

Data Sources

Zeyuan Xiong (Victor)

2/8/2021

Data sources

File name	Content	Source
Geneva47_UNTC	Geneva 1947 MFN Schedules	UNTC official website, Registration number A-814, Volume number 61. https://treaties.un.org/doc/Publication/UNTS/Volume%2061/v61.pdf
Annecy_UNTC	Annecy 1949 MFL Schedules	UNTC official website, Registration number A-814, Volume number 63. https://treaties.un.org/doc/Publication/UNTS/Volume%2063/v63.pdf
Torquay_UNTC	Torquay 1951 MFL Schedules	UNTC official website, Registration number A-814, Volume number 144. https://treaties.un.org/doc/Publication/UNTS/Volume%20144/v144.pdf
Geneva56_UNTC	Geneva 1956 MFL Schedules	UNTC official website, Registration number A-814, Volume number 245. https://treaties.un.org/doc/Publication/UNTS/Volume%20245/v245.pdf
Dillon_UNTC	Dillon 1960 MFL Schedules	UNTC official website, Registration number A-814, Volume number 440. https://treaties.un.org/doc/Publication/UNTS/Volume%20440/v440.pdf
Kennedy_UNTC	Kennedy 1964 MFL Schedules	UNTC official website, Registration number A-814, Volume number 624. https://treaties.un.org/doc/Publication/UNTS/Volume%20624/v624.pdf
Tokyo_UNTC	Tokyo 1979 MFL Schedules	UNTC official website, Registration number A-814, Volume number 1189. https://treaties.un.org/doc/Publication/UNTS/Volume%201189/v1189.pdf
Uruguay_UNTC_1	Uruguay 1988 MFL Schedules up to Ch. 63	UNTC official website, Registration number A-814, Volume number 1632. Note that Volume 1632 seems the same file as Volume 1633. https://treaties.un.org/doc/Publication/UNTS/Volume%201632/v1632.pdf
Uruguay_UNTC_2	Uruguay 1988 MFL Schedules, after Ch. 63	UNTC official website, Registration number A-814, Volume number 1634. Note that Volume 1634 seems the same file as Volume 1635. https://treaties.un.org/doc/Publication/UNTS/Volume%201634/v1634.pdf

File name	Content	Source
Torquay (black white)	Geneva47, Annecy, Torquay MFN Schedules combined	The hard copy was borrowed from the Library of University of Texas. We then scanned and digitized the copy. https://search.lib.utexas.edu/discovery/fulldisplay?context=L&vid=01UTAU_INST:SEARCH&search_scope=MyInst_and_CI&tab=Everything&docid=alma991056424989706011
Tariff Act of 1930 cleaner	Tariff schedule of 1930 Smoot-Hawley Tariff Act	Tariff Act of 1930: “AN ACT To provide revenue, to regulate commerce with foreign countries, to encourage the industries of the United States, to protect American labor, and for other purposes.” June 17, 1930 File: https://govtrackus.s3.amazonaws.com/legislink/pdf/stat/46/STATUTE-46-Pg590.pdf ; For citation information, search the following for “page: 590”: https://github.com/unitedstates/legisworks-historical-statutes/blob/master/data/046.yaml

Digitization Process

1. We began the digitization with “**Torquay (black white)**”, which is from a hard copy borrowed from the University of Texas library. We scanned the hard copy, conducted OCR and digitized the product descriptions in R. By running the R packages “*pdftools*”, we obtain an editable excel file of all the product descriptions in the book. This hard copy turns out to be a comprehensive schedule that contains the commitments made in the first three rounds combined (Geneva47, Annecy and Torquay). We began hand-coding the PDF files that Bob Staiger had found and shared with Kristy.
2. We later found more systematic data sources (the UNTC series) of each individual round (Geneva47 to Uruguay). We manually entered the tariff rates for each round based on the benchmark we constructed in Step 1.
3. Finally we entered the tariff data for Smoot-Hawley. Note that the official schedules for all rounds from Geneva47 to Dillon use the Smoot-Hawley paragraph/line numbers. We modified the list of descriptions (tariff lines) as we went to include all the products in Smoot-Hawley and the rounds up through Dillon (as of 2/8/2021, we are in the process of adding the final three GATT rounds). This is so far for the United States only. Note our dataset includes all products that are in Smoot-Hawley even if there were no commitments made under the GATT.
4. To check the reliability of our benchmark file, we compared the three individual rounds we have from the UNTC website: “**Geneva47_UNTC**”, “**Annecy_UNTC**”, and “**Torquay_UNTC**” with the consolidated version we initially found, namely “**Torquay (black white)**” and found no differences in tax rate value. However, we did observe discrepancies in descriptions for specific products, especially those lines involving tax intervals, e.g., Paragraph 53 in Geneva47 “corn or maize, 3 cents per lb, but no less than 10% but nor more than 20% ad val.”
5. The Excel file is then saved as csv (UTF-8) format and imported in R for cleaning and analysis.