

1989–1996 U.S. Tariff Data

Miguel Acosta
Federal Reserve Board

Lydia Cox*
University of Wisconsin – Madison

December 11, 2023

This document describes our digitization of the 1989–1996 U.S. Harmonized Tariff Schedule (HTS). We stop in 1996 because data from 1997 are available in machine-readable format from the United States International Trade Commission (USITC). [Feenstra, Romalis and Schott \(2002\)](#) (FRS, hereafter) have also digitized the 1989–1996 tariff rates. Our implementation has two main advantages, and two main disadvantages over that of FRS. The first advantage is that we have corrected tariff rates on about 10 percent of the products reported by FRS. The second advantage is that we have provided replication materials to go from the files available on the USITC’s website to our final dataset. The first disadvantage is that we have only digitized tariff MFN and column-2 tariff rates, and have thus ignored preferential tariff rates. The second disadvantage is that we do not have data from 1990. Users can weigh the advantages and disadvantages of each dataset.

Digitization Procedure We started from the USITC’s HTS Archive webpage, and used the “Basic Publications” for each year available before 1996: 1989 and 1991–1996.¹ We downloaded the PDFs for each chapter of the HTS in each of these years, converted them to text format using the Python wrapper of the `tabula-java`, and parsed the text into numeric tariff rates using more Python code. In many cases, formatting errors in the PDF tables led to missing tariff rates, so we added these manually by inspecting the PDF pages by hand. We benefited from having the dataset of FRS available to us for finding and correcting errors in earlier versions of our code. As FRS largely do, we exclude chapters 91 (watches), 98 (special classifications) and 99 (temporary rates). As seen in Table 1, our dataset has just over 62,000 products, starting at about 8,500 in 1989 and ending at nearly 10,000 in 1996.

Corrections to FRS Our dataset contains corrections to the tariff rates of about ten percent of the product-years in the dataset of FRS. There are three types of corrections, summarized in Table 1. The first, and most prominent corrections pertained to products with compound tariff rates, i.e., products with both an *ad valorem* and specific tariff. In these cases—which constitute just under 10 percent of products—FRS had the *ad valorem* component set to zero. The second type of correction involved adding tariff rates for just over 900 product-years that were not included in the dataset of FRS. Most often, we found that these cases were on pages of the PDFs with formatting issues.² Finally, there were nearly 2,000 product-years in the FRS dataset that did not appear in the tariff schedule for the corresponding year. Before 1996, this is very infrequent.³ However, the 1996 data includes 1,956 products that were in neither the 1996 tariff schedule, nor the 1996 U.S. import data ([Schott, 2008](#)).

References

- Feenstra, Robert C, John Romalis, and Peter K Schott, “U.S. Imports, Exports, and Tariff Data, 1989-2001,” Working Paper 9387, National Bureau of Economic Research December 2002.
- Schott, Peter K, “The relative sophistication of Chinese exports,” *Economic Policy*, 2008, 23 (53), 6–49.

*Acosta: miguel.acosta@frb.gov. Cox: lydia.cox@wisc.edu. We would be happy to hear of any mistakes/issues that you find when working with this dataset. The views expressed here are solely our responsibility, and should not be interpreted as reflecting the view of the Board of Governors of the Federal Reserve System, or of any other person associated with the Federal Reserve System.

¹The Archive is located at <https://www.usitc.gov/tata/hts/archive/list>.

²A frequent case was: The top of page n was printed at the bottom of page $n - 1$. In those cases, data from page n was typically missing.

³While we did not track down which year of the HTS each of these products may have come from, we did find that in some cases, the product may have come from a tariff schedule of a slightly different year. For example 9101.12.80 (listed in 1989) is not in the 1989 HTSA Basic Publication, nor in the 1989 HTSA Supplement 1, though it is in the 1990 Supplement 2 file.

Table 1: Summary of New Dataset and Corrections

Year	Products	Corrections		
		Compound Rates	Missing	Not in Schedule
1989	8469	849	96	3
1991	8535	827	106	10
1992	8563	845	117	3
1993	8615	881	166	6
1994	8640	883	174	0
1995	9655	1012	205	1
1996	9916	1036	41	1956
Total	62393	6333	905	1979