

# Overview

The code in this replication package reproduces all the results in “*Profits, Scale Economies and the Gains from Trade and Industrial Policy*” (Ahmad Lashkaripour and Volodymyr Lugovskyy).

The replication package code consists of projects: **Matlab** and **Stata**:

- (i) Project **Matlab** produces Figure 1 and Tables 1, 4, and 5. It also generates data points used to produce Figures 2 and 3 by project **Stata**. Project **Matlab** must be executed before project **Stata**. This project uses MATLAB.
- (ii) Project **Stata** produces Table 3 as well as Figures 2 and 3 from the paper. This project uses STATA.

Each project has its own set of master files. The replicator should expect the MATLAB code to run for approximately 7½ hours and the STATA code to run for approximately 1 hour on a regular desktop system.

## Data Availability and Provenance Statements

### Statement about Right

- ☒ I certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript.
- ☒ I certify that the authors of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

### Summary of Availability

- ☐ All data are publicly available.
- ☒ Some data cannot be made publicly available.
- ☐ No data can be made publicly available.

## Details on each Data Source

- **Transaction-level Colombian import data from DATAMYNE Inc**

- Stata/data/confidential\_data/datamyne/ColombiaImports2007\_1.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2007\_2.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2008.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2009\_1.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2009\_2.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2010\_1.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2010\_2.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2011\_1.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2011\_2.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2012\_1.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2012\_2.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2013\_1.dta
- Stata/data/confidential\_data/datamyne/ColombiaImports2013\_2.dta

DATAMYNE Inc provides transaction-level trade records from 43 countries including United States, Latin America, Asia and several European Union Member States (website: [datamyne.com](http://datamyne.com)). Their data is obtained from customs authorities and government agencies. Access to the transaction-level Colombian trade data was purchased from DATAMYNE Inc by Indiana University for Volodymyr Lugovskyy on May 15th, 2014. The data were available for manual online download in segments of 5,000 observations per download. For each year, Volodymyr Lugovskyy manually downloaded the data. For all years, except 2008, the annual data was split into two STATA files due to the file size constraint.<sup>1</sup> Our final dataset covers the universe of import transactions during 2007-2013 and contains more than 17 million observations. These data files cannot be shared per terms of agreement. However, the codebook for this data set is provided as part of this archive under `Stata/data/confidential_data/datamyne/list_of_variables.log`.

- **Monthly exchange rate data**

- `Stata/data/exchange_rate_report_boc/exchange_rate_report.csv`
- `Stata/data/exchange_rate_report_boc/exchange_rate_report.dta`

These datafiles report the exchange rate between different international currencies and the Colombian Peso as well as the US dollar. The datafile `exchange_rate_report.csv` contains monthly exchange rate data for various international currencies, downloaded from the Bank of Canada web portal. This data is sourced from REFINITIV (formerly THOMSON REUTERS). Around 2017, the Bank of Canada updated its exchange rate web portal following a revision to its exchange rate calculation methodology, discontinuing its publishing of daily exchange rates for a wide set of countries. The historical rates relevant to this project can be still downloaded as a legacy data file using the following link: [https://www.bankofcanada.ca/valet/observations/group/LEGACY\\_MONTHLY\\_RATES/csv](https://www.bankofcanada.ca/valet/observations/group/LEGACY_MONTHLY_RATES/csv).<sup>2</sup> The linked CSV datafile reports the end-of-month exchange rates for 68 foreign currencies quoted in Canadian dollars (CAD) from January 1999 to April 2017. The data file `exchange_rate_report.dta` is constructed by matching country names from the transaction-level import data (from DATAMYNE) with the corresponding exchange rate variable names in `exchange_rate_report.csv`. The manually constructed datafile, `cross_walk_xrate.dta`, reports the match between country names across the two datasets. The variable `xrate_usd` is calculated by dividing each exchange rate value in the raw CSV file by the value reported under `IEXM0101`, which is the US-dollar-to-CAD exchange rate. Likewise, the variable `xrate_col` is calculated by dividing each exchange rate value in the raw CSV file by the value reported under `IEXM3901`, which is the Colombian-Peso-to-CAD exchange rate.

- **Penn World Tables**

- `Stata/data/penn_world_tables/pwt90.dta`

This is a country-level database reporting national accounts data (including real GDP) between 1950 and 2017 (Feenstra, Inklaar, and Timmer, 2015; Groningen Growth and Development Centre, 2019). The datafile, `PWT 9.1`, can be downloaded via the following link: <https://www.rug.nl/ggdc/productivity/pwt/pwt-releases/pwt9.1>. A copy of this datafile is provided as part of this archive.

- **Concordance data files**

- `Stata/data/concordance/isic_hs6_WITS.csv`
- `Stata/data/concordance/isic_hs6.dta`
- `Stata/data/concordance/CORARAN-v44-oct-2017.xls`

---

<sup>1</sup> Saving original data as two separate files had no effect on the results, since, after removing redundant variables, all files were appended first into integrated annual files and then into a master file containing data for the entire 2007-2013 period.

<sup>2</sup> The above link downloads the datafile, `exchange_rate_report.csv`, under a different name, `LEGACY_MONTHLY_RATES.csv`.

- Stata/data/concordance/hs\_code\_change.dta

The concordance datafiles are used to match product codes across various classifications and years. The datafiles `isic_hs6_WITS.csv` and `isic_hs6.dta` report the concordance between 6-digit HS (Harmonized System) product codes (rev. 2007) and ISIC rev. 3 sectors. The datafile `isic_hs6_WITS.csv` is obtained from the World Bank's publicly accessible World Integrated Trade Solutions (WITS) web portal. The compressed folder containing the datafile, `isic_hs6_WITS.csv`, can be downloaded in ZIP format via the following link: [http://wits.worldbank.org/data/public/concordance/Concordance\\_H3\\_to\\_I3.zip](http://wits.worldbank.org/data/public/concordance/Concordance_H3_to_I3.zip).<sup>3</sup> The Stata file `isic_hs6.dta` is a compact version of `isic_hs6_WITS.csv` that contains 6-digit HS product codes and the corresponding ISIC rev 3 codes without the textual product description provided in the original CSV file. We use this compact datafile to partition our transaction-level import datafile for the industry-level demand estimation performed in Section 6 (results reported in Table 3). The datafiles `CORARAN-v44-oct-2017.xls` and `hs_code_change.dta` report the longitudinal concordance between HS8 product codes in Colombian trade records from DATAMYNE. These product codes were updated by the Colombian Statistical Agency (DANE) during the 2007-2013 time period. The Excel datafile `CORARAN-v44-oct-2017.xls` can be obtained from DANE's webpage via the following link: <https://www.dane.gov.co/files/sen/nomenclatura/tablasCorrelativas/CORARAN-v44-oct-2017.xls>. We used this Excel file to manually construct the STATA file `hs_code_change.dta` of matching product codes. A copy of these datafiles is provided as part of this archive.

## • World Bank Exporter Dynamics Database

- Stata/data/confidential/edd\_worldbank/CYH6\_manuf.dta

The datafile `CYH6_manuf.dta` is obtained from the Exporter Dynamics Database described in Fernandes et al. (2016). This database cannot be distributed (per terms of use) but consists of *Public Use Files (PUFs)* that are available to anyone agreeing to respect a core set of easy-to-meet conditions. The datafile `CYH6_manuf.dta` can be downloaded (in zip format) from the World Bank's micro-data web portal, <https://doi.org/10.48529/agcr-yt74>, subject to the acceptance of the terms of use.

## • World Input-Output Database (WIOD)

- Matlab/input/WIOD/WIOT2014\_Nov16\_ROW.xlsm
- Matlab/input/WIOD/WIOT2014.csv

These files contain data on sector-country flows of goods and services for the year 2014. `WIOT2014_Nov16_ROW.xlsm` is the original data file that was obtained from the World Input-Output Database 2016 Release (Timmer et al., 2015; Groningen Growth and Development Centre, 2016). This Excel file can be downloaded from <https://www.rug.nl/ggdc/valuechain/wiod/wiod-2016-release>, under "WIOT Tables Excel". Selecting "WIOT Tables Excel" downloads the full zip archive `WIOTS_in_EXCEL.zip` for the years 2000-2014. A copy of `WIOT2014_ROW_Apr12.xlsm` is provided as a part of this zip archive. The data are in the public domain and provided as part of this archive. The data in `WIOT2014.csv` is identical to the data in `WIOT2014_ROW_Apr12.xlsm` and is manually converted to the CSV format.

## • Manually constructed matrixes

- Matlab/input/Industry\_List.xlsx
- Matlab/input/Country\_List.xlsx
- Matlab/input/Agg\_S.csv

---

<sup>3</sup> The OECD web portal provides a similar concordance file in Excel format (<https://www.oecd.org/sti/ind/ConversionKeyBTDIxE4PUB.xlsx>).

- Matlab/input/Agg\_P.csv
- Matlab/input/Agg\_T.csv
- Stata/data/exchange\_rate\_report\_boc/cross\_walk\_xrate.dta

The datafiles `Industry_List.xlsx` and `Country_List.xlsx` contain the list of countries and industries used in our paper in the same order as they appear in Timmer et al. (2015) data. Both datafiles are manually constructed, and the datafile `Country_List.xlsx` is used to report the TEX file corresponding to Table 4. The datafile `AGG_S.csv` provide aggregation schemes for WIOD industries. The original Timmer et al. (2015) data has 56 industries. In our paper, inspired by Costinot and Rodriguez-Clare (2014), we aggregate them into 16 industries. The datafile `Agg_P.csv` contains the industry-level scale and trade elasticities used in our quantitative analysis under various specifications. Column 1 includes our industry-level trade elasticity estimates as reported in Table 3 of the paper. Column 2 includes our industry-level scale elasticity estimates as reported in Table 3 of the paper. Columns 3 includes trade elasticities estimated obtained with aggregate trade and tariff data, which are taken from Table 1 in Lashkaripour (2021). Column 4 includes our alternative industry-level trade elasticity estimates obtained from a two-ways fixed effects estimation, as reported in Table Q1 of the online appendix. Column 5 includes our alternative industry-level scale elasticity estimates obtained from a two-ways fixed effects estimation, as reported in Table Q1 of the online appendix. The data file `Agg_T.csv` includes applied tariff data which are constructed following the line-by-line instructions and code provided by Kucheryavyy et al. (2023). The first column corresponds to WIOD number of the country applying the tariff. The second column corresponds to the WIOD number of the country to which the tariff is applied. The third column corresponds to the industry number. And the last column corresponds to data on simple tariff line averages for the noted entry. All the manually constructed data files described above are provided as part of this archive. Lastly, the datafile `cross_walk_xrate.dta` is constructed manually by matching the country names in the Bank of Canada exchange rate report with names in the DATAMYNE datafiles. All these files are provided as part of this archive.

- The Data files in `Stata/data/internally_generated` contain intermediate output files that allow replicators to produce all the tables and figures without access to the confidential data not provided as part of this archive.

Data file	Source	Notes	Provided
Stata/data/confidential_data/... ColombiaImports2007_1.dta ColombiaImports2007_2.dta ColombiaImports2008.dta ColombiaImports2009_1.dta ColombiaImports2009_2.dta ColombiaImports2010_1.dta ColombiaImports2010_2.dta ColombiaImports2011_1.dta ColombiaImports2011_2.dta ColombiaImports2012_1.dta ColombiaImports2012_2.dta ColombiaImports2012_1.dta ColombiaImports2012_2.dta	DATAMYNE (www.datamyne.com)	Purchase required	No
Stata/data/exchange_rate_report_boc/... exchange_rate_report.csv exchange_rate_report.dta	Bank of Canada	Public domain	Yes
Stata/data/confidential/edd_worldbank/... CYH6_manuf.dta	World Bank EDD	License required	No

Stata/data/concordance/... <b>isic_hs6_WITS.csv</b> <b>isic_hs6.dta</b>	World Bank WITS	Public domain	Yes
Stata/data/concordance/... <b>hs_code_change.dta</b> <b>CORARAN-v44-oct-2017.xlsx</b>	DANE (Colombian Statistical Agency)	Public domain	Yes
Stata/data/penn_world_tables/... <b>pwt90.dta</b>	Penn World Tables v9	Public domain	Yes
Matlab/input/WIOD/... <b>WIOT2014_Nov16_ROW.xlsx</b> <b>WIOT2014.csv</b>	World Input-Output Database (WIOD)	Public domain	Yes
Matlab/input/... <b>Industry_List.xlsx</b> <b>Country_List.xlsx</b> <b>Agg_S.csv</b> <b>Agg_P.csv</b> <b>Agg_T.csv</b> Stata/data/exchange_rate_report_boc/... <b>cross_walk_xrate.dta</b>	Manually constructed from public data sources	Public domain	Yes

## Computational Requirements

### Software Requirements

- STATA (code was last run with version Stata/SE 17)
  - `colrspace` (distribution date: 2022-04-21)
  - `cleanplots` (distribution date: 2022-12-17)
  - `palettes` (distribution date: 2022-04-21)
  - `estout` (distribution date: 2023-02-12)
  - `outreg` (distribution date: 2015-09-18)
  - `ivreg2` (distribution date: 2022-05-10)
  - `gtools` (distribution date: 2022-12-05)
  - `hdfe` (distribution date: 2016-08-06)
  - `reghdfe` (distribution date: 2019-11-18)
  - `ivreghdfe` (distribution date: 2018-09-15)
  - `ftools` (distribution date: 2019-11-18)
  - `ranktest` (distribution date: 2020-09-29)
  - `winsor2` (distribution date: 2020-11-25)
  - `strdist` (distribution date: 2017-12-13)
  - `labutil` (distribution date: 2013-11-23)
  - `scheme-burd` (distribution date: 2013-03-21)
- MATLAB (code was last run with MATLAB Release 2023a)

## Memory and Runtime Requirements

- The approximate time needed to run the **Matlab** project on a regular desktop machine is 7 hours and 30 minutes.
- The approximate time needed to run the **Stata** project on a regular desktop machine is 1 hour.
- The code was last run and verified on a MacPro (late 2013) with MacOS version 12.6.5 with 3 GHz 8-Core Intel Xeon E5 and 64 GB of RAM.

## Description of Programs/Code

We begin by listing tables and figures in the paper and Online Appendix and identifying the corresponding file in the replication package that generates these results. Then, we list each file within the **Matlab** and **Stata** projects, discussing what each file does, and explaining how to replicate all the results in the paper and Online Appendix.

All the figures and tables one be generated by performing the following tasks:

1. Run the MATLAB file `Matlab/code/run_all_files.m` (make sure to set the current working directory in `run_all_files.m` to where the folder `Matlab` is located).
2. Then, run the STATA file `Stata/code/run-all-dofiles.do` (make sure to set the current working directory in `run-all-dofiles.do` to where the folder `Stata` is located).

MAIN PAPER			
Exhibit	Program	Output File	Notes
Figure 1	<code>Matlab/code/figure_1.m</code>	<code>Matlab/output/Figure_1A.png</code> <code>Matlab/output/Figure_1B.png</code>	
Table 2	<code>Matlab/code/table_2.m</code>	<code>Matlab/output/Table_2.tex</code>	
Table 3	<code>Stata/code/analysis/1-table_3.do</code>	<code>Stata/output/Table_3.tex</code>	To run all portions of this file, you will require access to the transaction-level import data from DATAMYNE (not provided as part of this archive). However, the file can generate <b>Table_3.tex</b> using the estimation results pre-saved in the folder <b>Stata/data/internally-generated</b> .
Table 4	<code>Matlab/code/table_4.m</code>	<code>Matlab/output/Table_4.tex</code>	
Table 5	<code>Matlab/code/table_5.m</code>	<code>Matlab/output/Table_5.tex</code>	
Figure 2	<code>Matlab/code/figure_2.m</code> <code>Stata/code/analysis/2-figure_2.do</code>	<code>Stata/output/Figure_2.pdf</code>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.

Figure 3	<b>Matlab/code/figure_3.m</b> <b>Stata/code/analysis/3-figure_3.do</b>	<b>Stata/output/Figure_3A.pdf</b> <b>Stata/output/Figure_3B.pdf</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.
<b>ONLINE APPENDIX</b>			
<b>Exhibit</b>	<b>Program</b>	<b>Output File</b>	<b>Notes</b>
Figure E.1	<b>Matlab/code/figure_E1.m</b>	<b>Matlab/output/Figure_E1.png</b>	
Figure H.1	<b>Matlab/code/apppnedix_H.m</b> <b>Stata/code/analysis/4-figure_H1.do</b>	<b>Stata/output/Figure_H1.png</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.
Figure H.2	<b>Matlab/code/apppnedix_H.m</b> <b>Stata/code/analysis/5-figure_H2.do</b>	<b>Stata/output/Figure_H2.png</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.
Figure H.3	<b>Matlab/code/apppnedix_H.m</b> <b>Stata/code/analysis/6-figure_H3.do</b>	<b>Stata/output/Figure_H3.png</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.
Table N.1	<b>Stata/code/analysis/7-table_N1.do</b>	<b>Stata/output/Table_N1.tex</b>	To run all portions of this file, you will require access to the transaction-level import data from DATAMYNE (not provided as part of this archive). However, the file can generate <b>Table_N1.tex</b> using the estimation results pre-saved in the folder <b>Stata/data/internally-generated</b> .
Figure O.1	<b>Stata/code/analysis/8-figure_O1.do</b>	<b>Stata/output/Figure_O1_a.png</b> <b>Stata/output/Figure_O1_b.png</b>	To run all portions of this file, you will require access to the transaction-level import data from DATAMYNE (not provided as part of this archive). However, the file can generate <b>Figure_O1_a.png</b> and <b>Figure_O1_b.png</b> using the estimation results pre-saved in the folder <b>Stata/data/internally-generated</b> .
Figure P.1	<b>Stata/code/analysis/9-figure_P1.do</b>	<b>Stata/output/Figure_P1_A.png</b> <b>Stata/output/Figure_P1_B.png</b> <b>Stata/output/Figure_P1_C.png</b>	To run all portions of this file, you will require access to the transaction-level import data from DATAMYNE (not provided as part of this archive). However, the file can generate <b>Figure_P1_A.png</b> ,

			<b>Figure_P1_B.png</b> , and <b>Figure_P1_C.png</b> using the estimation results pre-saved in the folder <b>Stata/data/internally-generated</b> .
Table Q.1	<b>Stata/code/analysis/10-table_Q1.do</b>	<b>Stata/output/Table_Q1.tex</b>	To run all portions of this file, you will require access to the transaction-level import data from DATAMYNE (not provided as part of this archive). However, the file can generate <b>Table_3.tex</b> using the estimation results pre-saved in the folder <b>Stata/data/internally-generated</b> .
Table R.1	<b>Stata/code/analysis/11-table_R1.do</b>	<b>Stata/output/Table_R1.tex</b>	To run all portions of this file, you will require access to the Exporter Dynamics Database (not provided as part of this archive). However, the file can generate <b>Table_R1.tex</b> using the estimation results pre-saved in the folder <b>Stata/data/internally-generated</b> .
Table S.1	<b>Stata/code/analysis/12-table_S1.do</b>	<b>Stata/output/Table_S1.tex</b>	To run all portions of this file, you will require access to the transaction-level import data from DATAMYNE (not provided as part of this archive). However, the file can generate <b>Table_S1.tex</b> using the estimation results pre-saved in the folder <b>Stata/data/internally-generated</b> .
Table V.1	<b>Matlab/code/table_V1.m</b>	<b>Matlab/output/Table_V1.tex</b>	
Figure W.1	<b>Matlab/code/apppnedix_W.m</b> <b>Stata/code/analysis/14-figure_W1.do</b>	<b>Stata/output/Figure_W1.eps</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.
Figure W.2	<b>Matlab/code/apppnedix_W.m</b> <b>Stata/code/analysis/15-figure_W2.do</b>	<b>Stata/output/Figure_W2.eps</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.
Figure X.1	<b>Matlab/code/figure_X1.m</b>	<b>Matlab/output/Figure_X1.png</b>	
Table Y.1	<b>Matlab/code/apppnedix_Y.m</b> <b>Stata/code/analysis/16-figure_Y1.do</b>	<b>Stata/output/Figure_Y1_a.png</b> <b>Stata/output/Figure_Y1_b.png</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.



Table Y.2	<b>Matlab/code/apppnedix_Y.m</b> <b>Stata/code/analysis/17-figure_Y2.do</b>	<b>Stata/output/Figure_Y2_a.png</b> <b>Stata/output/Figure_Y2_b.png</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.
Table Y.3	<b>Matlab/code/apppnedix_Y.m</b> <b>Stata/code/analysis/17-figure_Y3.do</b>	<b>Stata/output/Figure_Y3_a.png</b> <b>Stata/output/Figure_Y3_b.png</b>	The MATLAB file generates and saves the output data, which is then used by the Stata file to plot the figure in pdf format.
Figure Z.1	<b>Matlab/code/figure_Z1.m</b>	<b>Matlab/output/Figure_Z1A.eps</b> <b>Matlab/output/Figure_Z1B.eps</b>	

## Details for the Project **Matlab**

Running the following MATLAB file generates Figures 1, E1, and Z1 and Tables 2, 4, 5 and V1:

- **Matlab/code/run\_all\_files.m**

The same file generates the necessary data points to create Figures 2, 3, H1-H3, W1-W2, and Y1-Y3 in STATA. See also the following readme file for a full description of the MATLAB code:

- **Matlab/code/README.txt**

All the data needed to run the **Matlab** project are provided as part of this archive. To replicate results for each figure or table separately, run the following MATLAB files in any order desired.

<b>Matlab/code/figure_1.m</b>	Replicates Figure 1 in Section IV of the paper
<b>Matlab/code/table_2.m</b>	Replicates Table 2 in Section IV of the paper
<b>Matlab/code/table_4.m</b>	Replicates Table 4 in Section VII of the paper
<b>Matlab/code/table_5.m</b>	Replicates Table 5 in Section VII of the paper
<b>Matlab/code/figure_2.m</b>	Creates data points for Figure 2 in Section VII the paper (figure generated in STATA)
<b>Matlab/code/figure_3.m</b>	Creates data points for Figure 3 in Section VII the paper (figure generated in STATA)
<b>Matlab/code/figure_E1.m</b>	Replicates Figure E1 of the Online Appendix E
<b>Matlab/code/appendix_H.m</b>	Creates data points for Figures H1-H3 of the Online Appendix H (figures generated in STATA)
<b>Matlab/code/table_V1.m</b>	Replicates Table V1 of the Online Appendix V
<b>Matlab/code/appendix_W.m</b>	Creates data points for Figures W1-W2 of the Online Appendix W (figures generated in STATA)
<b>Matlab/code/figure_X1.m</b>	Replicates Figure X1 of the Online Appendix X
<b>Matlab/code/appendix_Y.m</b>	Creates data points for Figures Y1-Y3 of the Online Appendix Y (figures generated in STATA)
<b>Matlab/code/figure_Z1.m</b>	Replicates Figure Z1 in the Online Appendix Z

## Details for the Project **Stata**

Running the following STATA file generates Tables 3, N1, Q1, R1, and S1 as well as Figures 2, 3, H1-H2, O1, P1, S1, W1-W2, and Y1-Y3 in the paper and Online Appendix:

- **Stata/code/run-all-dofiles.do**

See also the following readme file for a full description of the MATLAB code:

- **Stata/code/README.txt**

Some of the data used by the **Stata** project is confidential and not provided as part of this archive. However, the necessary output to generate all the tables and figures (without data access) is saved in the folder `Stata/data/internally_generated` and provided as part of this archive. To replicate individual figures or tables separately, individual STATA do-files should be run in the order listed below after executing the **Matlab** project.

<code>Stata/code/analysis/1-table_3.do</code>	Replicates Table 3 in Section VI of the paper
<code>Stata/code/analysis/2-figure_2.do</code>	Replicates Figure 2 in Section VII of the paper
<code>Stata/code/analysis/3-figure_3.do</code>	Replicates Figure 3 in Section VII of the paper
<code>Stata/code/analysis/4-figure_H1.do</code>	Replicates Figure H1 of the Online Appendix H
<code>Stata/code/analysis/5-figure_H2.do</code>	Replicates Figure H2 of the Online Appendix H
<code>Stata/code/analysis/6-figure_H3.do</code>	Replicates Figure H3 of the Online Appendix H
<code>Stata/code/analysis/7-table_N1.do</code>	Replicates Table N1 of the Online Appendix N
<code>Stata/code/analysis/8-figure_O1.do</code>	Replicates Figure O1 of the Online Appendix O
<code>Stata/code/analysis/9-figure_P1.do</code>	Replicates Figure P1 of the Online Appendix P
<code>Stata/code/analysis/10-table_Q1.do</code>	Replicates Table Q1 of the Online Appendix Q
<code>Stata/code/analysis/11-table_R1.do</code>	Replicates Table R1 of the Online Appendix R
<code>Stata/code/analysis/12-table_S1.do</code>	Replicates Table S1 of the Online Appendix S
<code>Stata/code/analysis/13-figure_S1.do</code>	Replicates Figure S1 of the Online Appendix S
<code>Stata/code/analysis/14-figure_W1.do</code>	Replicates Figure W1 of the Online Appendix W
<code>Stata/code/analysis/15-figure_W2.do</code>	Replicates Figure W2 of the Online Appendix W
<code>Stata/code/analysis/16-figure_Y1.do</code>	Replicates Figure Y1 of the Online Appendix Y
<code>Stata/code/analysis/17-figure_Y2.do</code>	Replicates Figure Y2 of the Online Appendix Y
<code>Stata/code/analysis/18-figure_Y3.do</code>	Replicates Figure Y3 of the Online Appendix Y

## Instructions to Replicators

To run the **Matlab** project, execute the master m-file `Matlab/code/run_all_files.m`

- before executing the master m-file, edit **line 10** to set your local path to where the folder ‘Replication/Matlab’ is located in your computer.

To run the **Stata** project, execute the master do-file `Stata/code/run-all-dofiles.do`

- before executing the master do-file, edit **line 12** to set your local path to where the folder ‘Replication/Stata’ is located in your computer.
- If you do not have access the confidential trade data from DATAMYNE, assign to the global macro `access_to_datamyne` the value `no` instead of `yes`. In that case, you can still replicate all the tables and figures using the intermediate output files pre-saved in the folder `Stata/data/internally-generated`.
- If you do not have access the confidential data from the Exporter Dynamics Database, assign to the global macro `access_to_edd` the value `no` instead of `yes`. In that case, you can still replicate all the tables and figures using the intermediate output files pre-saved in the folder `Stata/data/internally-generated`.

## References

Bank of Canada. 2017. “Exchange Rates: Historical noon and closing rates,”  
<https://www.bankofcanada.ca/rates/exchange/legacy-noon-and-closing-rates/>

DANE (Colombian Statistical Agency). 2017. "Economics and Social Correlative Tables," <https://www.dane.gov.co/files/sen/nomenclatura/tablasCorrelativas/CORARAN-v44-oct-2017.xls>

Costinot, A., and A. Rodríguez-Clare. 2014. "Trade Theory with Numbers: Quantifying the Consequences of Globalization." In *Handbook of International Economics*. Vol. 4, ed. Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, Chapter 4, 197–261. Elsevier.

EDD by World Bank. 2016. "Exporter Dynamics Database version 2.0, 1997-2014," <https://doi.org/10.48529/agcr-yt74>

Feenstra, R., R., Inklaar and M. Timmer (2022), "Penn World Table version 9.1". Groningen Growth and Development Centre [distributor]. <https://www.rug.nl/ggdc/productivity/pwt/?lang=en> (accessed October 6<sup>th</sup> of 2022)

Fernandes, A., Freund, C., and M. Pierola. 2016. "Data for: Exporter Behavior, country Size and Stage of Development: Evidence from the Exporter Dynamics Database," *Journal of Development Economics*, 119, 121–137. <https://doi.org/10.1016/j.jdeveco.2015.10.007>

"International Database" by DATAMYNE Inc. <https://www.datamyne.com> (accessed between May 8<sup>th</sup> of 2014 and November 8<sup>th</sup> of 2016)

Kucheryavyy, K., G., Lyn, and A. Rodríguez-Clare. 2023. "Replication data for: Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale." *American Economic Journal: Macroeconomics* [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.1257/mac.20190156>

Lashkaripour, A., 2021. "Data for: Can Trade Taxes be a Major Source of Government Revenue?" *Journal of the European Economic Association*, 19 (5): 2399–2428. <https://doi.org/10.1093/jeea/jvaa058>

Timmer, M., E. Dietzenbacher, B. Los, R. Stehrer, and G. de Vries. 2015. "An Illustrated User Guide to the World Input-Output Database: The Case of Global Automotive Production." *Review of International Economics*, 23(3): 575– 605.

WIOD, 2021, "World Input-Output Database 2016 Release, 2000-2014," Groningen Growth and Development Centre [distributor]. <https://doi.org/10.34894/PJ2M1C>, DataverseNL, V2

World Integrated Trade Solutions (WITS) by the World Bank. "Product Concordance," [https://wits.worldbank.org/product\\_concordance.html](https://wits.worldbank.org/product_concordance.html)