estimates.[36] Correcting for both the sample selection and false interpolation issues in the WITS data shrinks the absolute value of the long-run tariff elasticity from -2.14 to -0.80 and that of the short-run elasticity from -0.62 to -0.46. For the long-run estimate, 64% of this change is driven by the correction for sample bias. Figure B.12 summarizes the results for the other time horizons and also highlights the

[36]The replicated coefficients in column (1) differ slightly from those reported in BLP because of adjustments to the sample for comparability with the GTD, which includes only HS6 products in the 1988/92 nomenclature. The adjusted BLP sample yields estimates nearly identical to those in the original paper (see Figure B.11).

importance of correcting for sample selection: while both corrections—for false interpolation and for selection—lead to lower absolute values of the tariff elasticities, including the full sample instead of the selected WITS sample yields the largest change in coefficients for all time horizons, i.e., $-\tilde{\sigma}_s < -\sigma_s < -\sigma$. This result suggests treatment heterogeneity across ijk observations that are included in the WITS sample relative to those that are not.

The results of BLP are sensitive to corrections for the selection bias in the tariff data. This is not surprising, given their identification strategy explicitly relies on small exporters, as major trade partners are excluded. Notably, the number of available observations nearly triples for the long-run analysis, rising from 7 million to 20 million, and increases by a factor of 2.5 in the short-run, from 21 million to 52 million observations.

One key take away of the results is that the pair–product observations omitted from the WITS sample react less to changes in tariffs. Potential explanations for these lower effects include the possibility that tariffs represent only a small fraction of total costs to exporters. For instance, Sequeira (2016), in the context of Mozambique, finds that a sizable reduction in statutory tariff rates has little impact on imports due to pervasive corruption. In such an environment, a substantial tariff liberalization translates into only small changes in trade costs, resulting in low estimates of the tariff elasticity. Costly rules of origin present another potential explanation for the low tariff elasticity. Preferential tariffs might not be used by all exporters due to the high costs of complying with rules of origin. In this case, a change in MFN tariffs would affect not only the treatment group (pairs with MFN relations) but also the control group. This overlap would make it impossible to identify a differential effect between the two groups. The costs of complying with the rules might be too high for low trade flows, but they may be worthwhile for larger trade flows, leading to a heterogeneous treatment effect when including the pair–product observations omitted from the WITS sample.

The lower first-stage coefficient would yield, all else equal, an attenuated $-\tilde{\sigma}_s$. Thus, the consistently larger absolute magnitude of the tariff elasticity estimated by BLP, $-\tilde{\sigma}_s < -\sigma_s$, arises from a lower reduced-form estimate when correcting for false interpolation in the WITS data. Misclassified pair–product observations are, in reality, part of an RTA and therefore exhibit high trade growth over extended periods due to gradual adjustment of trade following tariff liberalizations. Hence, for the misclassified observations, trade growth is large and positive, while simultaneously the instrument $\tilde{z}_{ijkt}$ is often negative as MFN tariffs decrease, leading to an overstated reduced form. Furthermore, the sample restriction for the long-run (1997–2008) further amplifies this bias, as many trade agreements were signed shortly before or during the early years of this period; if data were available for a longer period, the control group would include more country pairs whose trade had already stabilized following earlier tariff cuts.[37]

7 Conclusion

This paper demonstrates that a widely used tariff dataset suffers from significant nonclassical measurement error and systematic selection bias due to missing tariff observations. Many countries fail to report their tariff rates annually, with compliance systematically lower for preferential than for MFN tariffs. The World Bank's WITS, a primary source of tariff data, interpolates the missing preferential tariff rates with MFN tariff rates, creating artificial spikes in bilateral time series and introducing substantial measurement error. Moreover, WITS provides tariffs only for importers reporting trade to UN Comtrade, such that low-income countries are underrepresented and tariff rates for products with zero trade flows are excluded. To address these issues, I propose a new interpolation algorithm that accounts for misreporting and combine five data sources to create a global tariff dataset at the six-digit product level, covering 200 importers and their partners over 34 years. Reestimating three prominent trade elasticity studies with the corrected data reveals significant changes in their results, underscoring how important it is for researchers to use reliable tariff data.

The replication study highlights that while correcting tariff rate errors improves data quality, it does not resolve the trade elasticity puzzle—uncertainty about the magnitude of this parameter remains. Interestingly, correcting these errors increases the variance in the trade elasticity estimates across the three studies, suggesting that the challenges extend beyond data quality. These findings point to issues such as treatment heterogeneity across countries or sectors, aggregation biases, and a broader need for rigorous identification strategies in future research. A promising direction could involve moving away from pooled cross-country data and focusing instead on cases where tariff variation is clearly exogenous, as exemplified by studies on the Trump tariffs (e.g., Fajgelbaum et al. 2020) or the EU Eastern Enlargement (Sandkamp 2020).

While this paper advances the coverage of statutory tariff rates, a critical step toward understanding how tariffs impact international trade, most analyses ultimately require data on the *de facto* tariffs levied, which reflect the tariff rate that exporters actually face. Discrepancies between statutory and de facto tariff rates may arise from costly rules of origin, which often lead exporters to underutilize tariff preferences, resulting in higher effective tariffs (e.g., Conconi et al. 2018). Exemptions, such as de minimis thresholds (Fajgelbaum and Khandelwal 2024), and other barriers, including antidumping duties and safeguards not captured in this dataset (cf Bown and Crowley 2016, for an overview of other types of tariffs), further contribute to

[37] According to the WTO's RTA Database, the number of RTAs in force increased from 55 in 1997 to 179 in 2008.

these discrepancies. Accounting for the gap between statutory and de facto tariff rates is an important avenue for future research, offering a more accurate picture of the trade policy environment.

References

Alessandria, George A., Shafaat Khan, Armen Khederlarian, Kim J. Ruhl, and Joseph Steinberg (2024), "Trade-Policy Dynamics: Evidence from 60 Years of U.S.-China Trade." *Journal of Political Economy*, forthcoming.

Alfaro, Laura, Paola Conconi, Harald Fadinger, and Andrew F. Newman (2016), "Do Prices Determine Vertical Integration?" *Review of Economic Studies* 83.3, pp. 855–888.

Anderson, James E. and J. Peter Neary (1994), "Measuring the Restrictiveness of Trade Policy." *The World Bank Economic Review* 8.2, pp. 151–169.

Anderson, James E. and Eric Van Wincoop (2004), "Trade Costs." *Journal of Economic Literature* 42.3, pp. 691–751.

Anderson, James E. and Yoto V. Yotov (2022), "Estimating Gravity from the Short to the Long Run: A Simple Solution to the 'International Elasticity Puzzle'." *NBER Working Paper No. 30809.*

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare (2012), "New Trade Models, Same Old Gains?" *American Economic Review* 102.1, pp. 94–130.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple (2018), "Innovation and Production in the Global Economy." *American Economic Review* 108.8, pp. 2128–2173.

Bagwell, Kyle and Robert W. Staiger (2011), "What Do Trade Negotiators Negotiate About? Empirical Evidence from the World Trade Organization." *American Economic Review* 101.4, pp. 1238–1273.

Bagwell, Kyle, Robert W. Staiger, and Ali Yurukoglu (2021), "Quantitative Analysis of Multiparty Tariff Negotiations." *Econometrica* 89.4, pp. 1595–1631.

Baier, Scott L. and Jeffrey H. Bergstrand (2007), "Do Free Trade Agreements actually Increase Members' International Trade?" *Journal of International Economics* 71.1, pp. 72–95.

Baier, Scott L., Jeffrey H. Bergstrand, and Michael Feng (2014), "Economic Integration Agreements and the Margins of International Trade." *Journal of International Economics* 93.2, pp. 339–350.

Barari, Soubhik Barari and In Song Kim (2020), "Democracy and Trade Policy at the Product Level."

Besedes, Tibor, Tristan Kohl, and James Lake (2020), "Phase out tariffs, phase in trade?" *Journal of International Economics* 127.

Blanchard, Emily J. and Xenia Matschke (2015), "U.S. Multinationals and Preferential Market Access." *Review of Economics and Statisticss* 96.4, pp. 839–854.

Boehm, Christoph E., Andrei A. Levchenko, and Nitya Pandalai-Nayar (2023), "The Long and Short (Run) of Trade Elasticities." *American Economic Review* 113.4, pp. 861–905.

Boehm, Johannes, Swati Dhingra, and John Morrow (2022), "The Comparative Advantage of Firms." *Journal of Political Economy* 130.12, pp. 3025–3100.

Bouët, Antoine, Yvan Decreux, Lionel Fontagné, Sébastien Jean, and David Laborde (2008), "Assessing Applied Protection across the World." *Review of International Economics* 16.5, pp. 850–863.

Bound, John, Charles Brown, Greg J. Duncan, and Willard L. Rodgers (1994), "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data." *Journal of Labor Economics* 12.3, pp. 345–368.

Bound, John, Charles Brown, and Nancy Mathiowetz (2001), "Measurement Error in Survey Data." *Handbook of Econometrics.* Vol. 5. Elsevier Masson SAS, pp. 3705–3843.

Bound, John and Alan B. Krueger (1991), "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9.1, pp. 1–24.

Bown, Chad P. and Meredith A. Crowley (2016), "The Empirical Landscape of Trade Policy." *Handbook of Commercial Policy.* Ed. by Kyle Bagwell and Robert W. Staiger. Vol. 1A. Elsevier B.V. Chap. 1, pp. 3–108.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou (2020), "Geography, Transportation, and Endogenous Trade Costs." *Econometrica* 88.2, pp. 657–691.

Brandt, Loren, Johannes Van Biesebroeck, Luhang Wang, and Yifan Zhang (2019), "WTO Accession and Performance of Chinese Manufacturing Firms." *American Economic Review* 109.4, pp. 1616–1621.

Broda, Christian and David E. Weinstein (2006), "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121.2, pp. 541–585.

Burstein, Ariel and Jonathan Vogel (2017), "International Trade, Technology, and the Skill Premium." *Journal of Political Economy* 125.5, pp. 1356–1412.

Bustos, Paula (2011), "Trade Liberalization, Exports and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinean Firms." *American Economic Review* 101.1, pp. 304–340.

Caliendo, Lorenzo, Robert C. Feenstra, John Romalis, and Alan M. Taylor (2023), "Tariff Reductions, Heterogeneous Firms, and Welfare: Theory and Evidence for 1990-2010." *IMF Economic Review* forthcomin.

Caliendo, Lorenzo and Fernando Parro (2015), "Estimates of the Trade and Welfare Effects of NAFTA." *Review of Economic Studies* 82.1, pp. 1–44.

Caliendo, Lorenzo, Fernando Parro, Luca David Opromolla, and Alessandro Sforza (2021), "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement." *Journal of Political Economy* 129.12, pp. 3491–3545.

Conconi, Paola, Laura Puccio, Manuel García-Santana, Roberto Venturini, Laura Puccio, and Venturini (2018), "From Final Goods to Inputs: the Protectionist Effect of Rules of Origin." *American Economic Review* 108.8, pp. 2335–2365.

Conte, Maddalena, Pierre Cotterlaz, and Thierry Mayer (2022), *The CEPII Gravity Database*. Tech. rep. 2022-05.

Costinot, Arnaud and Andrés Rodríguez-Clare (2014), "Trade Theory with Numbers: Quantifying the Consequences of Globalization." *Handbook of International Economics*. Vol. 4. Elsevier B.V., pp. 197–261.

Disdier, Anne-Célia and Keith Head (2008), "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *The Review of Economics and Statistics* 90.1, pp. 37–48.

Dix-Carneiro, Rafael and Brian K. Kovak (2017), "Trade Liberalization and Regional Dynamics." *American Economic Review* 107.10, pp. 2908–2946.

Dür, Andreas, Leonardo Baccini, and Manfred Elsig (2014), "The Design of International Trade Agreements: Introducing a New Dataset." *Review of International Organizations* 9.3, pp. 353–375.

Ederington, Josh and Michele Ruta (2016), "Nontariff Measures and the World Trading System." *Handbook of Commercial Policy 1b*. Ed. by Kyle Bagwell and Robert W. Staiger. 1st ed. Vol. 1A. Elsevier B.V., pp. 211–277.

Egger, Peter H. and Mario Larch (2008), "Interdependent Preferential Trade Agreement Memberships: An Empirical Analysis." *Journal of International Economics* 76.2, pp. 384–399.

Fajgelbaum, Pablo D., Pinelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal (2020), "The Return to Protectionism." *The Quarterly Journal of Economics* 135.1, pp. 1–55.

Fajgelbaum, Pablo D. and Amit Khandelwal (2024), "The Value of De Minimis Imports." *NBER Working Paper No. 32607*.

Feenstra, Robert C. (1994), "New Product Varieties and the Measurement of International Prices." *The American Economic Review* 84.1, pp. 157–177.

Feenstra, Robert C. and John Romalis (2014), "International Prices and Endogenous Product Quality." *Quarterly Journal of Economics* 129.2, pp. 477–527.

Fink, Carsten, Aaditya Mattoo, and Ileana Cristina Neagu (2005), "Assessing the Impact of Communication Costs on International Trade." *Journal of International Economics* 67.2, pp. 428–445.

Fontangé, Lionel, Houssein Guimbard, and Gianluca Orefice (2022), "Tariff-Based Product-Level Trade Elasticities." *Journal of International Economics* 137, p. 103593.

Forero-Rojas, Alejandro, Hibret Maemir, Ana M. Fernandes, and Aaditya Mattoo (2018), "Developing Countries' Trade and Market Access in the European Union and the United States: Introducing Two New Databases." *mimeo*.

Gaulier, Guillaume and Soledad Zignago (2010), *BACI: International Trade Database at the Product-Level. The 1994-2007 Version*. Working Papers. CEPII.

Guimbard, Houssein, Sébastien Jean, Mondher Mimouni, and Xavier Pichot (2012), "MAcMap-HS6 2007, An Exhaustive and Consistent Measure of Applied Protection in 2007." *International Economics* 130, pp. 99–121.

Handley, Kyle and Nuno Limão (2017), "Policy Uncertainty, Trade and Welfare: Theory and Evidence for China and the United States." *American Economic Review* 107.9, pp. 2731–2783.

Head, Keith and Thierry Mayer (2014), "Gravity Equations: Workhorse,Toolkit, and Cookbook." *Handbook of International Economics*. Ed. by Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff. Vol. 4. Elsevier B.V. Chap. 3, pp. 131–195.

– (2019), "Brands in motion: How Frictions Shape Multinational Production." *American Economic Review* 109.9, pp. 3073–3124.

Hillberry, Russell and David Hummels (2013), "Trade Elasticity Parameters for a Computable General Equilibrium Model." *Handbook of Computable General Equilibrium Modeling*. Vol. 1. Elsevier, pp. 1213–1269.

Hofmann, Claudia, Alberto Osnago, and Michele Ruta (2017), "Horizontal Depth: A New Database on the Content of Preferential Trade Agreements." World Bank Policy Research Working Paper No. 7981.

Hummels, David (2001), "Toward a Geography of Trade Costs."

Imai, Kosuke and Teppei Yamamoto (2010), "Causal Inference with Differential Measurement Error: Nonparametric Identification and Sensitivity Analysis." *American Journal of Political Science* 54.2, pp. 543–560.

Imbs, Jean and Isabelle Mejean (2015), "Elasticity Optimism." *American Economic Journal: Macroeconomics* 7.3, pp. 43–83.

Jordà, Òscar (2005), "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95.1, pp. 161–182.

Kee, Hiau Looi, Alessandro Nicita, and Marcelo Olarreaga (2008), "Import Demand Elasticities and Trade Distortions." *Review of Economics and Statistics* 90.4, pp. 666–682.

Keller, Wolfgang and Stephen Ross Yeaple (2013), "The Gravity of Knowledge." *American Economic Review* 103.4, pp. 1414–1444.

Lashkaripour, Ahmad and Volodymyr Lugovskyy (2023), "Profits, Scale Economies, and the Gains from Trade and Industrial Policy." *American Economic Review* 113.10, pp. 2759–2808.

Ludema, Rodney D. and Anna Maria Mayda (2013), "Do Terms-of-Trade Effects Matter for Trade Agreements? Theory and Evidence from WTO Countries." *Quarterly Journal of Economics* 128.4, pp. 1837–1893.

Nicita, Alessandro, Marcelo Olarreaga, and Peri Silva (2018), "Cooperation in WTO's tariff waters?" *Journal of Political Economy* 126.3, pp. 1302–1338.

Ornelas, Emanuel (2016), "Special and Differential Treatment for Developing Countries." *Handbook of Commercial Policy*. Ed. by Kyle Bagwell and Robert W. Staiger. Vol. 1B. Elsevier B.V. Chap. 7, pp. 369–432.

Romalis, John (2007), "NAFTA's and CUSFTA's Impact on International Trade." *The Review of Economics and Statistics* 89.3, pp. 416–435.

Sandkamp, Alexander (2020), "The Trade Effects of Antidumping Duties: Evidence from the 2004 EU Enlargement." *Journal of International Economics* 123, p. 103307.

Schennach, Susanne (2022), "Measurement Systems." *Journal of Economic Literature* 60.4, pp. 1223–1263.

Sequeira, Sandra (2016), "Corruption, trade costs, and gains from tariff liberalization: Evidence from Southern Africa." *American Economic Review* 106.10, pp. 3029–3063.

Shapiro, Joseph S. (2021), "The Environmental Bias of Trade Policy." *Quarterly Journal of Economics* 136.2, pp. 831–886.

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

Smith, Adam (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*. McMaster University Archive for the History of Economic Thought.

Soderbery, Anson (2015), "Estimating import supply and demand elasticities: Analysis and implications." *Journal of International Economics* 96.1, pp. 1–17.

– (2018), "Trade Elasticities, Heterogeneity, and Optimal tariffs." *Journal of International Economics* 114, pp. 44–62.

UNCTAD (2011), "WITS User's Manual (Version 2.01)." January. URL: http://wits.worldbank.org/data/public/WITS%7B%5C_%7DUser%7B%5C_%7DManual.pdf.

World Bank (2019), *WITS Dataprovider*. URL: https://wits.worldbank.org/dataproviders.html (visited on 08/08/2019).

– (2018), *World Integrated Trade Solution (WITS)*. URL: https://wits.worldbank.org/ (visited on 10/29/2018).

A Survey of Papers Using WITS Tariff Data

In November 2023, I used Google Scholar to compile a list of all relevant papers using WITS tariff data. As researchers have many ways of referring to the WITS tariff data, it was not possible to perform a Google Scholar search that contained all possible combinations. Therefore, I began by searching for all published papers since 2010 in leading journals (*American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*) containing the word "tariffs" anywhere in the full text and downloaded the corresponding paper to build a corpus of papers that potentially use the WITS data. This query gives more than 600 distinct papers.[38]

Next, I searched within these papers for at least one of the following keywords: "WITS", "world integrated trade solution", "W.I.T.S.", "UNCTAD-Trains", "Trains", "T.R.A.I.N.S.", "Trade Analysis and Information System", and "UNCTAD". I then manually checked all the resulting papers one by one to remove false positives; for example, Shapiro (2021) uses WITS for nontariff barrier data. This process identified 21 papers published in leading journals using WITS data.

I repeated the process for publications in second-tier journals, including the *Review of Economics and Statistics*, *AEJ: Economic Policy*, *AEJ: Microeconomics*, *Macroeconomics*, *AEJ: Applied Economics*, *Journal of the European Economic Association*, and the *Journal of International Economics*. As expected, most papers using WITS data were from the *Journal of International Economics*, the leading field journal in international economics (76 papers). Additionally, 13 papers were published in various *AEJ* journals, 14 in the *Review of Economics and Statistics*, and one in the *Journal of the European Economic Association*. The full list of papers is available upon request.

Not all papers identified through this exercise necessarily rely on biased tariff data. The extent of potential errors depends on how the authors downloaded the tariffs (e.g., via WITS' download tool or bulk download option) and the cleaning steps taken to address missing tariff rates. Table A summarizes the data cleaning steps for papers published in top-five journals

[38]Initially, I tried following Disdier and Head (2008) by using EconLit to generate a corpus of full texts with "tariffs". However, this approach missed many relevant papers, so I switched to Google Scholar.

since 2010 and evaluates whether, and to what extent, the resulting tariff rates remain biased. This analysis is based on a review of the papers, their online appendices, and any replication files available on the journals' websites. For papers published in second-tier journals, this analysis was not feasible due to the large number of papers, exceeding 100 articles.

Table A1. Papers Using WITS Data and Published in Top-Five Journals since 2010

Paper	Tariffs	Aggr.	Ctries.	Years	Data Cleaning Steps	Remaining Issues
Trade Elasticity						
⚠ Arkolakis et al. (2018)	MFN & pref	pairs	62 ctries.	1999	Replace missing MFN tariffs at the ijk-level with the average MFN tariff over ik. Replace preferential tariffs with zero for all ijk if an RTA exists between ij.	Imputed preferential tariffs are mismeasured. Accurate preferential tariffs require a comprehensive list of RTAs. The average observed MFN tariffs contains selection bias, which is carried over to unreported observations.
✗ Brancaccio et al. (2020)	MFN & pref	pairs	world-wide	2010–2016	Focus analysis only on bulk commodities. Use the minimum of MFN and preferential tariff rates, where applicable. Compute weighted average tariffs to get pair-level tariffs.	Use tariffs that are affected by false interpolation and selection issues. Selection biases the pair average, over-representing ijk observations with high trade volumes and most likely lower tariffs.
⚠ Boehm et al. (2023)	MFN & pref	6-digit (HS)	world-wide	1995–2018	Drop specific tariffs from the analysis. Replace all intra-EU tariffs with zero. Exclude tariffs missing at the HS6 level due to selection, unless MFN equals zero.	Use tariffs that are affected by false interpolation and selection issues. Only within EU tariffs are corrected.
⚠ Caliendo and Parro (2015)	MFN & pref	20 sec. (based on WIOD)	16 ctries.	1993	Replace missing years by using reported tariffs from the following sequence of years (in order): 1992, 1991, 1990, 1989, 1994, 1995.	Download tariff data at the industry-level using the "effectively applied tariff," which is an average of HS6 products that are affected by false interpolation and selection issues. Biased sectoral averages from other years are carried over, containing both selection and measurement errors.
Product Quality						
✓ Feenstra and Romalis (2014)	MFN & pref	4-digit (SITC)	world-wide	1984–2011	Cleaning procedure conducted at the HS6-level, separately for preferential and MFN tariffs.	No major issues identified.
Effects of Bilateral Trade Policy						
✓ Conconi et al. (2018)	MFN & NAFTA-pref	6-digit (HS)	CAN, MEX, USA	1991, 2003	For Mexican NAFTA preferential tariffs, they use preferential tariffs from 2004. Exclude countries with RTAs in baseline analysis. In a robustness check, they include countries that signed an RTA by 2003 and for which Mexico reports preferential tariffs to WITS. Countries that might have an RTA but for which Mexico does not report preferential tariffs remain excluded.	No major issues identified. Potential selection bias arises from the exclusion of RTA members for which preferential tariffs are not reported.
ℹ Dix-Carneiro and Kovak (2017)	MFN	44 sec. (based on nat. class.)	BRA	1995, 2000, 2005, 2010	For the main specification, tariff changes come from national sources; for a robustness check, WITS Data are used. Calculate post-liberalization tariff changes for Brazil using MFN tariffs at the HS6-level. Brazil reports MFN tariffs annually. Preferential tariffs are not accounted for, and it is unclear whether this was intentional.	No major issues identified. Preferential tariffs are missing, but it is unclear whether only MFN tariffs were intentionally used.
★ Handley and Limão (2017)	MFN & Col. 2	6-digit (HS)	USA	2000, 2005	The complete tariff schedule for the US is downloaded; no interpolation is needed.	No major issues identified.

— *Continued from previous page* —

Paper	Tariffs	Aggr.	Ctries.	Years	Data Cleaning Steps	Remaining Issues
Multilateral Trade Policy						
Bagwell and Staiger (2011)	appl. & bound MFN	6-digit (HS)	16 recent WTO ctries.	1995, 2002	For products with tariff quotas, tariffs are replaced with the within-quota tariff binding.	No major issues identified. Applied and bound MFN tariffs may suffer from selection bias, as indicated by the particularly low number of observations.
Ludema and Mayda (2013)	appl. MFN	6-digit (HS)	36 ctries.	1995–2000	Applied MFN tariffs are downloaded from WITS. Tariffs are averaged over the period 1995–2000.	No major issues identified. Applied and bound MFN tariffs may suffer from selection bias, as indicated by the particularly low number of observations. *"The main advantage of the six-digit HS data set is its very fine level of disaggregation (more than 1,000 sectors per country with upward of 4,000 sectors for several of them) and the extensive country coverage."* (p. 1853).
Nicita et al. (2018)	appl. & bound MFN	6-digit (HS)	100 importer	2006	Applied MFN tariffs are downloaded from WITS, and bound MFN tariffs come from the WTO.	No major issues identified. Applied and bound MFN tariffs may suffer from selection bias, as indicated by the particularly low number of observations.
Firm-Level Studies						
Alfaro et al. (2016)	MFN	4-digit (SIC)	200 c's (imposed & faced)	2004	Replace missing MFN tariffs with data from 2003, 2002, 2005, and 2006 using TRAINS. For observations still missing, use data from the WTO. To calculate the share of trade covered by MFN tariffs, compute the fraction of imports sourced from countries without a trade agreement.	No major issues identified. Applied MFN tariffs may suffer from selection bias. The share of trade covered by MFN tariffs would be more accurate if preferential tariffs were included.
Boehm et al. (2022)	MFN & pref	3-digit (nat. class.)	IND	2000–2010	Replace remaining missing observations with MFN tariffs at the same level. Compute weighted tariffs to aggregate to industry-level.	Use tariffs that are affected by false interpolation and selection issues. Potentially authors exacerbate false interpolation, as MFN tariffs are applied to all observations without tariff information: For ijk observations that are not traded but could have a preferential tariff, the authors incorrectly assume that MFN applies.
Brandt et al. (2019)	MFN & pref	4-digit (nat. class.)	CHN	1994–2007	Aggregate tariffs at the 8-digit level of the HS classification to China's Industrial Classification (CIC) system using unweighted trade averages. Tariffs are described generically as "import tariff rates," without explicit details on the type of tariff.	No major issues identified: China has few preferential trade agreements and regularly reports preferential tariffs during the observation period. It is unclear whether the analysis uses tariffs affected by selection bias.
Bustos (2011)	MFN	4-digit (ISIC)	BRA, ARG	1991–1996	Uses initial MFN tariffs at 9-digit level. Compute import-weighted tariffs to aggregate to industry-level. Preferential tariff reductions are based on legal texts defining them.	No major issues identified. It is unclear whether the analysis uses tariffs affected by selection bias.
✗ Keller and Yeaple (2013)	MFN & pref	2-digit (SIC)	faced by US exporters	1994, 1999, 2004	No imputation methods are specified.	Use tariffs that are affected by false interpolation and selection issues.

— *Continued from previous page* —

Paper	Tariffs	Aggr.	Ctries.	Years	Data Cleaning Steps	Remaining Issues
Tariffs Used for Constructing Counterfactuals or Moments						
⚠ Bagwell et al. (2021)	MFN & pref	49 industries (SITC)	AUS, CAN, EUN, JPN, KOR, USA, 5 agg.	1990; 2000	If tariff data are unavailable for 1990 or 2000, it is borrowed from the closest available year. For European countries, they calculate Euro-zone common import tariffs and apply them product-wise to each country. For a given importing country (region) and product category, if the import tariff is missing for a specific partner, assume the MFN tariff is applied to that partner. They assume all trade is free if an RTA is in place (based on CEPII RTA data).	Imputed preferential tariffs are mismeasured. Accurate preferential tariffs require a comprehensive list of RTAs. The average observed MFN tariffs contains selection bias, which is carried over to unreported observations.
⚠ Burstein and Vogel (2017)	MFN & pref	2-digit (ISIC)	60 ctries.	2007	Use the first available tariff observation from the following sequence of years: 2007, 2006, 2008, 2005, 2009, 2004, and 2003 (at the two-digit manufacturing ISIC sector level). If there are no tariffs available between 2003–09, they use importer-exporter-sector triplets with observed tariffs to project these tariffs on an exporter-sector fixed effect and an importer-sector fixed effect. Predicted tariffs from this regression are then used to fill missing observations.	Download tariff data at the industry-level using the "effectively applied tariff," which is an average of HS6 products that are affected by false interpolation and selection issues. Biased sectoral averages from other years are used, carrying over mistakes across years.
⚠ Caliendo et al. (2021)	MFN & pref	pairs	EU15, new EU member states	2003	Use trade-weighted averages from WITS at the pair level. Replace tariffs for Cyprus and Hungary with their 2002 tariffs. Replace tariffs for Latvia with its 2001 tariffs.	Download tariff data at the country-level using the "effectively applied tariff," which is an average of HS6 products that are affected by false interpolation and selection issues. Biased sectoral averages from other years are used, carrying over mistakes across years.
ⓘ Head and Mayer (2019)	MFN & pref	Cars and parts[39]	worldwide	2000–2016	They fill missing data via linear interpolation. If data are missing for the most recent year, they use the last available year. Use preferential tariffs when available; otherwise, use MFN tariffs. They manually add the following tariff schedules for RTAs: EUN-KOR, USA-KOR, CAN-KOR, KOR-PER, KOR-TUR, KOR-AUS, KOR-NZL, KOR-VNT, KOR-CHN, JPN-PER, JPN-AUS, COL-USA, COL-CAN, COL-EUN, and others.	No major issues identified, assuming all missing preferential tariff schedules in WITS data were manually filled in.
⚠ Lashkaripour and Lugovskyy (2023)	MFN & pref	22 sec. (WIOD)	worldwide	2014	Download tariffs at ISIC3-level. Impute missing tariffs using the closest available data. Infer within-EU tariffs as zero and apply the same external tariffs for all EU partners. Replace remaining missing tariffs due to selection with the average tariff by reporter (importer)-sector across all partners. Replace any remaining missing tariffs (still due to selection) with the average tariff by sector across all reporters and partners.	Download tariff data at the industry-level using the "effectively applied tariff," which is an average of HS6 products that are affected by false interpolation and selection issues. Biased sectoral averages from other years are used, carrying over mistakes across years.

Note: The table contains all the papers featuring WITS data and published in the top-five economics journals (*American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*) since 2010. When MFN tariff rates are not further specified, the respective authors use the applied MFN tariffs. The description of the cleaning steps is based on a review of the original paper, and its online appendices, where available through the journal's website. Cntrs. refers to countries, Agg. is aggregation level. Legend of the icons: ✗ do not address data issues, relying on raw data as provided; ⚠ handle selection issues with flawed tariffs, carrying errors across years; ? might be biased by potential selection bias but unclear; ⓘ rely on legal texts for information about tariff cuts; ★ use accurate raw data; ✓ correctly fill in missing data at the HS6 level for missing tariffs, effectively addressing data issues.

[39] Included HS-codes: 8703, 8706, 8707, 8708, 840733, 840734, 840820, 840991, 840999.

Online Appendix for

Missing Tariffs

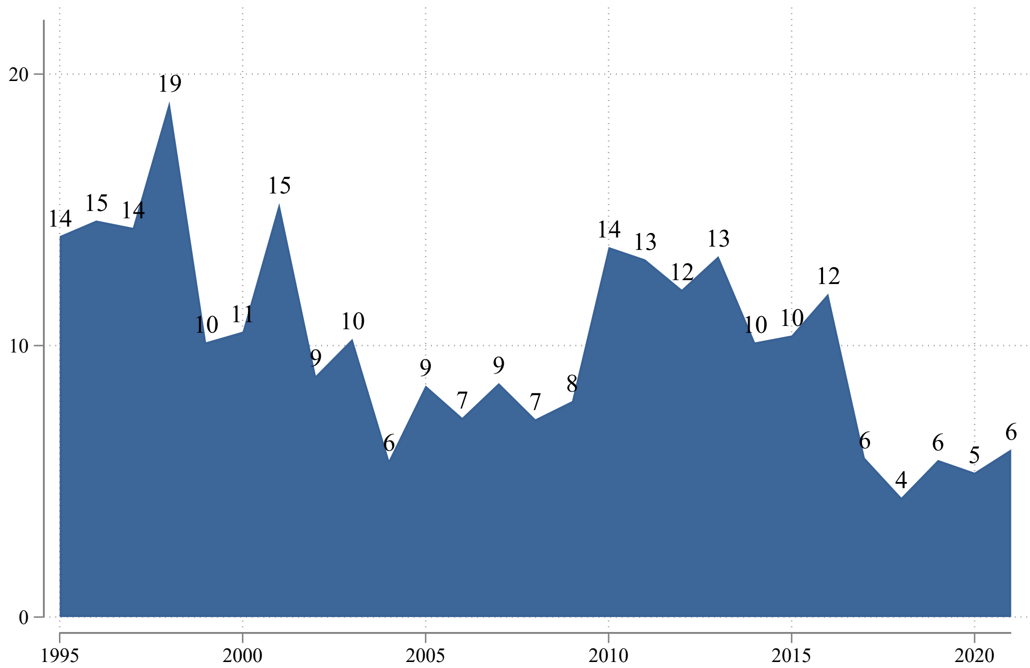
Feodora A. Teti

ifo Institute, Princeton University, University Munich, CEPR, CESifo

December 19, 2024

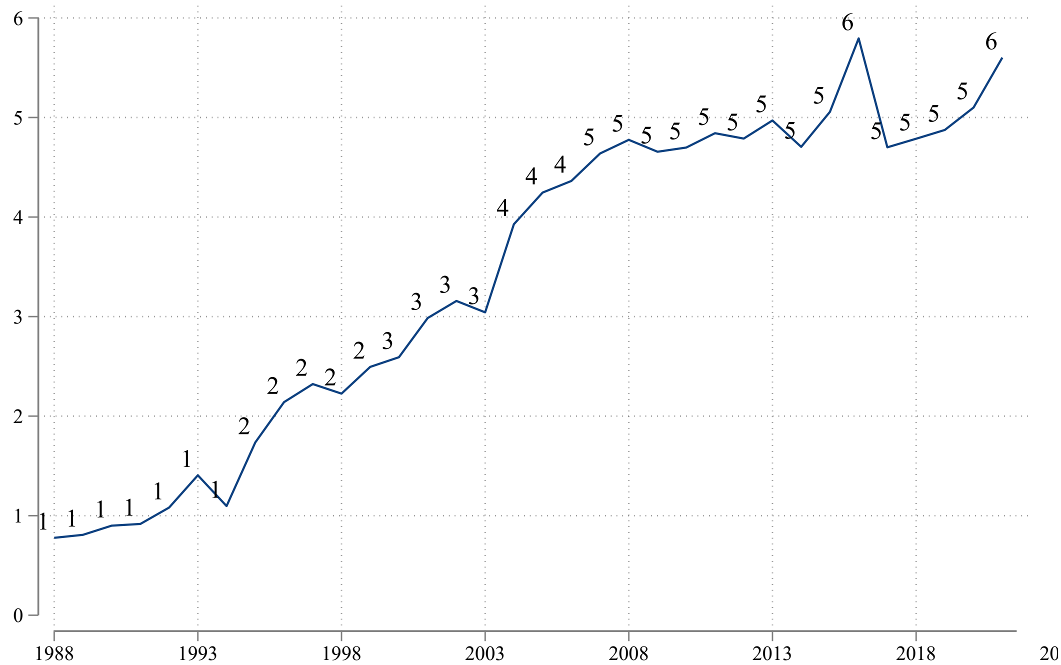
B Additional Material

Figure B.1. Share of Imports within RTAs with Falsely Interpolated Tariffs (EU included)

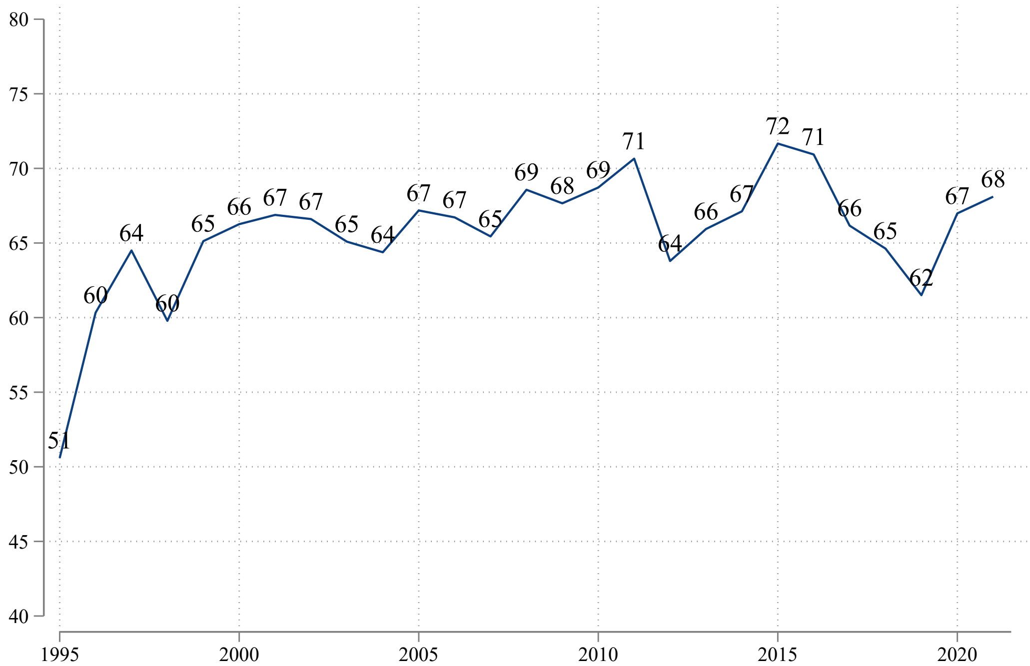


Note: The graph display the share of imports within RTAs impacted by falsely interpolated tariffs. Trade flows within the European Union are included. Trade flows for which MFN tariffs are zero are excluded. The share is smaller than in Figure 2 because tariffs within the EU are missing, hence, never falsely interpolated but inflate the denominator.

Figure B.2. The Extent of Selection Bias in WITS over Time



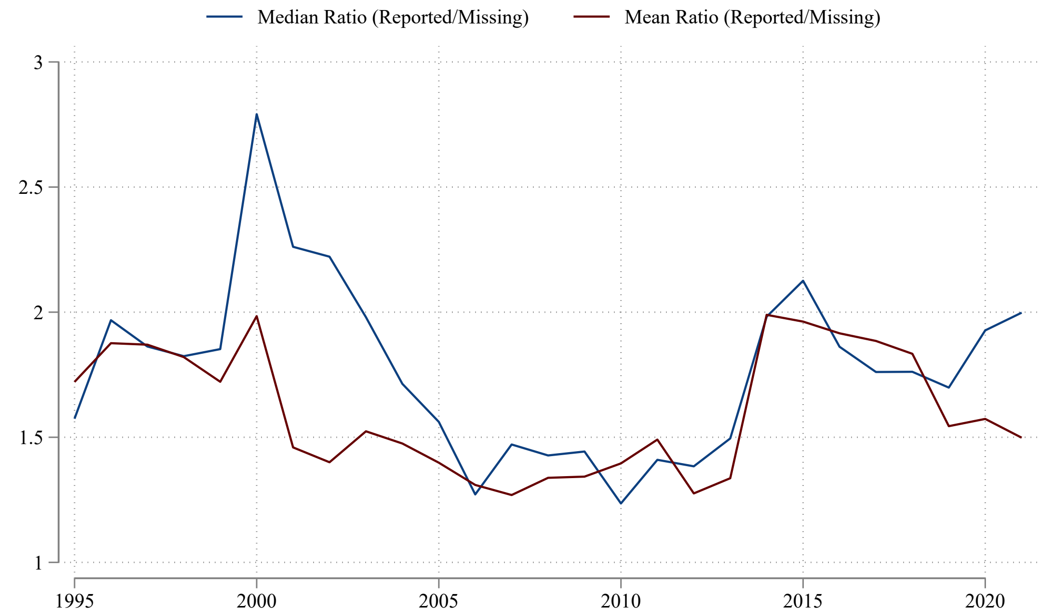
(a) Share of Reported Observations in WITS vs. new GTD



(b) Share of Worldwide Imports Reported in WITS

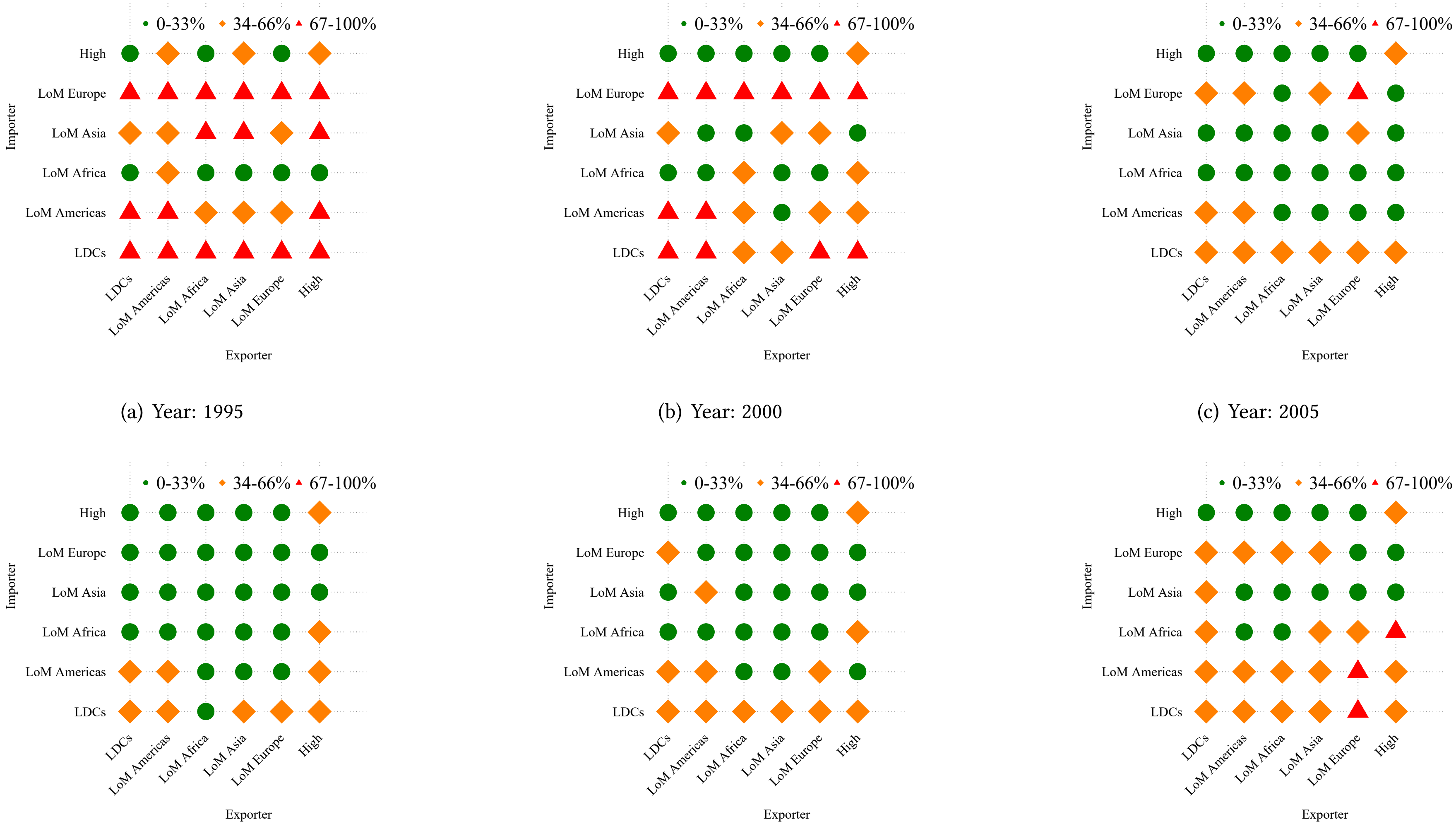
Note: Panel (a) shows the share of observations reported in the WITS database relative to the total number of observations available in the new Global Trade Database (GTD) from 1988 to 2021. Panel (b) reports the share of global imports, based on BACI trade data, for which WITS provides any tariff information, including both accurate tariffs and falsely interpolated tariff rates, between 1995 and 2021.

FIGURE B.3. RATIO OF REPORTED TO MISSING TRADE FLOWS, 1995–2021



Note: The graph shows the ratio of median and mean trade flows between reported and missing observations from 1995 to 2021. The ratios remain consistently above 1, indicating that reported trade flows are substantially larger than the missing ones across the entire period.

Figure B.4. Share of Total Trade with Missing or Falsely Interpolated Tariffs by Income Groups over Time



(a) Year: 1995

(b) Year: 2000

(c) Year: 2005

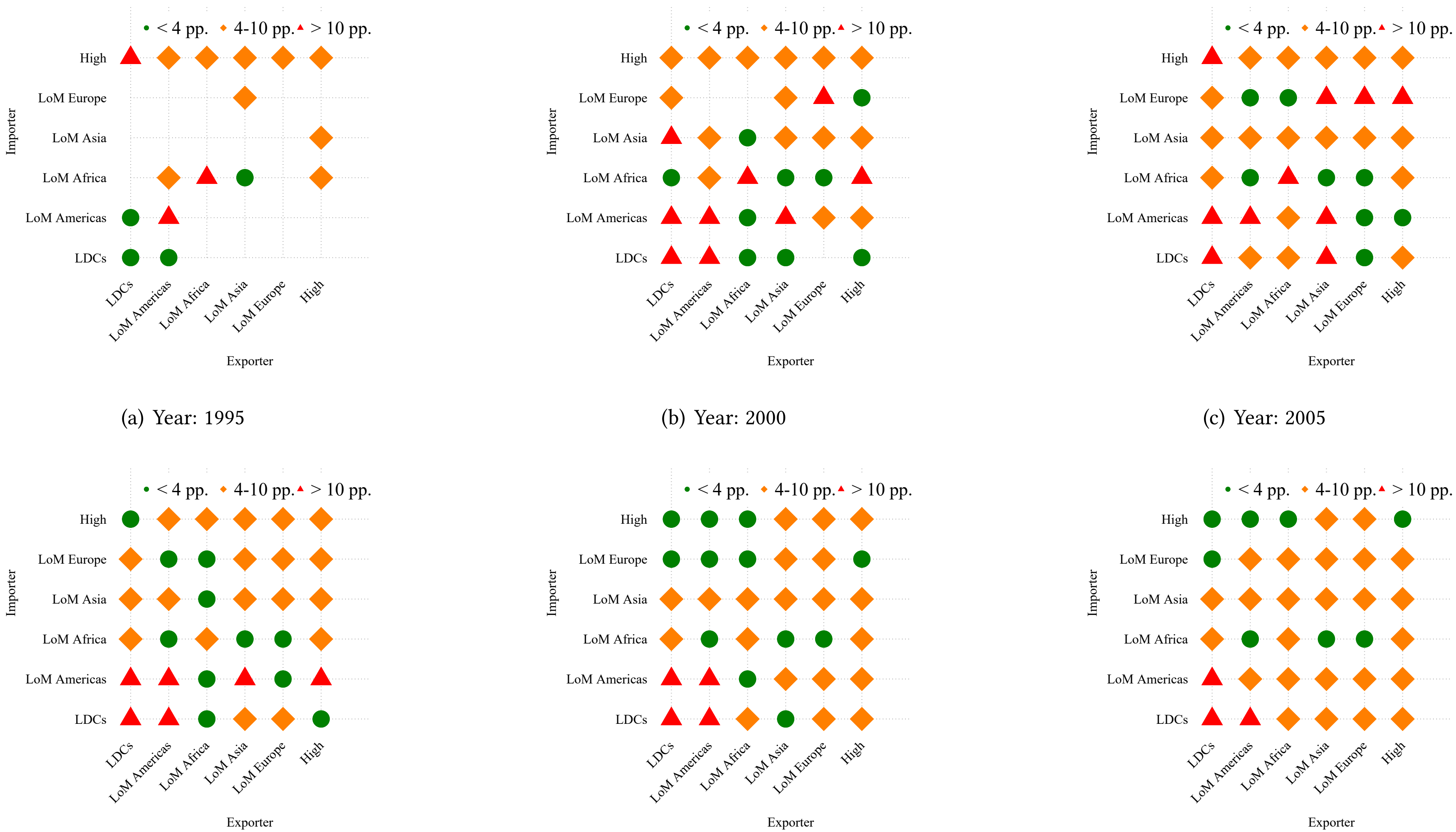
(d) Year: 2010

(e) Year: 2015

(f) Year: 2020

Note: The graph shows the share of worldwide imports for which tariffs are corrupted (falsely interpolated or missing tariffs) , disaggregated by income groups over time. Income groups follow the World Bank classification: least developed countries (LDCs), regional low- or middle-income (LoM) countries, and high-income countries.

FIGURE B.5. SHARE OF AVERAGE INTERPOLATION ERROR BY INCOME GROUPS OVER TIME



(a) Year: 1995

(b) Year: 2000

(c) Year: 2005

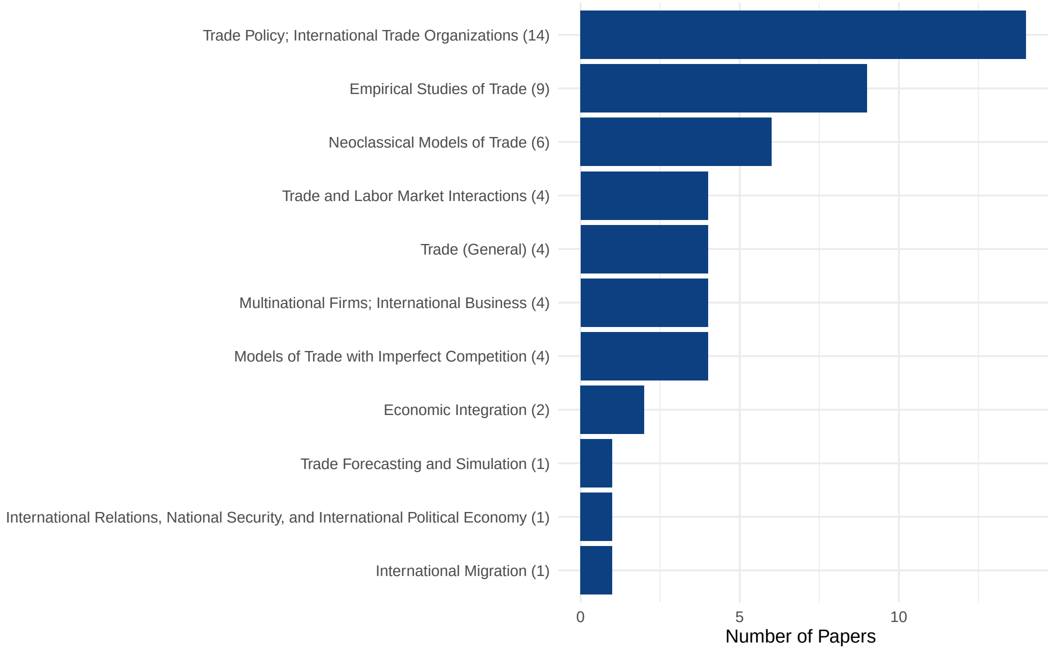
(d) Year: 2010

(e) Year: 2015

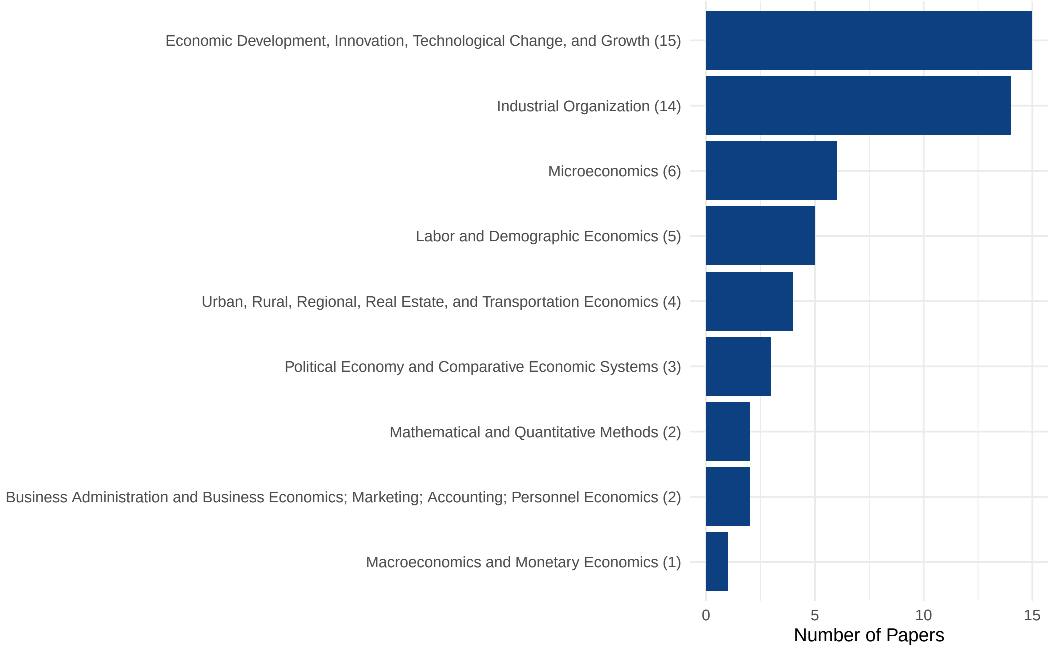
(f) Year: 2020

Note: The graph shows the average interpolation error (in pp.), disaggregated by income groups over time. Income groups follow the World Bank classification: least developed countries (LDCs), regional low- or middle-income (LoM) countries, and high-income countries.

FIGURE B.6. ANALYSIS OF JEL CODES OF PAPERS USING WITS DATA AND PUBLISHED IN TOP-FIVE JOURNALS



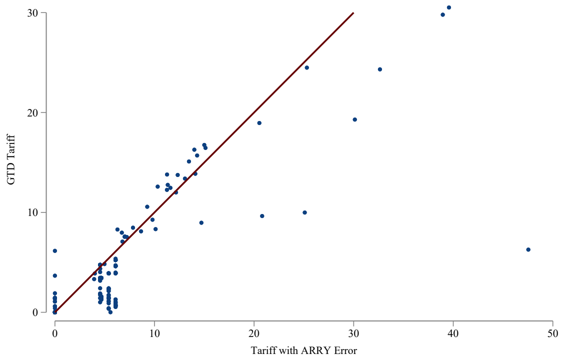
(a) Within International Economics



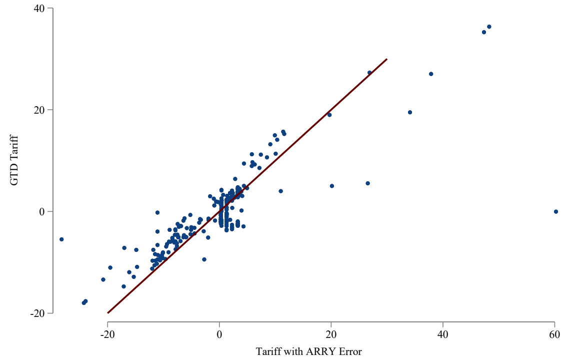
(b) Other Fields

Note: The graphs show the distribution of JEL codes of the 21 papers that use WITS data; the total number of papers with the respective JEL code is in parentheses.

FIGURE B.7. MEASUREMENT ERROR IN ARRY



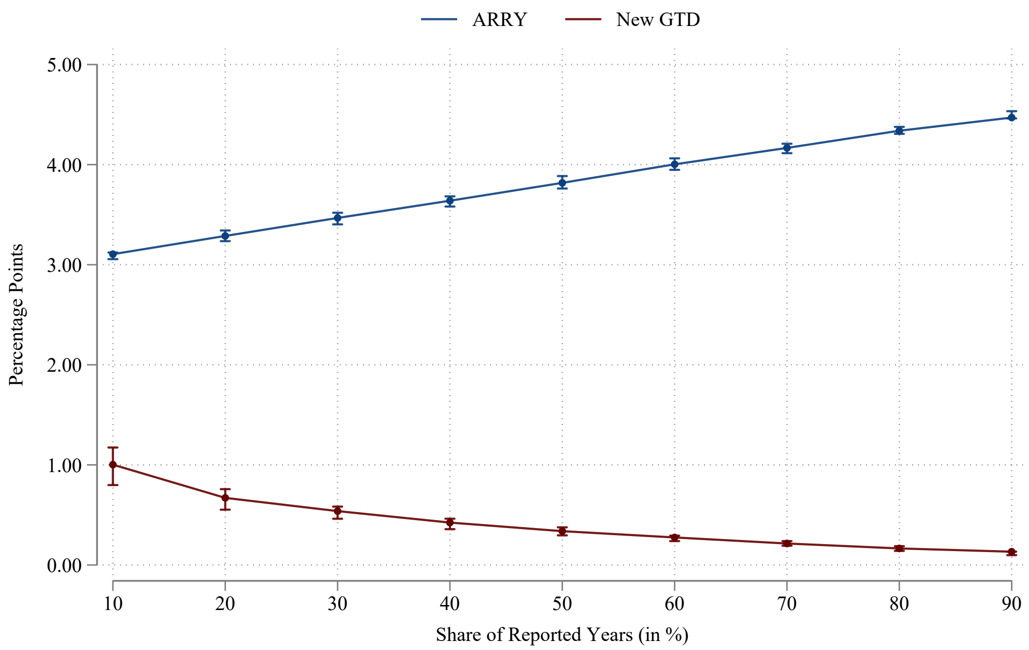
(a) Tariffs (in %)



(b) De-meaned Tariffs ($\tau - \dot{\tau}$)

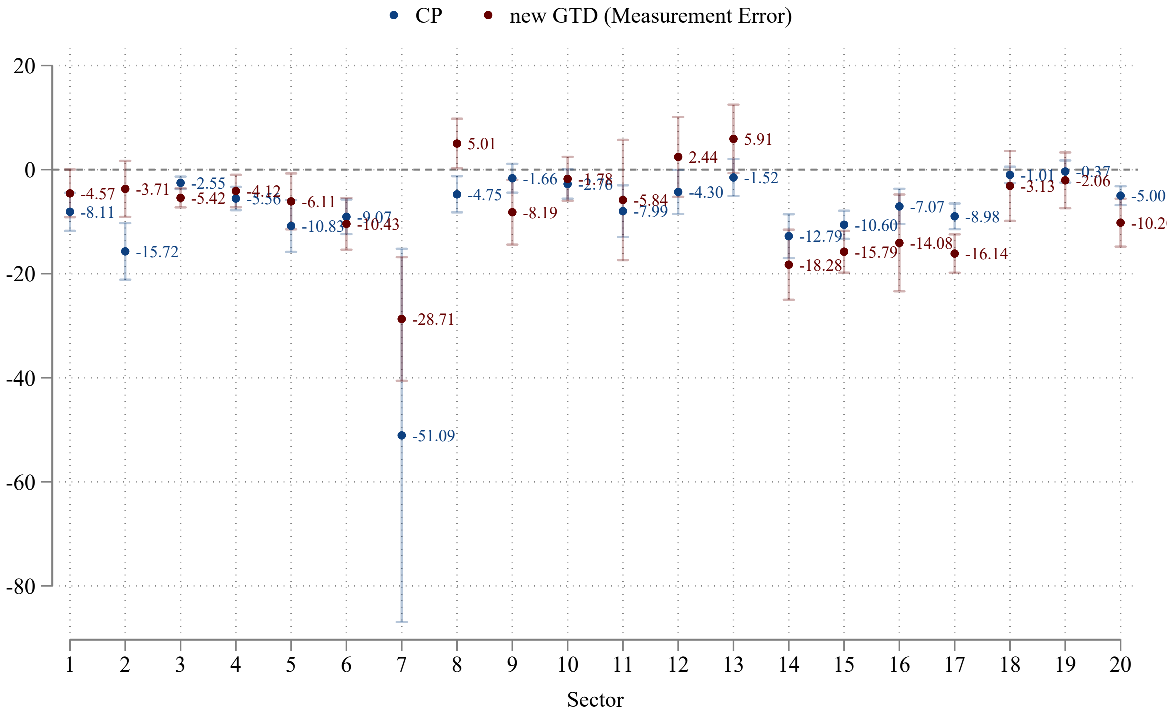
Note: The graph compares the simple average of pair-level tariffs calculated from the GTD tariff rates and from the ones used in ARRY. Panel (a) shows tariff rates in %; Panel (b) gives the de-meaned tariffs $\tau_{ij} - \dot{\tau}_{ij}$ with $\dot{\tau}_{ij} = \tau_{ij} - \bar{\tau}_i - \bar{\tau}_j + \bar{\tau}_{ij}$). The figure clearly shows that the error aligns more closely with classical measurement error, as the corrupted data cluster around the 45-degree line rather than falling below it. The importer and exporter fixed effects amplify this issue by re-centering the measurement error.

FIGURE B.8. MEAN ABSOLUTE ERROR: NEW GTD VS. ARRY-INTERPOLATION METHOD



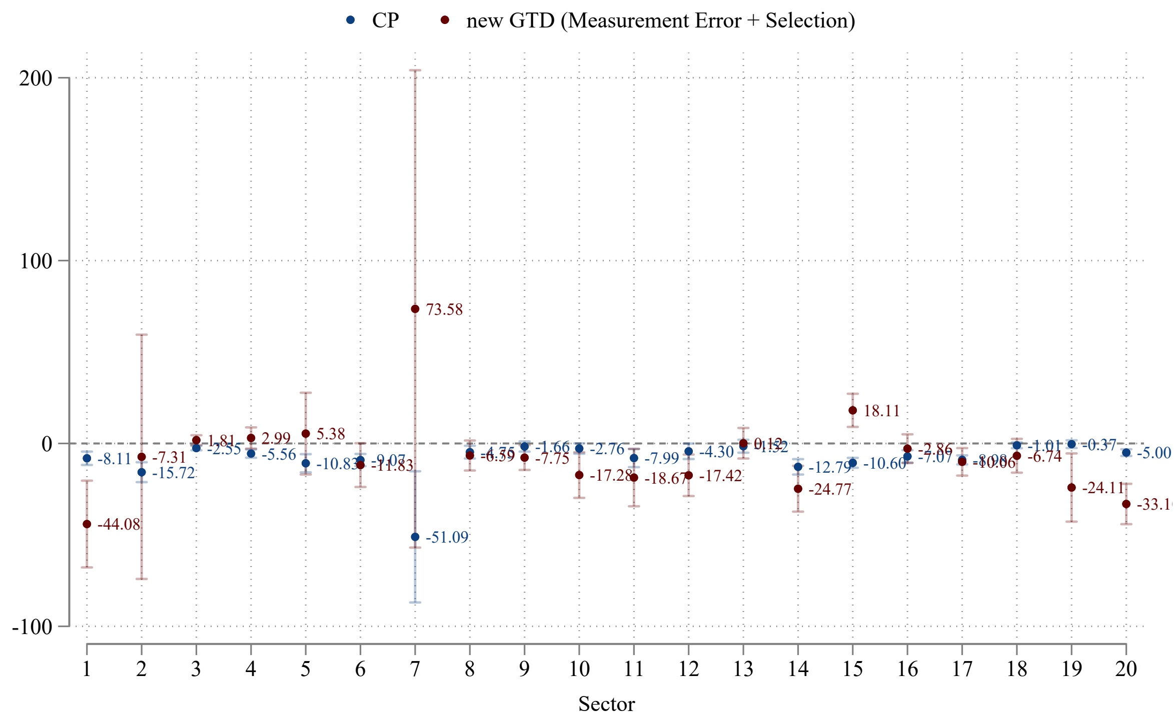
Note: The mean absolute error (MAE) and the 25th and 75th percentiles are reported in percentage points. The graph compares ARRY's interpolation method, which sets tariffs for all products to zero when a trade agreement starts, with the GTD. See the main text for details on the data used in the simulation exercise.

Figure B.9. Elasticities by WIOD Sector: Caliendo and Parro (2015): Correcting for False Interpolation



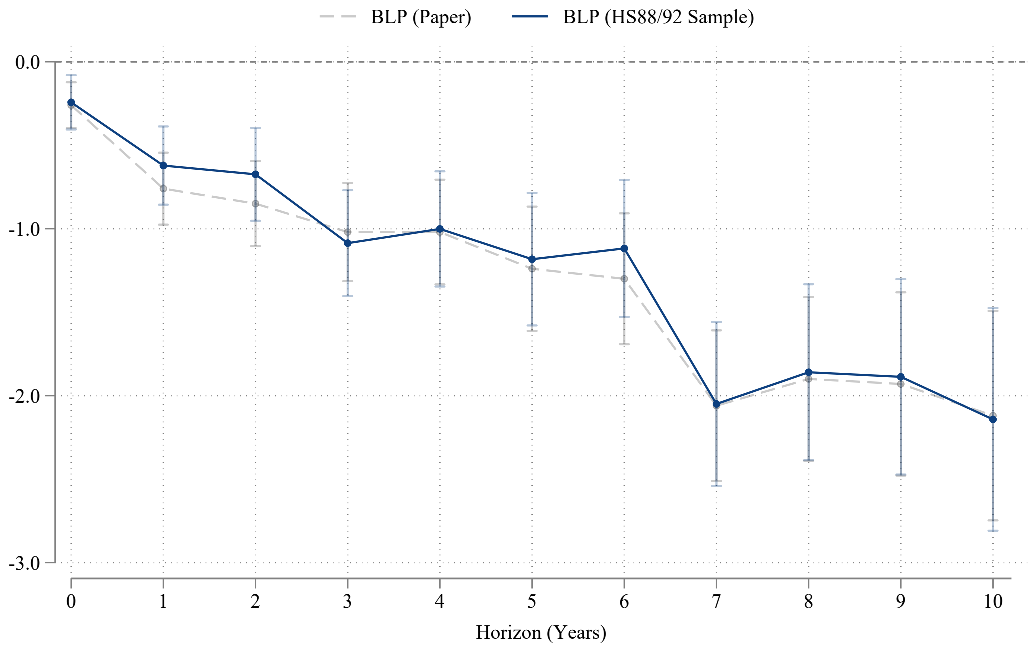
Note: The graph shows the results of estimating Equation (5) separately for all WIOD sectors. The replicated coefficients from CP are displayed in blue and give the estimates when I use the tariff data that contains both the false interpolation and positive selection ($-\tilde{\sigma}_s$). The results when I use the GTD keeping only observations that are also included in CP and hence correct only for false interpolation are shown in red ($-\sigma_s$).

Figure B.10. Elasticities by WIOD Sector: Caliendo and Parro (2015): Correcting for False Interpolation and Selection



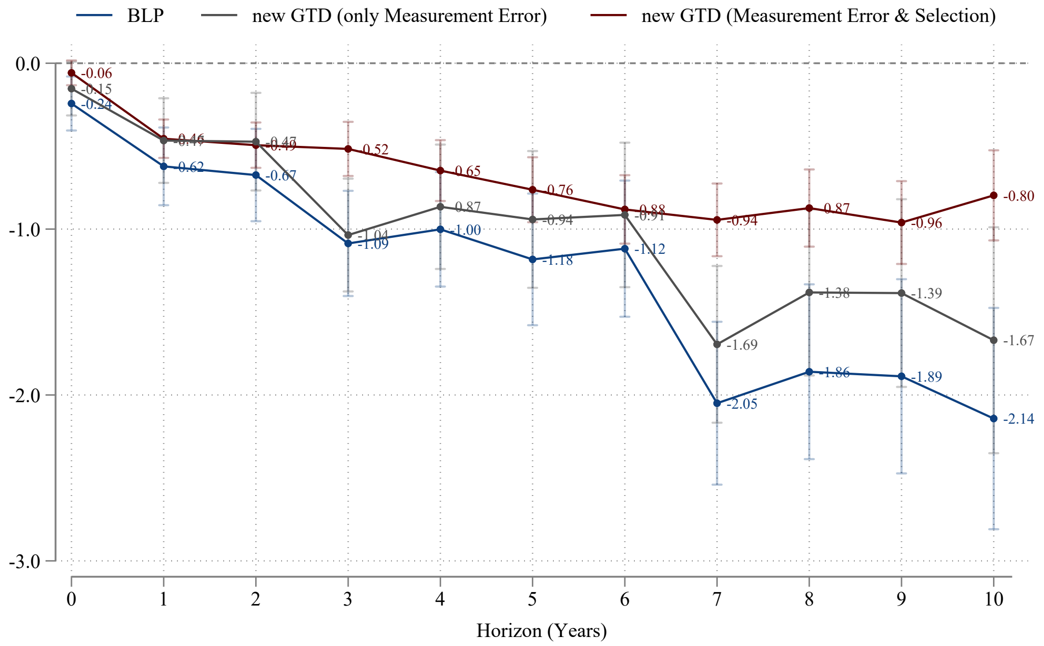
Note: The graph shows the results of estimating Equation (5) separately for all WIOD sectors. The replicated coefficients from CP are displayed in blue and give the estimates when I use the tariff data that contains both the false interpolation and positive selection ($-\tilde{\sigma}_s$). The results when I use the GTD and hence correct for both issues present in the WITS data are shown in red ($-\sigma$).

Figure B.11. Replication of Boehm et al. (2023) with Smaller Sample



Note: The gray line shows the coefficients that result when I replicate BLP's analysis using their provided data and code, which are identical to the results reported in their paper. The dark blue line reports the coefficients when I rerun the same analysis on the same data but trim the sample by keeping only HS6 products in the 1988/92 nomenclature. The differences are minimal. All regressions include importer–HS4–year, exporter–HS4–year, importer–exporter–HS4 fixed effects, and one lag of the changes in tariffs and trade as pretrend controls.

Figure B.12. Reestimating Boehm et al. (2023): Different Horizons



Note: The graph compares the tariff elasticity over different horizons when I calculate it with the WITS data and the GTD data. The blue line shows the results using the WITS data ($-\tilde{\sigma}_s$). The gray line shows the estimates accounting only for false interpolation ($-\sigma_s$), while the red line shows the estimates when also correcting the data for the positive selection in the WITS data ($-\sigma$). All regressions include importer–HS4–year, exporter–HS4–year, importer–exporter–HS4 fixed effects, and one lag of the changes in tariffs and trade as pretrend controls.

Table B.1. First Stage in Boehm et al. (2023)

	(1)	(2)	(3)	(4)	(5)	(6)
		$h=1$			$h=10$	
z_t	0.82***	0.74***	0.75***	0.68***	0.64***	0.56***
	(0.005)	(0.003)	(0.002)	(0.008)	(0.007)	(0.003)
N (in Mio.)	21	21	52	7	7	20
	BLP	GTD-BLP sample	GTD	BLP	GTD-BLP sample	GTD

Note: The table shows the first stage using the BLP data (Columns (1) and (4)), when correcting for false interpolation but keeping the sample constant (Columns (2) and (5)), and when correcting for false interpolation and sample selection (Columns (3) and (6)) at horizon $h=1$ and $h=10$. All regressions include importer–HS4–year, exporter–HS4–year, importer–exporter–HS4 fixed effects, and one lag of the changes in tariffs and trade as pretrend controls.

C Interpolation Algorithm

This section outlines the interpolation algorithm, which is applied separately to MFN and preferential tariffs. The algorithm infers missing tariff rates from reported values while accounting for key trade policy features, such as gradual tariff phase-outs and the dynamics of multiple, sequentially deepening agreements. Appendices D and E describe the data sources and preprocessing steps for both RTA and raw tariff data.

C.1 Interpolating MFN Tariffs

For MFN tariffs, I construct a matrix containing all importer–HS6 product–year combinations, as these are the dimensions along which MFN tariffs vary. I then merge the reported MFN tariffs and interpolate the missing observations as follows: Rather than using linear interpolation between available observations, each missing tariff is set equal to the nearest preceding observation. If no preceding observation exists, the missing MFN tariff is set equal to the nearest subsequent observation. Figure 4 in the main text illustrates this procedure. This approach accounts for anecdotal evidence suggesting that countries are more likely to update schedules following significant tariff changes and aligns with the methodology in Caliendo et al. (2023).

Hence, the final dataset containing fully interpolated MFN tariffs is a balanced panel uniquely identified by importer, HS6 product, and year. For importing countries that report tariffs for all 5,018 products at least once across all years, this results in a total of 170,612 observations (5,018 products multiplied by 34 years). For some countries, such as certain non-WTO members or least developed countries, the number of products for which MFN tariff rates are reported at least once is slightly lower than this.

C.2 Interpolating Preferential Tariffs

I now outline the interpolation algorithm for preferential tariffs. The interpolation process relies on complete observations of MFN tariffs for every importer–product–year observation, as these are used to interpolate preferential tariffs that are being phased-out.

Extend Time-Dimension The reported preferential tariffs are merged into a matrix containing all importer–exporter–HS6 product–year combinations, covering 200 importers, 199 exporters, and 5,000 products over the years 1988–2021. Fully interpolated MFN tariffs are also added to the matrix, as they are essential for subsequent interpolation of preferential tariffs. Additionally, the dataset includes information on RTAs, i.e., the agreement name, year of entry into force, year of withdrawal, year of full implementation, a dummy variable indicating whether phase-outs is allowed, and a dummy variable specifying whether the country pair has only a single agreement or multiple agreements between the entire period, i.e., 1988–2021.

To make the process computationally easier, I keep only those ijk observations for which, at least once over all years, a preferential tariff strictly smaller than the MFN tariff, i.e., $t_{ijkt} < t_{ikt}$, is reported. The resulting dataset is then split by products into batches of 500 products to further simplify the computation.

Then the country pair–year observations are split into two categories: single agreement (with and without phase-outs) vs. sequentially deepening agreements (with and without phase-outs).

Single RTA For country pairs with only a single trade agreement, the interpolation of preferential tariffs is relatively straightforward and follows the procedure illustrated in Figure C.1.

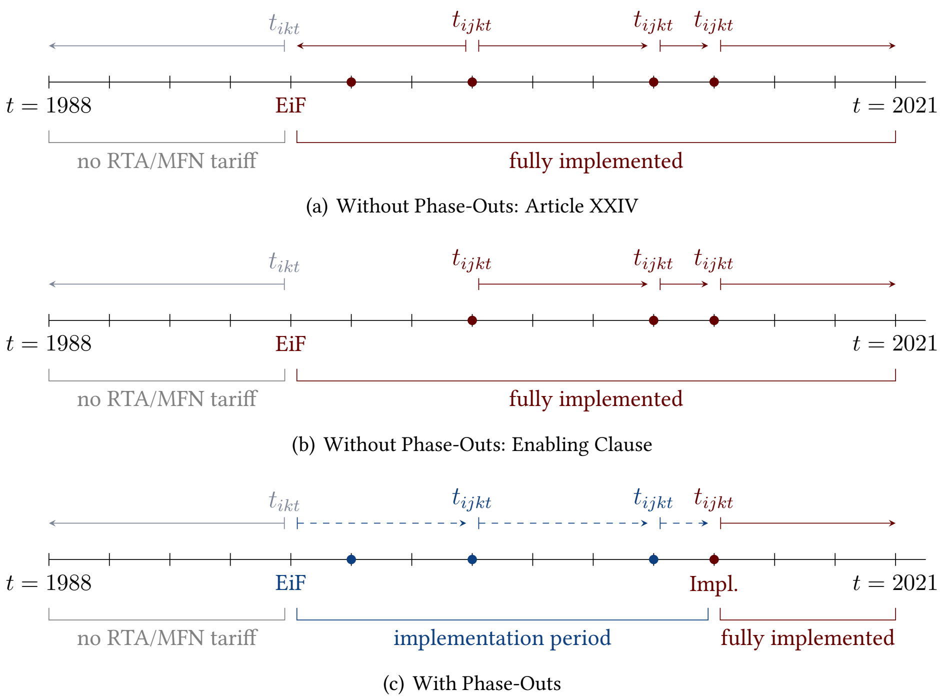


FIGURE C.1. IMPUTATION ALGORITHM FOR PAIRS WITH A SINGLE RTA

Note: The solid line represents missing observations filled by forward and backward filling, while the dashed line represents linear interpolation. "EiF" stands for entry into force, "Impl." for year or full implementation.

For agreements without phase-outs, the algorithm first fills in the reported preferential tariffs both backward and forward for all years following the year of entry into force, as the entry into force and full implementation years coincide. This approach is applied to Article XXIV agreements under the assumption that a preferential tariff, once reported, is in effect starting from the agreement's entry into force year onward (Figure C.1, Panel (a)).

Many nonreciprocal agreements, which constitute the majority of enabling clause agreements, undergo regular extensions of their scope. For instance, the European Union expanded its program for least developed countries in 2001 by introducing the "Everything but Arms" (EBA) initiative, which grants preferential access for all products except arms. This program operates alongside other nonreciprocal agreements, such as the Generalized System of Preferences, which covers a broader range of countries but with less extensive product scope. However, the RTA data do not allow differentiation between various types of nonreciprocal agreements; only the presence of a nonreciprocal agreement is observable. Consequently, sequential deepening within enabling clause agreements cannot be directly tracked. To address this limitation, the algorithm only performs forward interpolation for agreements notified under the enabling clause, refraining from backward interpolation. This ensures that the algorithm effectively captures potential sequential deepening within enabling clause agreements while avoiding the false assignment of low preferential tariffs to years when they were not yet in place. This process is illustrated in Panel (b).

The procedure described above fills in the red segments shown in Figure C.1, corresponding to years in which agreements are fully implemented. For agreements without phase-out periods, this completes the interpolation. However, for agreements with phase-out periods, the algorithm interpolates linearly for years within the implementation period (blue segment). If no preferential tariff rates are reported during this period, the algorithm would determine the missing tariff rates by linearly interpolating between the MFN tariff rate t_{ikt} applicable in $t = eif - 1$ (the year prior to the agreement's entry into force) and the preferential tariff rate t_{ijkt} observed in $t = impl$ (the year of full implementation) (Panel (c)). The algorithm's accuracy naturally improves with more comprehensive reporting. Performance is significantly enhanced when countries report tariff data for at least one year soon after the entry into force, as this enables the algorithm to more accurately determine which products (k) are immediately reduced to zero and which are subject to longer phase-out periods.

Sequential Deepening of RTAs The sequential deepening of RTAs—i.e., trade agreements becoming progressively deeper over time and further liberalizing trade—is a well-documented phenomenon, as demonstrated by Dür et al. (2014) and Hofmann et al. (2017), who construct RTA data measuring the depth and content of trade agreements.

With the sequential deepening of RTAs, it is essential to assign each reported preferential tariff for product k to its corresponding trade agreement. Note that the correct corresponding trade agreement might change over time. The algorithm achieves this by identifying the

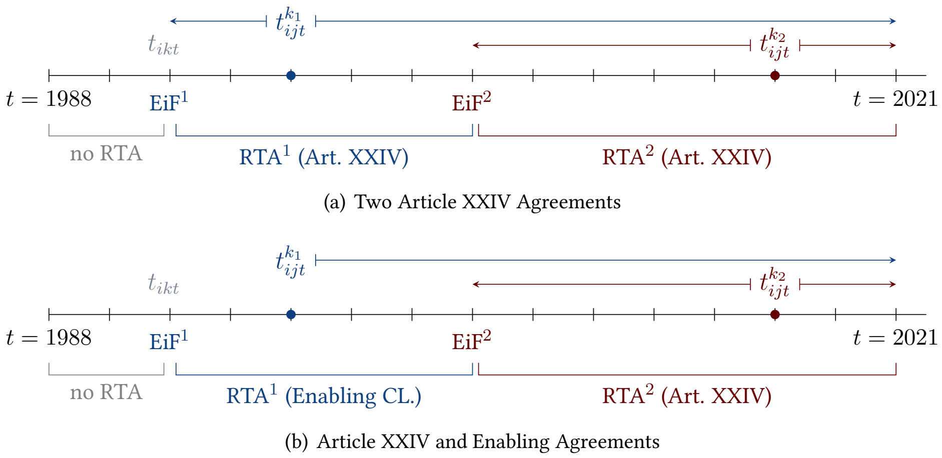


FIGURE C.2. IMPUTATION ALGORITHM FOR PAIRS WITH SEQUENTIAL RTAS AND NO PHASE-OUTS
Note: Blue represents the first RTA, which may be notified under either Article XXIV (Panel (a)) or the Enabling Clause (Panel (b)), while the second RTA is shown in red. "EiF" stands for entry into force.

first year in which a preferential tariff for product k is observed. In that year, the algorithm determines which RTA is in force and assigns its entry into force year to the preferential tariff rate for product k. This assignment method allows the entry into force years to vary by product k.

For example, Panel (a) of Figure C.2 illustrates a case where country i reports a preferential tariff for product k_1 while the first RTA (RTA^1) is in effect. Accordingly, the algorithm assigns the entry into force year for the first RTA to product k_1. In contrast, for product k_2, country i only reports a preferential tariff for the first time when the second RTA (RTA^2) is already in effect. Typically, in that year, country i would also report the preferential tariff rate for product k_1, but this is omitted from the figure for simplicity.

Thus, for product k_1, the algorithm assumes that the preferential tariff can be interpolated over the entire period from $t =$ EIF^1 to the last available year (or, if applicable, until no RTA is in place for country pair ij). However, for product k_2, the algorithm assumes that the preferential tariff applies only from $t =$ EIF^2 onward, without extending back to $t =$ EIF^1. It is then interpolated for the entire period from $t =$ EIF^2 to the last available year (or, if applicable, until no RTA is in place for country pair ij). As in the case of single RTAs, for agreements notified under the enabling clause, I carry reported tariff rates forward but do not interpolate backward.

For country pairs with sequential deepening RTAs that are gradually phased in, it is essential to assign each product-level preferential tariff rate both an entry into force year and an implementation year corresponding to the relevant RTA. For instance, if country i reports a preferential tariff rate for product k^1 during one of the years when RTA^1 is in force for the first time, the algorithm assigns the entry into force year of RTA^1 and the corresponding year of implementation to product k^1. The missing years are then interpolated in the same manner as if it were a single RTA.

This approach reflects the idea that countries generally adhere to pre-existing liberalization plans when deepening trade policy relations. The trade relationship between the United States and Canada offers a compelling example of the rationale behind this assumption. Most tariff eliminations in this relationship were governed by the Canada–United States Free Trade Agreement (CUFTA), with NAFTA subsequently inheriting CUFTA's tariff schedules, including the final year of implementation.

D RTA Data

To interpolate the non-reported preferential tariffs, ideally, I would like to map every reported preferential tariff to its corresponding regional trade agreement (RTA) and extract, for each RTA, information on the year of entry into force, whether tariffs are phased out, and, if so, the year of implementation. However, such detailed information is not readily available in a single source. Therefore, I construct a new dataset by combining information from five distinct sources.

For bilateral trade agreements, I primarily rely on DESTA (Dür et al. 2014), which provides a comprehensive list of bilateral agreements and tracks all accessions and withdrawals. However, as DESTA does not distinguish between active and inactive agreements, I supplement this information with the WTO RTA Database. Additionally, I cross-check these sources using Mario Larch's RTA database (Egger and Larch 2008), which, while useful for validation, contains only a dummy variable indicating the existence of an RTA without differentiating between agreements. Conflicts arising from these sources are reviewed and resolved on a case-by-case basis.

DESTA, the WTO RTA Database, and Mario Larch's database do not include nonreciprocal trade arrangements, such as the African Growth and Opportunity Act (AGOA) or the General System of Preferences (GSP), which offer unilateral tariff preferences to low-income countries. These arrangements were particularly prevalent in the early 1990s. For example, in 1995, 22% of all country pairs were receiving preferences through a nonreciprocal trade arrangement. To capture these agreements, I use the NSF's-Kellog Institute Data Base on Economic Integration Agreements (EIAs) compiled by Scott Baier and Jeffrey Bergstrand (Baier et al. 2014), supplemented by the WTO's database on preferential trade agreements. This allows me to construct a comprehensive dataset of nonreciprocal trade arrangements.

Addressing the sequential deepening of RTAs requires distinguishing between RTAs for specific country pairs and tracking their evolution over time. The goal is to construct a coherent list of trade agreements that reflects changes in agreement names, especially when these indicate a potential shift in the scope of products covered by tariff reductions. For instance, in the trade relationship between the United States and Mexico, the dataset should include the GSP until 1993, transitioning to NAFTA from 1994 onward to capture the significant expansion in product coverage. Conversely, the USMCA, which entered into force in 2020, should not be coded, as it primarily introduced nontariff changes, such as stricter rules of origin, without affecting preferential tariffs. Achieving this requires information from all five databases, as no single source provides all specific type of information needed.

D.1 Sources

I combine data from DESTA (Dür et al. 2014), the WTO RTA database, Mario Larch's RTA database (Egger and Larch 2008), Baier and Bergstrand's NSF-Kellogg Institute Data Base on Economic Integration Agreements (Baier et al. 2014), and the WTO's database on preferential trade agreements. Below, I describe the download procedure for each database.

The primary data bases used in this paper are:

DESTA (Design of Trade Agreements Database) The data from the Design of Trade Agreements (DESTA) database are downloaded from https://www.designoftradeagreements.org/downloads/. I use Version 2.2, released in 2023, and the files were retrieved on April 16, 2024. Specifically, the following components are downloaded:

- *List of Treaties* (Version 2.2),
- *DESTA Database* (Version 2.2), and
- *Indices* (Version 2.2).

WTO RTA Database The data from the WTO Regional Trade Agreements (RTA) database are obtained from https://rtais.wto.org/UI/PublicMaintainRTAHome.aspx. The data were downloaded on July 26, 2024. The following steps are used to retrieve the data:

1. Select the option "Export all RTAs" from the website interface.
2. Save the resulting file as an Excel table.

Mario Larch's RTA Database The data from Mario Larch's Regional Trade Agreements (RTA) database are obtained from https://www.ewf.uni-bayreuth.de/en/research/RTA-data/index.html. The version used is current as of July 12, 2024, and the data were downloaded on July 30, 2024.

NSF-Kellogg Institute Database on Economic Integration Agreements The data from the NSF-Kellogg Institute Database on Economic Integration Agreements are obtained from https://kellogg.nd.edu/nsf-kellogg-institute-data-base-economic-integration-agreements. The version used is July 2021, and the data were downloaded on February 16, 2024. Nonreciprocal trade arrangements are identified by filtering the database to include only observations where $\text{EIA} = 1$.

WTO Preferential Trade Agreements (PTA) Database The data from the WTO Preferential Trade Agreements (PTA) database are retrieved from http://ptadb.wto.org/ptaList.aspx. The procedure involves the following steps:

1. Sort the list of PTAs by clicking on "Initial Entry Into Force" to order them by entry date (EiF).
2. Download the resulting Excel file of the list and save it as `WTO-PTAlist_2024_08_15.xlsx`.
3. Cross-check the data with Baier and Bergstrand's database to identify missing Generalized System of Preferences (GSP) granting countries.
4. For PTAs missing in Baier and Bergstrand's database:
   (a) Go to http://ptadb.wto.org/ptaList.aspx.
   (b) Click on the respective PTA name (e.g., "Duty-free treatment for African LDCs – Morocco").
   (c) Select the "Beneficiaries" option and download the list of beneficiaries by clicking on the Excel icon.
   (d) Save the data for the missing GSP granter in the folder: `GSPgranting/BBmissing-GSPgranting_2024_08_15.xlsx`.

The resulting file compactly combines all missing information and is used for coding purposes.

D.2 Combining Five Different RTA Datasets

DESTA–WTO RTA Database The DESTA dataset provides the largest coverage of bilateral trade agreements. Its main advantage is that it includes agreements beyond those officially notified to the WTO, making it the most comprehensive source for bilateral trade agreements. Additionally, DESTA comprehensively codes all accessions and withdrawals of countries. However, DESTA only contains a list of agreements without indicating whether they are still in force. This omission is significant, as many agreements have been superseded or replaced over time.

To address this shortcoming, I merge the DESTA data with the WTO RTA Database. Unlike DESTA, the WTO RTA Database does not provide bilateral trade agreements but instead includes trade agreements identified by their names. To match trade agreements between DESTA and the WTO RTA Database, I perform fuzzy string matching. Specifically, I find the closest match for all trade agreements that share the same year of entry into force in both datasets.

The matching process is iterative: when exact matches are unavailable, I relax the matching criteria to include agreements with entry-into-force dates that differ by up to two years, following the sequence $eif + 1, eif - 1, eif + 2, eif - 2$. I also consider the ratification year as a fallback matching criterion. In cases where the WTO data provide two separate dates for goods and services agreements, I retain only the entry-into-force date for goods. This choice reflects the focus of my analysis on tariff data, where goods-specific dates are the most relevant.

Matching the two databases reveals several inconsistencies in DESTA, which I correct manually on a case-by-case basis (details are available upon request). For example, three countries—Chinese Taipei, Kosovo, and Palestine—are not included in DESTA. Some agreements, such as the *Central America–EC* agreement, are missing pairs, often due to uncoded accessions (e.g., the accession of Croatia). Others, like the *EFTA–Colombia* agreement, have incorrect entry-into-force dates. Additionally, several older agreements, including those between the EC and Egypt or Tunisia, are entirely absent from DESTA.

Cross-Validation I also cross-check the DESTA–WTO data with Mario Larch's RTA Database, which reveals additional discrepancies that I correct manually. The rationale is to construct the most comprehensive dataset on RTAs possible: whenever a preferential tariff appears in the tariff data, it can be reliably linked to a corresponding trade agreement. To this end, I add all pair–year observations where Mario Larch's data indicates the presence of an agreement not recorded in either DESTA or the WTO data.

In most cases, these are older agreements that are eventually superseded by newer agreements, which are recorded in the DESTA–WTO data. My objective is to utilize any available information on reported preferential tariffs while avoiding interpolation in years where the RTA information is less reliable. To address this, I treat these agreements as if they were new. Specifically, I interpolate only for the years in which Mario Larch's data indicates the presence of an agreement and no other data source provides conflicting information. I use the first available year in Mario Larch's data where the agreement is coded as active as entry-into-force year and assume that these agreements are not phased-in. I repeat the same procedure for bilateral agreements using Baier and Bergstrand's EIA database.

Adding Nonreciprocal Trade Arrangements The current DESTA–WTO–Larch/EIA dataset does not include any nonreciprocal trade arrangements. I construct these using Baier and Bergstrand's EIA database, which serves as the primary source. Since this information ends in 2017, I supplement and cross-check it with the WTO's PTA Database to include more recent years. For the Baier and Bergstrand data, I use the earliest available year as the entry into force, while for the WTO PTA data, I rely on the provided entry-into-force year. For all nonreciprocal trade arrangements, I assume that they are not phased-out.

There are two major shortcomings. First, Baier and Bergstrand's GSP data do not differentiate between nonreciprocal trade arrangement schemes that countries might be eligible for, which can vary significantly in product scope (e.g., AGOA versus GSP). Second, while the data account for full graduation—such as the EU revoking China's GSP eligibility in 2015—there is, to the best of my knowledge, no comprehensive and reliable information on product-level graduation.

With these data at hand, the final dataset contains all trade agreements, their names, entry-into-force years, and information on the final year by which tariff phase-outs must be completed. The information on these variables is compiled sequentially, prioritizing sources in the following order: WTO RTA Database, DESTA, Mario Larch's RTA Database, Baier and Bergstrand's EIA Database, and the WTO PTA Database.

Constructing of the Final Dataset The current version of the data contains all agreements that were ever in place (the data are identified by pair–agreement–year). This is not what I ultimately need, which is a coherent list of trade agreements that reflects changes in agreement names when these signal a potential change in the scope of products covered by tariff reductions. To determine the product scope of trade agreements, I distinguish between two types: agreements notified under Article XXIV of the GATT/WTO and those notified under the Enabling Clause. The key difference is that Article XXIV agreements are required to liberalize *substantially all trade*, while Enabling Clause agreements are not. Accordingly, I add information on whether an agreement covers substantially all trade.

To classify agreements, I proceed as follows. First, I use the WTO RTA Database, which indicates how each agreement was notified. For agreements not included in the WTO RTA Database but contained in DESTA, I rely on DESTA's depth variable. I classify agreements as Article XXIV if they include provisions stipulating the full elimination of tariffs. For agreements in Mario Larch's database that are not part of the DESTA–WTO data, I assume "Partial Scope Agreements" are notified under the Enabling Clause, while all others fall under Article XXIV. Similarly, for Baier and Bergstrand's EIA data, I classify agreements as Enabling Clause if they are labeled as "Preferential Trade Agreements" by Baier and Bergstrand. Finally, nonreciprocal trade arrangements are always classified as Enabling Clause agreements, as they are never notified under Article XXIV.

Ideally, I would have detailed information linking each preferential tariff to its originating trade agreement (e.g., whether it is governed by RTA^1 or RTA^2). With such data, I could directly map the preferential tariffs to the corresponding RTA at the agreement level. However, since this information is not available, I focus on preserving only the changes in trade agreements over time that are relevant for tariff modifications. Therefore, I want to keep for every pair–year observations one single agreement.

I begin by dropping agreements that have been superseded or from which countries have withdrawn. For each country pair–year observation, I retain only those agreements notified under Article XXIV of the GATT. When multiple Article XXIV agreements exist for the same pair–year, I prioritize the deepest agreement based on the depth variable provided by DESTA. If multiple Article XXIV agreements have the same depth, I keep the earliest agreement.

This approach preserves the objective of capturing significant updates to agreements that reflect substantial deepening, potentially extending tariff liberalizations to new sectors. For example, agreements between the European Community (EC) and Eastern European countries prior to their EU accession excluded agriculture, whereas within the EU, all tariffs are eliminated. The algorithm identifies this transition as a deepening event, as DESTA assigns different depth measures to these agreements. Conversely, as noted in the main text, agreements without major changes—such as the USMCA, which primarily introduced a minor adjustment to the rules of origin—do not alter the depth variable. In such cases, the algorithm retains NAFTA as the relevant agreement for the US–Mexico relationship, as no substantial deepening in tariff reductions occurred under the USMCA.

E Tariff Data

To address the missing tariff data, I begin by combining various primary sources. I will now outline the process of merging these sources, accessing the raw data, and performing the necessary data cleaning steps. The objective is to produce an unbalanced but uniquely identified $ijkt$ dataset for MFN and preferential tariffs, each compiled separately, and harmonized into a single nomenclature—HS1988/92.

E.1 Non Ad Valorem Tariffs

Another complication arises due to non-ad valorem tariffs. Regardless of the type of tariff—bound, MFN or preferential—a tariff can take one of two forms, i.e., ad valorem (for example, 8%) and non-ad valorem (for example, 1.22 USD/kg or 1.22 USD/kg + 8%). In theory, it is possible to convert non-ad valorem tariffs into ad valorem equivalents (AVEs) by dividing the non-ad valorem element of the tariff by the value of the product per unit. In practice, calculating ad valorem equivalents (AVEs) is very complicated; see Bouët et al. (2008) for a more detailed discussion.

Trains and the ITC offer AVEs for specific tariffs; however, they use different methods, yielding different AVEs.[40] For the IDB data, the phase-out schedules, and the data provided by national authorities, AVEs are simply not available.[41]

For most empirical analyses, it might be sensible to exclude non-ad valorem tariffs altogether since they vary even if no change in tariffs occurs. To facilitate this, the GTD flags all AVEs based on information provided by Trains and the ITC. For 2017, the available data suggests that less than 2% of country pair–products are subject to non-ad valorem tariffs. However, this most likely understates the true share because Trains/ITC record only the non-ad valorem tariffs for which Trains/ITC calculate AVEs. Whenever countries calculate AVEs themselves and end up reporting only ad valorem tariffs, it is impossible to trace back for which product the reported tariff is a true ad valorem tariff or is instead a non-ad valorem tariff. The EU is a prominent example of a jurisdiction that engages in this practice. Unfortunately, there is no way of systematically identifying all products with non-ad valorem tariffs. For practitioners, one solution might be to exclude extreme values of tariffs, e.g., those that suggest rates high as 150% or higher, as these have a high probability of being AVEs.

[40] From 2010 onward, the ITC has delivered tariffs to Trains. The only differences between the Trains and ITC data for the years 2010 to 2017 are observable for the ad valorem equivalents of non-ad valorem tariffs, as Trains and the ITC do not use the same method for computing the AVEs.

[41] The IDB ignores specific tariffs altogether, resulting, yet again, in missing data; a specific tariff of 1.22 USD is recorded as missing and a mixed tariff of 1.22 USD + 8% is recorded as equal to 8%.

To maintain consistency, I include only AVEs calculated by TRAINS. AVEs are also used to interpolate missing tariff rates. Since it is not possible to track the products for which the underlying tariffs are non-ad valorem tariffs, winsorizing offers the most effective way to exclude them.

E.2 Combining Reported MFN Tariffs

For MFN tariffs, I combine information from TRAINS, IDB, the ITC, and national sources for the United States and the European Union. For TRAINS, I use ad valorem equivalents (AVEs), while for all other sources, tariff lines or HS6 products containing non-ad valorem tariffs are excluded. Before merging these datasets, all individual data files are, where applicable, aggregated to the 6-digit level using unweighted averages and harmonized to the HS1988/92 nomenclature through concordance tables available via WITS (see Appendix G for details on accessing these tables). Additionally, all country ISO3 codes are standardized to the official ISO 3166 classification (https://www.iso.org/iso-3166-country-codes.html). Countries not included in the ISO 3166 list are dropped from the dataset.

All reported MFN tariffs from primary sources are combined and filled in according to the following order: if the TRAINS MFN tariffs are available, I use them first; otherwise, I use IDB, followed by ITC, and finally national sources. Section F outlines how exactly I access the respective raw data and what cleaning steps were performed.

The final dataset for the non-interpolated MFN tariffs is an unbalanced panel uniquely identified by importer, product (HS6 level), and year. Using this raw MFN tariff data, I then interpolate the missing MFN tariffs following the methodology outlined in Section C.

E.3 Combining Reported Preferential Tariffs

For preferential tariffs, I combine information from the phase-out schedules provided by the WTO RTA Database, TRAINS, the ITC, and national sources for the United States and the European Union. For TRAINS, I use ad valorem equivalents (AVEs), while for all other sources, tariff lines or HS6 products containing non-ad valorem tariffs are excluded. Before merging these datasets, all individual data files are, where applicable, aggregated to the 6-digit level using unweighted averages and harmonized to the HS1988/92 nomenclature through concordance tables available via WITS (see Appendix G for details on accessing these tables). Additionally, all country ISO3 codes are standardized to the official ISO 3166 classification (https://www.iso.org/iso-3166-country-codes.html). Countries not included in the ISO 3166 list are dropped from the dataset.

All reported preferential tariffs from primary sources are combined and filled in according to the following order: if the WTO RTA phase-out schedules are available, I use them first; otherwise, I use TRAINS, followed by ITC, and finally national sources.

Multiple RTAs To address multiple trade agreements, I assume that when multiple preferential tariff rates exist for ijk, the lowest rate is the one applied. At this stage, I also account for the possibility that countries do not report all preferential tariff schedules when multiple agreements exist, such as the case of the GSP and bilateral agreements for the EU and the Baltic countries prior to their accession to the EU, as discussed in the main text. To address this, I drop all $ijkt$ observations where the reported preferential tariff increases over time as this indicates misreporting.

Cross-Validation of Existence of Preferential Tariff Next, I add the RTA data and cross-validate that the reported preferential tariffs are actual preferential tariffs, i.e., that they correspond to an existing trade agreement. In only 7% of $ijkt$ cases, I do not find a corresponding agreement. For a subset of these cases, I verified that the reported "preferential" tariffs are indeed identical to MFN tariffs. I drop all preferential tariffs that I cannot cross-validate. By defining the presence of an RTA in the broadest possible manner—incorporating all available datasets to capture as many agreements as feasible—I am confident that this approach avoids omitting any valid RTAs. The alternative—over-interpolating the tariffs—appears to carry greater risk.

The final dataset for the non-interpolated preferential tariffs is an unbalanced panel uniquely identified by importer, exporter, product (HS6 level), and year. Using this raw preferential tariff data, I then interpolate the missing MFN tariffs following the methodology outlined in Section C.

F Accessing Reported Tariff Data

WITS provides three types of tariff rates: raw tariff schedules at the national tariff line level, lightly processed tariff rates at the HS6 level, "effectively applied tariffs " at the HS6 level as well as at other levels of aggregation. This section outlines the contents of each version and the corresponding download procedures.

F.1 Raw Data at the National Tariff Line Level

The raw tariff line data is the most detailed tariff data available, organized by importing country (reporter) and tariff schedule, including MFN and specific RTAs. National tariff lines are often more disaggregated than the 6-digit HS level. However, data availability depends on voluntary reporting by countries, leading to gaps in time series coverage. The dataset includes both non-ad valorem tariffs (e.g., 5$ per kg) and their ad valorem equivalents (AVEs). Bulk downloads are not supported, so the data must be downloaded separately for each importer–year combination, but all reported tariff schedules for that combination are included. Unfortunately, the nomenclature in which tariffs are reported is not explicitly specified in this version of the data.

Japan's tariff schedules, used in Section 4.3 to validate the interpolation algorithm, rely on this type of data. For consistency, I assumed that Japan adopts the most recent nomenclature—HS1988/92 until 1995, HS1996 from 1996 onward, and so forth.

Download To download the raw tariff line data, log in at https://wits.worldbank.org/ and navigate to *Quick Search*. Select *Tariff - View and Export Raw Data*, and then specify the following options:

- Data Source: *Trains-Total (Incl. AVE)*,
- Market: Desired importer,
- Year: Desired year,
- Duty Code: *All Duty Codes*,
- Estimation Method: *UNCTAD Method*.

Save the query and download the results in CSV format. The same procedure can be followed using *WTO-IDB* as the data source, which includes MFN and preferential tariffs (but no AVEs), or *WTO-CTS*, which provides bound MFN tariffs. Note that results must be downloaded separately for each importer–year combination.

F.2 Lightly Processed Tariff Rates (HS6 Level)

While the raw tariff schedules provide the most detailed data, they are impractical for putting together cross-country tariff rates due to the lack of a bulk download function. WITS offers a lightly processed version of the tariff line data, aggregated to the HS6 level using unweighted averages. This is available for both MFN and preferential tariffs. Additionally, WITS provides a version where AVEs are included for both MFN and preferential tariffs.

For MFN tariffs, the identifying dimensions are importer, HS6 code (in mixed nomenclature), and year. For preferential tariffs, the identifying dimensions are importer, tariff schedule, HS6 code (in mixed nomenclature), and year. The tariff schedule refers to the labels or names used by reporters to categorize tariff types, such as "General System of Preferences" or "Preferences for Mexico." However, these labels are often ambiguous and not explicitly linked to specific RTAs, limiting their practical usefulness. At this stage, the partner may represent a country aggregate (e.g., ASEAN) rather than a single country.

Download To download lightly processed tariff rates, log in at https://wits.worldbank.org/, navigate to *Advanced Query* → *Bulk Download (TRAINS)*, and define a query name. You will need to create four separate queries, one for each tariff type. For each query, select *MFN Applied Rates (including AVE)* as the tariff type and use the *UNCTAD Method* for estimation. The four tariff types available for download are:

(i) MFN Applied Rates,

(ii) MFN Applied Rates (including AVE),

(iii) Preferential Rates, and

(iv) Preferential Rates (including AVE).

Download the results. Once the query is complete, navigate to *Results* → *Download and View Results*, click on *Download*, and save the files. Note that WITS provides a separate file for each reporter–year combination and tariff type (e.g., `Pref_H0_AUS_1991.zip`). Hence, one ends up with many thousands of files containing the preferential tariffs.

Two auxiliary files are also required. The first, `TRAINS Preference Beneficiaries`, provides a concordance between country aggregates and their individual countries (e.g., ASEAN includes Brunei, Indonesia, Malaysia, the Philippines, Singapore, Vietnam, and Thailand). However, this file uses numerical codes rather than ISO3 country codes. To download, log in at https://wits.worldbank.org/, navigate to *Support Materials* → *TRAINS Tariff Measures and Preference Beneficiaries*, and save the file in CSV format.

The second file, `Country List`, contains a concordance between the numerical country codes used in WITS and the corresponding ISO codes or country names. This file is available under *Support Materials* → *WITS Reference Data* → *Countries*. Note that WITS does not use official ISO3 codes as per the ISO 3166 classification (https://www.iso.org/iso-3166-country-codes.html).

Cleaning Steps The cleaning process begins by appending all individual country-level files (e.g., `Pref_H0_AUS_1991.zip`) for the same tariff type, resulting in four large datasets: `Pref_all.dta`, `PrefAVE_all.dta`, `MFN_all.dta`, and `MFNAVE_all.dta`.

Next, the AVE datasets are merged with their corresponding non-AVE datasets to produce two combined datasets. The first dataset contains preferential tariffs and includes the following dimensions: importer, tariff schedule, HS6 product code, year, and two tariff variables: the preferential tariff rate (`pref`) and its ad valorem equivalent (`prefAVE`). The second dataset contains MFN tariffs and includes importer, HS6 product code, year, and two tariff variables: the MFN tariff rate (`mfn`) and its ad valorem equivalent (`mfnAVE`).

Next, I add ISO codes. For MFN tariffs, this process is straightforward: the numerical country codes of the reporting countries from the datasets are first matched to ISO3 codes using the `Country List` file. These ISO3 codes, which do not conform to the official ISO 3166 classification, are subsequently converted to the official ISO3 codes.

For preferential tariffs, an additional step is required: the tariff schedule is first disaggregated to individual countries using the `TRAINS Preference Beneficiaries` file, and ISO3 codes are subsequently added. This results in a dataset where `importer-exporter-HS6-year` is not uniquely identified, as some importers have multiple trade agreements through which the same exporter can access preferential tariffs. In such cases, the lowest available preferential tariff rate is assumed to apply. To address potential underreporting of preferential tariff schedules in instances of multiple agreements—such as the coexistence of the Generalized System of Preferences (GSP) and bilateral agreements for the EU and the Baltic countries before their EU accession—I treat all $ijkt$ observations where where reported preferential tariff rates increase over time as missing, as this pattern indicates likely failure to report all tariff schedules.

Finally, all HS6 products, which are reported in mixed nomenclatures, are harmonized to the HS88/92 classification.

The cleaned datasets based on TRAINS data, containing information on raw tariffs (i.e., non-interpolated tariff rates) for both MFN and preferential tariffs, are uniquely identified by importer (and exporter, in the case of preferential tariffs), HS6 product code (using the HS88/92 nomenclature), and year. All other primary sources have been processed such that they match these identifying dimensions, as outlined below.

F.3 WTO-IDB

In addition to tariff information collected from UNCTAD (TRAINS), WITS also provides tariff data collected by the WTO. Unfortunately, these data are not available in a lightly processed version, as is the case with TRAINS, through bulk download. Instead, users must access the data using the "Tariff and Trade Analysis" tool. However, tariff rates obtained through this tool suffer from the issues discussed in this paper, i.e., positive selection and false interpolation.

For MFN tariffs, which lack a partner dimension, the data remain usable. Positive selection can be, at least partially, addressed by downloading the data at the HS6 level and retaining the first non-missing MFN tariff for each `importer-HS6-year`. Additionally, I do not use WTO-IDB as the primary source but instead rely on TRAINS, using WTO-IDB data only to fill missing values. False interpolation is not an issue for MFN tariffs.

Conversely, preferential tariff rates from WTO-IDB are unusable due to false interpolation. These rates are accessed under the tariff type "Effectively Applied Rates," which has been used in prior work despite its limitations (e.g., Lashkaripour and Lugovskyy (2023)).

Download To download WTO-IDB tariff data, log in at https://wits.worldbank.org/, navigate to *Advanced Query* → *Tariff and Trade Analysis*, and define a query name and description, choosing *WTO-IDB* as data source. Specify the following options:

- Importers/Reporters: All countries,
- Products: All 6-digit HS codes,
- Exporters/Partners: All countries,
- Year: One year at a time (starting with 1995, the first available year).
- Tariff Type: *MFN Applied Rates*, *Effectively Applied rates*, and *Preferential rates*

Once the parameters are defined, click *Submit.* After the query is complete, navigate to *Results* → *Download and View Results* → *Download Data*, and save the file. Repeat this process year by year for all years (1995–2021), as downloading multiple years at once may cause the process to fail.

The resulting data are structured at the `importer-HS6 (mixed nomenclature)-year` level and include duty types such as MFN, AHS, and preferential tariffs. Among these, MFN tariffs are the only duty type that can be reliably used without concerns about data accuracy or flaws.

Cleaning Steps The cleaning process begins by appending the data for all years and retaining only observations corresponding to MFN tariffs. Numerical country codes for the reporting countries are then merged with ISO3 codes using the `Country List` file, and these are subsequently converted to the official ISO 3166 classification. For each importer–HS6 product–year, only the first non-missing MFN tariff is retained. Finally, all HS6 product codes are harmonized to the HS88/92 classification to ensure consistency across years and countries.

F.4 ITC Market Access Map

The ITC Market Access Map provides an additional source of raw tariff data, comparable to the raw tariff schedules available through WITS (see Section F.1). The data are structured by `importer-tariff line-agreement` and include all reported preferential tariff schedules as well as MFN and bound MFN tariffs, where available.

Download To download the data, visit https://www.macmap.org/, click on *Download Data*, and log in. Navigate to *Bulk Download* and select the following options: *Tariff*, *Applied Tariffs*, *Effectively Applied by Partner*, *By Each Agreement*, and *NTLC*. Then, specify the following parameters:

- Reporter: Select batches of 25 countries (repeat until all reporters are downloaded),
- Year: Select all available years (availability depends on the chosen reporters).

The downloaded files will be organized by `reporter-tariff line-year-agreement` and saved as multiple small files. Each file corresponds to a specific reporter (e.g., `MacMap-AFG_2007_Tariff_NTLC_agr.txt`).

Repeat this process for all reporting countries, saving the files for each batch of 25 countries.

You will also need the data file `Country Codes`, which can be found under *Download Reference Datasets* on the ITC Market Access Map website. This file provides a mapping between numeric country codes and ISO3 codes; however, note that the ISO3 codes in this file are not aligned with the official ISO 3166 classification and require minor adjustments.

Cleaning Steps The cleaning process starts by assigning ISO codes to both importers and exporters using the `Country Codes` file. All reporter files are then combined for each year, and the data are aggregated to the HS6 level and converted to the HS88/92 nomenclature. Since the ITC Market Access Map data do not specify the nomenclature used, and countries may report using different classifications (e.g., HS1996 in 2010 instead of HS2007), the appropriate nomenclature is identified by comparing the number of matches achieved under each classification. Specifically, HS1988/92, HS1996, HS2002, and subsequent nomenclatures are tested, with the one yielding the highest number of matches assigned to the importing country for that year. Finally, all non-ad valorem tariff lines are excluded from the dataset.

F.5 Phase-Out Schedules

The WTO provides detailed phase-out schedules for a subset of trade agreements, accessible through the WTO RTA Database. This database is organized by agreement, and clicking on an agreement entry reveals various tabs, depending on availability. Every agreement at least contains the tab labeled *Basic Information*, which contains details such as the agreement name and original signatories. For many agreements, an additional tab labeled *Trade-Related Data* is available, which includes the phase-out schedules for the respective trade agreement. For example, the Türkiye–Montenegro agreement, accessible at https://rtais.wto.org/UI/PublicShowRTAIDCard.aspx?rtaid=708, contains multiple tabs, including *Trade-Related Data*, where the schedules for each signatory are provided.

Download Since a bulk download option is not available, the manual process requires navigating to https://rtais.wto.org/, selecting *Explore the Data*, followed by *RTAs in Force*, and individually accessing each trade agreement. Given the large number of agreements, this approach is highly impractical.

Instead, I extracted phase-out schedules for all agreements by automating the process. Using the RTA-ID, which is embedded in the URL of each agreement, I looped through the IDs to systematically retrieve the data without requiring manual interaction.

The download was completed in June 2019, based on the version of the data available at that time.

Cleaning Steps The raw tariff schedules available through the WTO RTA Database are not standardized, requiring manual harmonization to ensure that all Excel files follow an identical structure, including consistent variable names and other key dimensions.

Country codes are assigned using the official ISO3 classification. Non-ad valorem tariff lines are excluded from the dataset. The data are then aggregated to the HS6 level and converted to the HS1988/92 nomenclature.

Since these data include full phase-out schedules, the last available year often extends far into the future. For example, the EFTA–Ecuador agreement, which entered into force in 2020, includes tariff cuts planned through 2036. To maintain consistency with my sample, I keep only data up to 2021, the final year of the sample period.

G Other Data

G.1 Concordance Tables

All concordance tables are downloaded from WITS. Proceed as follows:

1. Go to https://wits.worldbank.org/ and log in.
2. Click on *Support Materials*.
3. Select *Product Nomenclatures and Concordances*.
4. Choose the relevant *Product Concordances* (e.g., HS 1996) and *Concordances* (e.g., HS1988/92).
5. Click *Download*, name the job, and wait for processing.
6. Navigate to *Results* → *Download and View Results* → *Download Data*.

G.2 Baci-Trade Data

BACI is an international trade database developed by CEPII (Gaulier and Zignago 2010). It provides detailed bilateral trade flows at the product level using the Harmonized System (HS) nomenclature. The data are reconstructed and reconciled to ensure accuracy and consistency. The data are available at https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=37. The version used in this paper is April 9, 2024, and I use the HS92 classification.

G.3 Country ISO3 Codes

All country ISO3 codes are standardized to the official ISO 3166 classification, obtained from https://www.iso.org/obp/ui/#search and downloaded in February 2020. The table titled *officially assigned codes* was copied into Excel for further processing. Countries not included in the ISO 3166 list were excluded from the dataset.

G.4 Included Countries

Table G.1. Included Countries

Country	Income Group
Afghanistan	LDCs
Albania	LoM Europe
Algeria	LoM Americas
Angola	LDCs
Anguilla	LoM Africa
Antigua and Barbuda	LoM Africa
Argentina	LoM Africa
Armenia	LoM Asia
Aruba	LoM Africa
Australia	High
Austria	High
Azerbaijan	LoM Asia
Bahamas	LoM Africa
Bahrain	LoM Asia
Bangladesh	LDCs
Barbados	LoM Africa
Belarus	LoM Europe
Belgium	High
Belize	LoM Africa
Benin	LDCs
Bermuda	LoM Africa
Bhutan	LDCs
Bolivia	LoM Africa
Bosnia and Herzegovina	LoM Europe
Botswana	LoM Americas
Brazil	LoM Africa
Brunei Darussalam	LoM Asia
Bulgaria	High
Burkina Faso	LDCs
Burundi	LDCs
Cabo Verde	LoM Americas
Cambodia	LDCs
Cameroon	LoM Americas

Continued on next page

Continued from previous page

Country	Income Group
Canada	High
Cayman Islands	LoM Africa
Central African Republic	LDCs
Chad	LDCs
Chile	LoM Africa
China	LoM Asia
Colombia	LoM Africa
Comoros	LDCs
Congo	LDCs
Congo	LoM Americas
Cook Islands	LoM Asia
Costa Rica	LoM Africa
Croatia	High
Cuba	LoM Africa
Cyprus	High
Czechia	High
Côte d'Ivoire	LoM Americas
Denmark	High
Djibouti	LDCs
Dominica	LoM Africa
Dominican Republic	LoM Africa
Ecuador	LoM Africa
Egypt	LoM Americas
El Salvador	LoM Africa
Equatorial Guinea	LDCs
Eritrea	LDCs
Estonia	High
Ethiopia	LDCs
Fiji	LoM Asia
Finland	High
France	High
French Polynesia	LoM Asia
Gabon	LoM Americas

Continued on next page

Continued from previous page

Country	**Income Group**
Gambia	LDCs
Georgia	LoM Asia
Germany	High
Ghana	LoM Americas
Gibraltar	LoM Europe
Greece	High
Grenada	LoM Africa
Guatemala	LoM Africa
Guinea	LDCs
Guinea-Bissau	LDCs
Guyana	LoM Africa
Haiti	LDCs
Honduras	LoM Africa
Hong Kong	High
Hungary	High
Iceland	High
India	LoM Asia
Indonesia	LoM Asia
Iran	LoM Asia
Ireland	High
Israel	High
Italy	High
Jamaica	LoM Africa
Japan	High
Jordan	LoM Asia
Kazakhstan	LoM Asia
Kenya	LoM Americas
Kiribati	LDCs
Korea	High
Kuwait	LoM Asia
Kyrgyzstan	LoM Asia
Lao People's Democratic Republic	LDCs
Latvia	High

Continued on next page

Continued from previous page

Country	**Income Group**
Lebanon	LoM Asia
Lesotho	LDCs
Liberia	LDCs
Libya	LoM Americas
Liechtenstein	High
Lithuania	High
Luxembourg	High
Macao	High
Macedonia	LoM Europe
Madagascar	LDCs
Malawi	LDCs
Malaysia	LoM Asia
Maldives	LoM Asia
Mali	LDCs
Malta	High
Mauritania	LDCs
Mauritius	LoM Americas
Mayotte	LoM Americas
Mexico	LoM Africa
Micronesia	LoM Asia
Moldova	LoM Europe
Mongolia	LoM Asia
Montenegro	LoM Europe
Montserrat	LoM Africa
Morocco	LoM Americas
Mozambique	LDCs
Myanmar	LDCs
Namibia	LoM Americas
Nauru	LoM Asia
Nepal	LDCs
Netherlands	High
New Zealand	High
Nicaragua	LoM Africa

Continued on next page

Continued from previous page

Country	Income Group
Niger	LDCs
Nigeria	LoM Americas
Norway	High
Oman	LoM Asia
Pakistan	LoM Asia
Palau	LoM Asia
Palestine, State of	LoM Asia
Panama	LoM Africa
Papua New Guinea	LoM Asia
Paraguay	LoM Africa
Peru	LoM Africa
Philippines	LoM Asia
Poland	High
Portugal	High
Qatar	LoM Asia
Romania	High
Russian Federation	LoM Europe
Rwanda	LDCs
Saint Kitts and Nevis	LoM Africa
Saint Lucia	LoM Africa
Saint Pierre and Miquelon	LoM Africa
Saint Vincent and the Grenadines	LoM Africa
Samoa	LDCs
Sao Tome and Principe	LDCs
Saudi Arabia	LoM Asia
Senegal	LDCs
Serbia	LoM Europe
Seychelles	LoM Americas
Sierra Leone	LDCs
Singapore	High
Slovakia	High
Slovenia	High
Solomon Islands	LDCs

Continued on next page

Continued from previous page

Country	**Income Group**
South Africa	LoM Americas
Spain	High
Sri Lanka	LoM Asia
Sudan	LDCs
Suriname	LoM Africa
Swaziland	LoM Americas
Sweden	High
Switzerland	High
Syrian Arab Republic	LoM Asia
Taiwan	High
Tajikistan	LoM Asia
Tanzania, United Republic of	LDCs
Thailand	LoM Asia
Timor-Leste	LDCs
Togo	LDCs
Tonga	LoM Asia
Trinidad and Tobago	LoM Africa
Tunisia	LoM Americas
Turkey	LoM Asia
Turkmenistan	LoM Asia
Tuvalu	LDCs
Uganda	LDCs
Ukraine	LoM Europe
United Arab Emirates	LoM Asia
United Kingdom of Great Britain and Northern Ireland	High
United States of America	High
Uruguay	LoM Africa
Uzbekistan	LoM Asia
Vanuatu	LDCs
Venezuela	LoM Africa
Viet Nam	LoM Asia
Wallis and Futuna	LoM Asia
Yemen	LDCs

Continued on next page

Continued from previous page

Country	Income Group
Zambia	LDCs
Zimbabwe	LoM Americas

Note: The table includes all importing countries and their corresponding income groups. Income groups follow the World Bank classification: least developed countries (LDCs), regional low- or middle-income (LoM) countries, and high-income countries.

G.5 "Effectively Applied Tariff" in WITS

To download the so-called "Effectively Applied Tariff" in WITS, which is widely used by researchers, proceed as follows:

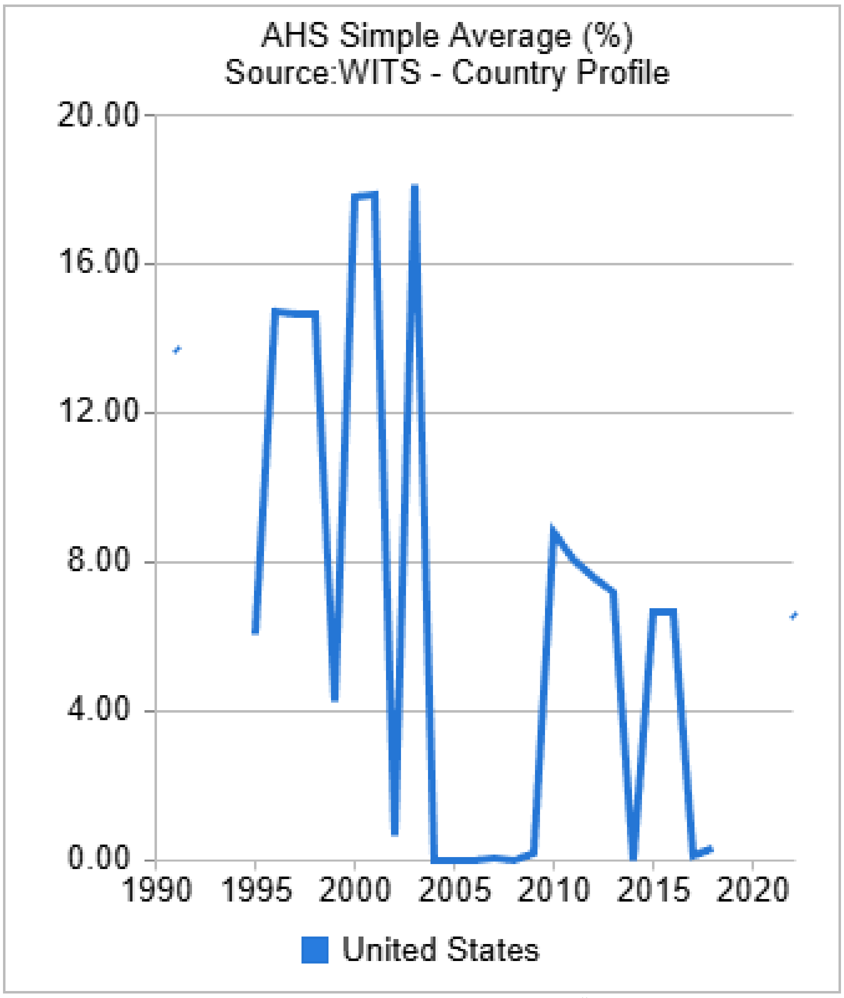
Log in at https://wits.worldbank.org/, navigate to *Advanced Query* → *Tariff and Trade Analysis*, and define a query name and description. Choose either *WTO-IDB* or *TRAINS* as the data source (most researchers used TRAINS) and specify the following options:

- Importers/Reporters: All countries,
- Products: All 6-digit HS codes,
- Exporters/Partners: All countries,
- Year: All relevant years,
- Tariff Type: *Effectively Applied Rates*.

If a higher level of product aggregation is selected (e.g., ISIC at the 4-digit level), WITS uses the effectively applied tariff at the HS6 level—which is subject to false interpolation and selection bias—and calculates unweighted averages to concord it to the more aggregated ISIC (4-digit) level.

To access country-level indicators, navigate to https://wits.worldbank.org/referencedata.html, then proceed as follows: *Trade Stats* → *AHS Simple Average* → click on *By Indicator*, select *Importer = Mexico* and *Partner = United States*. This process generates the graph depicted in Figure G.1.

FIGURE G.1. "EFFECTIVELY APPLIED TARIFF" IN WITS



Note: The graphs shows the unweighted average of the so-called "Effectively Applied Tariff" between Mexico and the U.S. available through WITS.