Global Tariff Database (GTD)

Variable Documentation and User Guide

Feodora A. Teti

Version: v_beta1-2024-12

January 2025

For inquiries or feedback, please contact:
teti@ifo.de

Introduction

The Global Tariff Database (GTD) is a comprehensive dataset providing detailed information on *bilateral statutory tariff rates* across *200 importing and exporting* countries over a 34-year period *(1988–2021)*. The GTD is based on the *HS6-level* classification (HS88/92 Nomenclature) and includes both (applied) Most-Favored-Nation (MFN) and preferential tariff rates, resulting in a dataset with more than 6.9 billion observations.

The GTD does *not include* certain types of tariffs or trade measures beyond statutory tariff rates. For example, safeguard measures, anti-dumping duties, and tariffs introduced during the US–China trade war are not part of this dataset. Researchers requiring such information should consult *other* specialized data sources.

This user guide focuses exclusively on the *aggregated version* of the GTD, which organizes the data at higher aggregation levels beyond the HS6 classification and is available on my website at https://feodora-teti.weebly.com/. Specifically, it covers aggregated tariff data by country pairs and sectors. These aggregated versions provide a more accessible and practical perspective for researchers analyzing broad trends and patterns in international trade policy.

The guide offers a detailed overview of the variables included in the aggregated GTD, along with their definitions and descriptions. It also provides essential notes on data coverage and limitations, including practical advice on handling outliers and extreme values. Furthermore, it outlines proper citation practices for those using the GTD in academic or professional contexts.

Citation

If you use the Global Tariff Database (GTD) in your research, please cite it as follows:

"Feodora Teti's Global Tariff Database (v_beta1-2024-12) from Teti (2024)"

The corresponding reference is:
Teti, Feodora A. (2024), *Missing Tariffs*, CESifo Working Papers No. 11590.

Feedback

The Global Tariff Database (GTD) is currently in its *beta version* (`v_beta1-2024-12`). While every effort has been made to ensure accuracy, errors or inconsistencies may still be present.

I kindly ask users to *report any issues* they encounter while working with the data. Your feedback is invaluable in improving the dataset and refining the algorithm used to compile it. Additionally, I welcome *suggestions* for new features or enhancements that could make the GTD more useful for your research.

If you would like to share *your feedback* or have any questions, please feel free to reach out.

Feodora Teti
Email: `teti@ifo.de`

Variable Description

This section provides a *detailed description* of the variables included in the aggregated version of the GTD. Trade-weighted tariffs are constructed using trade data from CEPIIÄs BACI starting from 1995.[1] The dataset contains the following variables:

Variable	Description
`iso1`	Importer (ISO 3)
`iso2`	Exporter (ISO 3)
`year`	Year (1988–2021)
`sector`	Respective sector dimension (if applicable)
`tariff`	Unweighted average bilateral tariff, varies by i–j (or i–j–sector, if applicable)
`mfn`	Unweighted average MFN tariff, varies by i (or i–sector, if applicable)
`tariff_w`	Trade-weighted average bilateral tariff (BACI, contemporaneous trade values), varies by i–j (or i–j–sector)
`mfn_w`	Trade-weighted average MFN tariff (BACI, contemporaneous trade values), varies by i (or i–sector)

[1] Gaulier, G. and Zignago, S. (2010) BACI: International Trade Database at the Product-Level. The 1994-2007 Version. CEPII Working Paper, N°2010-23.

`tariff95`	Unweighted average bilateral tariff with winsorized outliers (replace tariff = p_{95} if tariff $> p_{95}$), varies by i–j (or i–j–sector)
`mfn95`	Unweighted average MFN tariff with winsorized outliers (replace MFN = p_{95} if MFN $> p_{95}$), varies by i (or i–sector)
`tariff95_w`	Trade-weighted average bilateral tariff with winsorized outliers (replace tariff = p_{95} if tariff $> p_{95}$; BACI, contemporaneous trade values), varies by i–j (or i–j–sector)
`mfn95_w`	Trade-weighted (BACI, contemporaneous trade values) average MFN with winsorized outliers (replace MFN = p_{95} if MFN $> p_{95}$; BACI, contemporaneous trade values), varies by i (or i–sector)

Outliers and Recommendations

The tariff data provided in the Global Tariff Database (GTD) includes some extreme outliers at the HS6 level. For instance, in 2021, the maximum recorded tariff rate is 3000%, while the mean is 7.4% and the median is 5.0%. Such extreme values can significantly impact statistical analyses and are often the result of ad valorem equivalents of non-ad valorem tariffs.

To address this issue, the aggregated version of the GTD includes outlier-corrected tariff variables. These variables have been Windsorized by capping extreme values above the 95th percentile at the 95th percentile threshold. In my experience, using the Windsorized variables yield more reliable results, making them the preferred choice for most research applications.

Algorithm

For a detailed description of the filling algorithm, see "*Missing Tariffs*" (Teti, 2024). The main text and the online appendix provides a detailed explanation of the algorithm's steps and methodology.