**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**data\_847**

**country:** The country from which the item was imported. This variable also includes “Other British Possessions”, “Other Foreign Countries”, and “Others” as three additional sources, which we keep here for illustrative purposes, but do not include in our regression analyses.

**good\_name:** This variable details the good name that is used for our consistent balanced panel analysis.

**name**: This gives the name of the 847 items, which are then consolidated to produce our 258 consistent series “good\_name”.

**year**: The year in question. From 1924 to 1938.

**prnation**: A dummy indicating if the country was a member of the British Empire or Commonwealth.

**sheetnum:** An indicator of the SITC 3 digit code associated with the product name.

**sheetname:** The SITC 3 digit code name.

**SITC:** A character vector indicating the SITC 3 digit code associated with the product in question.

**broad\_category:** Our broad classification of products: Agricultural, Revenue, Manufacturing, and Raw Materials

**value**: The total value of imports into Britain of a product from a country in a year.

**tariff:** The ad valorem tariff applied to the product-country-year in question. Measured as per cent.

**exclude:** A dummy indicating whether this specific item is to be excluded from the consistent series of 258 goods.

**item\_number:** A numeric variable that maps to the numeric labels of the “name” variable.

**good\_number:** A numeric variable that maps to the numeric labels of the “good\_name” variable.

**broad\_category\_number:** A numeric variable that maps to the numeric labels of the “broad\_category” variable.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**data\_258**

**good\_name:** The name identifier for the good based on our consistent classification scheme.

**year:** The year in question.

**prnation:** As above.

**sheetnum:** As above.

**sheetname:** As above.

**broad\_category:** As above.

**value:** As above; it is the sum of value for the various “name” items that comprise each of the good\_names.

**tariff:** Since data\_258 is an aggregated version of data\_847, we have multiple product items contained within each “good\_name”. We take the unweighted average of the tariff rates on each item (“name”) to be the tariff rate for each “good\_name” reported in these data.

**tariffw:** Instead of taking a simple average, as in “tariff”, this variable is created by taking a weighted average where the weights are given by the total value of trade for each of the products in question. For example, let a good\_name called A be composed of two items from the data\_847 data, B and C. If the tariff on B was 10% and the tariff on C was 6% this would mean the “tariff” variable would be 8%. However, if B represented 90% of the total value of good\_name A, and therefore C represents 10%, the “tariff\_w” variable would instead be: 0.9(10)+0.1(6) = 9.6%. This variable is used to calculate the weighted average tariffs used in the analysis of Table 3.

**erate:** Nominal exchange rates. Calculated as annual averages of closing daily exchange rates taken from Global Financial Data.

**gdpgbp:** Nominal GDP taken from Klasing and Milionis (2014), adjusted for interwar borders using the adjustment coefficients from Broadberry and Klein (2012).

**good\_number:** A numeric variable that maps to the numeric labels of the “good\_name” variable.

**broad\_category\_number:** A numeric variable that maps to the numeric labels of the “broad\_category” variable.

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

The “data\_258” dataset is created by collapsing the “data\_847” dataset based on the “good\_name” variable. Most variables do not vary within each “good\_name” classification, so they map directly between the two datasets. The value and tariff variables represent exceptions to this. In the case of values we just sum all of the values within each “good\_name” category. For tariffs we either take the sample mean (for “tariff”) or weighted (by total value over the entire sample period) mean (for “tariff”) of the tariff variable within the “good\_name” grouping. To collapse these data we used the aggregate() function in R.