

README

Data and Code for “Trade Diversion and Jobs: Evidence from Mexican Municipalities”

by Carlos Rodríguez-Castelán, Emmanuel Vazquez and
Hernan Winkler¹
(May 2026)

Overview

This document provides information needed to replicate the results in “Trade Diversion and Jobs: Evidence from Mexican Municipalities”.

The folders included in this replication package are the following:

- **Data**
- **Do_files**
- **Outputs**

Data, Do_files, and Outputs should be in the same folder (e.g. the folder they were archived in) to run the project do-files below.

Computational requirements

Software Requirements: Stata version 14 or higher is required. The guidelines below, sourced from <https://www.stata.com/products/compatible-operatingsystems/>, outline the following computational requirements:

Platforms

Stata for Windows®

- Windows 11*
- Windows 10 *
- Windows Server 2022, 2019, 2016, 2012R2 *

* Stata requires 64-bit Windows for x86-64 processors made by Intel® or AMD (Core i3 equivalent or better)

[Find out if your OS is 64-bit compliant.](#)

Stata for Mac®

- Mac with Apple Silicon or Intel processors
- macOS 11.0 (Big Sur) or newer for Macs with Apple Silicon and macOS 10.13 (High Sierra) or newer for Macs with Intel processors

Stata for Linux

- Any 64-bit (Core i3 equivalent or better) running Linux
- Minimum requirements include the GNU C library (glibc) 2.17 or better and libcurl4
 - Check the output of **ldd -v** within a terminal
- For xstata, you need to have GTK 2.24 installed

Hardware requirements

Package	Memory	Disk space
Stata/MP	4 GB	2 GB
Stata/SE	2 GB	2 GB
Stata/BE	1 GB	2 GB

Stata for Linux requires a video card that can display thousands of colors or more (16-bit or 24-bit color)

¹ Rodríguez-Castelán and Winkler are with the World Bank, and Vazquez is with CEDLAS (IIE, FCE) – Universidad Nacional de La Plata. Corresponding author for the reproducibility package: Emmanuel Vazquez (evazquez@cedlas.org).

Packages needed for tables and figures: carryforward, gtools/fcollapse utilities, outreg2, shp2dta, spmap, and xtevent. All user-written programs required by the code are included in the ado folder (Do_files/ado), including auxiliary routines used by spmap and xtevent.

Memory and Runtime Requirements

The analyses can be run on a standard desktop computer. The replicator should expect the code to run for about ~10 minutes.

Folders and Contents

A. Data

The Data folder contains the following datasets, organized in subfolders according to the underlying source of data:

- **Subfolder “fajgelbaum”**
 - o “m_flow_hs10_fm.dta”: The tariff database from the replication package of the paper Fajgelbaum, P., P. Goldberg, P. Kennedy, A. Khandelwal, and D. Taglioni (2024). "The US-China Trade War and Global Reallocations." *American Economic Review: Insights* 6 (2): 295–312. Retrieved on 06/03/24 from <https://akhandelwal8.github.io/research/> by selecting “Replication” under the corresponding paper.
- **Subfolder “datamexico”**
 - o “Exports to US by product and municipality from 2006 to 2024.csv”: Data Mexico / Secretaría de Economía customs records from the INEGI-Balanza Comercial de Mercancías de México (BCMM). The file contains Mexican export values in U.S. dollars by year, municipality, HS product, and destination country, filtered to exports to the United States. Downloaded from the Data Mexico platform on 05/17/2024 using the Viz Builder tool (<https://www.economia.gob.mx/datamexico/en/vizbuilder>) and selecting “Foreign trade (BCMM) by municipality.”
 - o “Total Exports by product and municipality from 2006 to 2024.csv”: Data Mexico / Secretaría de Economía customs records from INEGI-BCMM. The file contains Mexican export values in U.S. dollars by year, municipality, and HS product for all destination countries. Downloaded from the Data Mexico platform on 05/17/2024 using the Viz Builder tool (<https://www.economia.gob.mx/datamexico/en/vizbuilder>) and selecting “Foreign trade (BCMM) by municipality.”
 - o “Total Exports by municipality from 2006 to 2024.csv”: Data Mexico / Secretaría de Economía customs records from INEGI-BCMM. The file contains annual municipal export totals in U.S. dollars across products and destination countries. Downloaded from the Data Mexico platform on 06/13/2024 using the Viz Builder tool (<https://www.economia.gob.mx/datamexico/en/vizbuilder>) and selecting “Foreign trade (BCMM) by municipality.”
 - o “Imports from CH by product and municipality from 2006 to 2024.csv”: Data Mexico / Secretaría de Economía customs records from INEGI-BCMM. The file contains Mexican import values in U.S. dollars by year, municipality, HS product, and origin country, filtered to imports from China. Downloaded from the Data Mexico platform on 11/29/2024 using the Viz

Builder tool (<https://www.economia.gob.mx/datamexico/en/vizbuilder>) and selecting “Foreign trade (BCMM) by municipality.”

- Subfolder “castelan et al - 2024”

- o “castelan_vazquez_winkler_2024.dta”: Reduced municipality-year extract from the database used in Rodríguez-Castelán, Vazquez, and Winkler (2024), “Exports and Income Inequality in Mexico,” *Journal of Development Studies* 60(7): 1131-1159, DOI: 10.1080/00220388.2024.2337373. The extract was shared by the authors on 03/02/2025 and contains the municipality identifier, year, total population, rural population share, literacy among the population aged 15 or older, and log median per capita income. Total population, rural population share, and literacy were constructed from Mexican population censuses and population counts. Total population corresponds to the municipal population count; rural population share measures the share of municipal residents living in rural localities; and literacy measures the share of the population aged 15 or older that can read and write. Log median per capita income corresponds to the logarithm of municipal median per capita income in real 2014 Mexican pesos. This income measure was generated using the poverty mapping methodology of Elbers, Lanjouw, and Lanjouw (2003), as implemented for Mexico in Enamorado, López-Calva, Rodríguez-Castelán, and Winkler (2016). Nominal incomes were deflated using the average National Consumer Price Index from INEGI, with December 2010 as the base period. For additional details on the underlying database and the construction of the income measures, see Enamorado, López-Calva, Rodríguez-Castelán, and Winkler (2016), as referenced in Rodríguez-Castelán, Vazquez, and Winkler (2024).

- Subfolder “ipums”

- o “ipums_original.dta”: Customized IPUMS International extract for Mexico containing harmonized person-level microdata from the 2010 Population and Housing Census, the 2015 Intercensal Survey, and the 2020 Population and Housing Census. Source data are from INEGI and harmonized/distributed by IPUMS International. Retrieved on 01/07/2025 from <https://international.ipums.org/international-action/variables/group> through the IPUMS extract system by selecting the Mexico 2010, 2015, and 2020 samples and the following variables: From Harmonized variables - Household - Technical household: COUNTRY, YEAR, SAMPLE, SERIAL and HHWT; From Harmonized variables - Household – Geography (F-N): GEO1_MX, GEO2_MX, GEO2_MX2010, GEO2_MX2015 and GEO2_MX2020; From Harmonized variables - Household - Household Economic: REMITT; From Harmonized variables - Person - Technical Person: PERNUM and PERWT; From Harmonized variables - Person – Demographic: AGE and SEX; From Harmonized variables - Person – Education: EDATTAIN; From Harmonized variables - Person – Work: EMPSTAT, CLASSWK and HRSWORK1; From Harmonized variables - Person - Occupation, Industry: OCC and INDGEN; From Harmonized variables - Person - Income: INCEARN; From Harmonized variables - Person - Migration: MIG1_5_MX and MIG2_5_MX; From Source variables – Person – Other Person: MX2015A_HLTHBENS, MX2015A_RETIRE, MX2020A_HLTHBENS and MX2020A_RETIRE.

- Subfolder “covid”

- o “Casos Diarios Municipio Confirmados 20230625.csv”: Municipality-level daily confirmed COVID-19 cases from Dirección General de Epidemiología, Secretaría de Salud, Gobierno de México. Retrieved on 09/27/2024 from <https://datos.covid-19.conacyt.mx/#DownZCSV> by selecting “Casos diarios por Municipio-Confirmados-Descargar”

- o “Casos Diarios Municipio Defunciones 20230625.csv”: Municipality-level daily COVID-19

deaths from Dirección General de Epidemiología, Secretaría de Salud, Gobierno de México. Retrieved on 09/27/2024 from <https://datos.covid-19.conacyt.mx/#DownZCSV> by selecting “Casos diarios por Municipio-Defunciones-Descargar”.

- Subfolder “censos economicos”

- o “SAIC Exporta 2025828 114239722.csv”: Extract from the INEGI Sistema Automatizado de Información Censal (SAIC), based on the Economic Censuses. The file contains municipality-level totals by census year and economic activity, including employed personnel, value added, fixed assets, and other economic variables. Retrieved on 08/28/2025 from the INEGI SAIC Interactive Tabulation System (<https://www.inegi.org.mx/app/saic/default.html>). The extract was generated by selecting the 2003, 2008, 2013, 2018, and 2023 census years; manually selecting all municipalities within each state from the geographic area menu; selecting “all sectors” and “with totals” from the economic activity menu; and selecting all available variables from the census variable menu. The “Suma de Estratos” option was selected in the employed personnel stratum submenu. The resulting query was then run by clicking “Consult” and exported as a CSV file using the “Export as CSV” option.

- Subfolder “shapes”

- o “areas_geoestadisticas_municipales.*”: INEGI municipal geographic boundary shapefiles from the Marco Geoestadístico, September 2019. The files contain municipal geo-statistical polygons and identifiers used to draw the sample map. The original source is INEGI’s Marco Geoestadístico webpage (<https://www.inegi.org.mx/temas/mg/>). The local shapefile components are timestamped 03/30/2019.

- o “coords.dta” and “data.dta”: Stata map files generated from “areas_geoestadisticas_municipales.*” with shp2dta/spmap. These files contain the polygon coordinates and associated municipal identifiers used for Appendix Figure A2.

- Subfolder “own production”

- o Intermediate and analysis-ready Stata datasets produced by the do-files from the raw inputs above. These files are regenerated by running 0_master.do and are included to make the replication workflow transparent and inspectable.

B. Do_files

Contains Stata .do files that replicate the data analysis (main paper and appendix).

“0_master.do”: **To run all the analysis, open this do file and set the path in line 29 to point to the folder on your computer** that contains this archive. Running this .do file calls all the .do files needed to create the tables and figures in the paper, which are the following:

- “1_process_ustariffs_chn.do”: Build product-year US tariffs on China from Fajgelbaum et al. (2024).
- “2_process_ustariffs_mex.do”: Build product-year US tariffs on Mexico from Fajgelbaum et al. (2024).
- “3_estimates_exposure.do”: Build municipal exposure measures using baseline exports and tariff schedules.
- “4_estimates_exposure_cdm.do”: Build City of Mexico aggregate exposure measures.
- “5_muni_trade.do”: Build municipal and product-level trade databases used in the paper.
- “6_process_ipums.do”: Process IPUMS census data into municipality-period labor-market

measures.

- “7_process_covid.do”: Process COVID-19 controls at municipality-period level.
- “8_process_ce.do”: Process Economic Census data at municipality-period level.
- “9_analysis_product_level.do”: Produce product-level Table 1 and Figure 2 data.
- “10_analysis_municipality_level.do”: Produce municipality-level tables and Figure 3.
- “11_maps.do”: Produce Appendix Figure A2 map.

C. Outputs

The .do files export tables and figures to the *Outputs* folder. This folder contains .xlsx files reproducing all tables in the paper generated as part of the data analysis, as well as files reproducing figures either in image format (.png) or, in specific cases, in spreadsheet format containing the underlying data. The name of each exported output file matches the corresponding table or figure in the paper.

Some figures in the paper are classified as non-analytical outputs. These include items that report data directly from cited sources or involve simple arithmetic operations following procedures explicitly described in the table notes. For these items, no replication code is provided.

The tables below list all tables and figures in the paper and indicate, for each item, whether it is reproduced by code and, when applicable, the corresponding program and lines where it is generated.

C.1. Tables

Table #	Program	Lines
Main Tables		
1	9_analysis_product_level.do	46-116
2	10_analysis_municipality_level.do	297-316
3	10_analysis_municipality_level.do	317-353
4	10_analysis_municipality_level.do	355-385
5	10_analysis_municipality_level.do	386-405
6	10_analysis_municipality_level.do	407-426
7	10_analysis_municipality_level.do	427-484
Appendix		
A1	10_analysis_municipality_level.do	486-515
A2	10_analysis_municipality_level.do	517-583
A3	10_analysis_municipality_level.do	584-603
A4	10_analysis_municipality_level.do	604-638
A5	10_analysis_municipality_level.do	639-666

C.2. Figures

Figure #	Program	Lines
Main Figures		
1	Non-analytical output	
2	9_analysis_product_level.do	21-43
3	10_analysis_municipality_level.do	668-719
Appendix		
A1	Non-analytical output	
A2	11_maps.do	36-41

References

Enamorado, T., López-Calva, L. F., Rodríguez-Castelán, C., & Winkler, H. (2016). Income inequality and violent crime: Evidence from Mexico's drug war. *Journal of Development Economics*, 120(May), 128–143. doi:10.1016/j.jdeveco.2015.12.004.

Fajgelbaum, P., Goldberg, P., Kennedy, P., Khandelwal, A. & Taglioni, D. (2024). The US-China Trade War and Global Reallocations. *American Economic Review: Insights*, 6 (2), 295–312. doi: 10.1257/aeri.20230094.

Rodríguez-Castelán, C., Vazquez, E. and Winkler, H. (2026) "Trade Diversion and Jobs: Evidence from Mexican Municipalities". World Bank Policy Research Working Paper.

Rodríguez-Castelán, C., Vazquez, E. and Winkler, H. (2024), "Exports and Income Inequality in Mexico," *Journal of Development Studies* 60 (7): 1131-1159, DOI: 10.1080/00220388.2024.2337373.

Data Availability Statement

“Trade Diversion and Jobs: Evidence from Mexican Municipalities”

All files included in the Data folder can be redistributed as part of this replication package. The Fajgelbaum et al. tariff data, Data Mexico foreign-trade files, Mexican COVID-19 files, INEGI SAIC extract, and INEGI municipal shapefiles are publicly available from the sources listed above. The IPUMS file is a public-use custom extract obtained from IPUMS International, which requires users to comply with IPUMS terms of use. The `castelan_vazquez_winkler_2024.dta` file is an author-prepared extract from Rodriguez-Castelan, Vazquez, and Winkler (2024) and can be reshared with this package. Files in the "own production" subfolder are derived datasets generated by the do-files from the raw inputs documented in this section.