1. General Description

This folder contains the data and codes to replicate the results in the paper "**Trade with Nominal Rigidities: Understanding the Unemployment and Welfare Effects of the China Shock**", by Andrés Rodríguez-Clare, Mauricio Ulate, and Jose P. Vasquez, to be published at the *Journal of Political Economy*. This README is associated with the following data repository:

Rodriguez-Clare, Andres; Ulate, Mauricio; Vasquez, Jose P., 2025, "Replication Data for: "Trade with Nominal Rigidities: Understanding the Unemployment and Welfare Effects of the China Shock"", Harvard Dataverse, https://doi.org/10.7910/DVN/R96X0Z.

The replication package contains three folders: "Empirics", "DataCalibration" and "Model". The folder "Empirics" performs the empirical analyzes discussed in Section 2 of the paper. The folder "DataCalibration" performs the construction process for the inputs required in order to run the model. Finally, the folder "Model" contains the Matlab codes to run the algorithms to solve the model and produce the theoretical figures and tables in the paper. Below, we have a section carefully explaining the purpose, structure, inputs, and outputs of each of the main three folders, and how each folder fits into the bigger picture of the paper.

1.1. Computational requirements

The overall time required to run the code is between 1 and 6 days depending on the computer specifications of the replicator. A single-core computer should suffice (i.e., no cluster is required), since we almost never parallelize computations. The only exception is the file "NEST_Main2.m" in the "Model" folder, for which the baseline code uses 44 cores, but we explain below how this can be changed to use less cores (going as low as a single core) in Section 4 of this readme. The memory and storage requirements should not be out of the ordinary, so we are not detailing anything special here. However, we specify that the computing environment where most of the analyses were run is as follows: "11th Gen Intel(R) Core(TM) i7-1195G7 @ 2.90GHz 1.80 GHz".

1.2. Data Availability Statement

We state that we had legitimate access to the data mentioned below for our research and that we have the right to redistribute the data that is included in this replication package.

2. Empirical Analysis ("Empirics" Folder)

2.1. Stata

All the codes in this folder are written and implemented using the program Stata. Stata is proprietary statistical software available for purchase from the following webpage:

https://www.stata.com/

We ran our codes in Stata 18 MP, but older versions of the software should work similarly well.

Note on the command "cd .." at the start of the main code file: Our "0-MASTER" .do file in this project begins with the command "cd ..", which moves the current working directory up one level from the file's

location. This ensures that relative file paths in the code work as intended regardless of where the entire replication folder is located in a given computer. For consistent replication, please place the files in a directory structure that matches the project's organization (i.e., one main folder which contains three subfolders labeled: "codes", "figures", "raw_data").

2.2. Codes Folder

The "codes" folder contains seven scripts. The file "0-MASTER.do" is a central code that runs all other scripts and generates all the exhibits. This is a description of the other six scripts:

- 1-ipums_acs.do: Creates a panel of employment indicators by CZ from 2006 to 2020 following the data construction from ADH (2013).
- 2-unemp_pop.do: Creates a panel of employment indicators by CZ from 2006 to 2020. Using LAUS and SEER data.
- 3-coefs_graphs_decadal.do: Runs regressions in the spirit of ADH 21 for 2006-2020 but using the data construction and exposure measures from ADH 13. Produces the main Figures 1 and 6, and Appendix Figure A1.
- 4-cps1990_rigmeasures.do: Creates the downward nominal wage rigidity measures.
- 5-dnwr_figures.do: Produces the main Figure 2, and Appendix Figures A2, A3, A4, A5, A6, A7, A8.
- subfile_ind1990dd.do: This do file was taken from Prof. David Dorn's data webpage (https://www.ddorn.net/data.htm). It recodes the ind1990 variable into ind1990dd. This crosswalk code helps to replicate the employment classification of manufacturing and non-manufacturing.

The following table provides a concordance between the codes ran by the "master" file, their output, and the corresponding paper exhibit:

Program	Output	Paper Figure/Table #
1-ipums_acs.do	temp/workfile_china_RUV.dta	
2-unemp_pop.do	temp/unemp_pop.dta	
3-coefs_graphs_decadal.do:	temp/models_coefs.dta	Figure_1A Figure_1B Figure_1C Figure_6A Figure_6B Figure_6C Figure_6D Figure_A1
4-cps1990_rigmeasures.do:	temp/Jo_state_level_dnwr_proc.dta temp/cps1990_rigmeasures.dta temp/cps_rigmeasures.dta	
5-dnwr_figures.do:	temp/state_workfile_china	Figure_2A Figure_2B Figure_A2A Figure_A2B Figure_A3A Figure_A3B Figure_A4A Figure_A4B Figure_A5A Figure_A5B Figure_A6A Figure_A6B Figure_A7A Figure_A7B Figure_A8A

		Figure_A8B
subfile_ind1990dd.do		

2.3. Raw Data Folder

The “raw_data” folder stores our main datasets. These files are publicly available and can be downloaded following the instructions in this Readme.

2.3.1 Files taken from Prof. Dorn's webpage

Files: raw_data\workfile_china.dta, raw_data\cw_puma2000_czone.dta, raw_data\cw_puma2010_czone.dta, and raw_data\cw_cty_czone.dta.

We take those four files from Prof. David Dorn's webpage.

Script 1 uses raw_data\workfile_china.dta, which is taken from the replication package of David Autor, David Dorn, and Gordon Hanson. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." American Economic Review, 103(6), 2121-2168, 2013. It contains the data to replicate the main results of their paper (except for Tables 1 and 2). To get it, one has to download [P3] and go to the dta folder.

Script 1 also uses David Dorn's crosswalk files from PUMAs to CZ. The files are named raw_data\cw_puma2000_czone.dta and raw_data\cw_puma2010_czone.dta. To download them, one must download them from [E5] and [E6], respectively.

Finally, in script 2, to go from the county level to the CZ level, we follow David Autor and David Dorn. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." American Economic Review, 103(5), 1553-1597, 2013. So we have to download [E8] to get raw_data\cw_cty_czone.dta.

2.3.2. ACS data

File: raw_data/ipums_2005_2021.dta,

Script 1 uses data from the American Community Surveys (ACS) from 2005 to 2021: raw_data/ipums_2005_2021.dta. You can download the file from the IPUMS (IPUMS, 2025) website (here) following the query in raw_data/ipums_2005_2021_query.txt. We pooled the 2005-2021 ACS 1-year samples for people between 16 and 64 years of age.

2.3.3. SEER data

Files: raw_data\us.1969_2022.singleages.adjusted.txt.

To construct our dataset for 1990-2020 unemployment and population at the CZ level, script 2 takes the population between 15 and 65 years old for each county from the Surveillance, Epidemiology, and End Results (SEER) Program (Surveillance, Epidemiology, and End Results (SEER) Program, 2025). We select County-level Population Files -Single-year Age Groups and download the .gz file for "1969-2022 White, Black, Other" (specifically, the "All States Combined" file). That is how we get raw_data\us.1969_2022.singleages.adjusted.txt.

2.3.4. LAUS data

Files: Those in the raw_data\LAU folder

To construct our dataset for 1990-2020 unemployment and population at the CZ level, Script 2 takes the unemployment data for each county from the Local Area Unemployment Statistics (LAUS) (Local Area Unemployment Statistics (LAUS), 2025). We select all the counties for each state (after choosing Counties and equivalents) and save them in .xlsx files named after each state. Since you can only download 200 counties per selection, we had to save Texas counties in Texas1.xlsx and Texas2.xlsx. Those files are stored in our raw_data\LAU folder.

2.3.5. CPS data

Files: raw_data/cps_86_90.dta, raw_data/morg86.dta, raw_data/morg90.dta.

Script 4 uses monthly labor force micro-data from the U.S. Current Population Survey (CPS) for 1986-1990. The file's name is raw_data/cps_86_90.dta, and we downloaded it from the IPUMS CPS (IPUMS, 2025) webpage following the query in raw_data/cps86_90_query.txt. For the data processing procedures, we followed Yoon J. Jo, 2024, "Beyond the Spike at Zero: Understanding Nominal Wage Rigidity through Empirical and Model-Based Approaches," forthcoming in the Journal of International Economics.

Script 4 also uses the CPS Merged Outgoing Rotation Group Earnings (NBER CPS, 2025) data for 1986-1990. We download it from NBER. Specifically, from the Stata dta files link, we take our raw_data/morg86.dta - raw_data/morg90.dta files. These files are merged into the CPS to obtain the hourly wage allocation flag variable.

2.3.6. Data provided by Dr. Yoon Joo Jo

File: raw_data\Jo_state_level_dnwr.dta

Script 4 uses raw_data\Jo_state_level_dnwr.dta for rigidity measures. This is a dataset provided by Dr. Yoon Joo Jo from her paper: Yoon J. Jo, 2024. "Beyond the Spike at Zero: Understanding Nominal Wage Rigidity through Empirical and Model-Based Approaches," forthcoming in the Journal of International Economics. The data is publicly available on her website.

2.3.7. Data from the model simulations

File: raw_data/ToPlotwithADH2021.xlsx

Script 3 uses raw_data/ToPlotwithADH2021.xlsx. This file is an output from the model simulations, given the calibration of the different "long" vs. "short" shocks. Section 4 of this readme explains how to obtain these model outputs, which are produced by “NEST_ProduceFigure6ModelPart”.

2.3.8. Note: Large files

Three raw data files are larger than 150 MB. We already mentioned them above in their relevant subsection, but we explained them here again. The files are as follows:

1. For *codes/1-ipums_acs.do*, the raw_data/ipums_2005_2021.dta. However, a codebook file is available in raw_data/ipums_2005_2021_query.txt. This document contains the query requested from the IPUMS USA (IPUMS, 2025) webpage. A log-in is necessary to download the data. After an account is created, data can be downloaded from the "Select Data"->"Select samples" window. Tick on all ACS files (we use the 1% sample).
2. For *codes/2-unemp_pop.do*, the raw_data/us.1969_2022.singleages.adjusted.

3. For *codes/4-cps1990_rigmeasures.do*, the raw_data/cps_86_90.dta. However, the query is available in the file raw_data/cps_86_90_query.txt. The data is taken from the IPUMS CPS webpage.

2.4. Figures Folder

The "figures" folder contains the following files:

1. Figure_1A.png: Panel (a) of Figure 1, titled "Effects of the China Shock on employment and non-employment" and subtitle "Tot. employment/Pop." Produced by the code "3-coefs_graphs_decadal.do"
2. Figure_1B.png: Panel (b) of Figure 1, titled "Effects of the China Shock on employment and non-employment" and subtitle "NILF/Pop." Produced by the code "3-coefs_graphs_decadal.do"
3. Figure_1C.png: Panel (c) of Figure 1, titled "Effects of the China Shock on employment and non-employment" and subtitle "Unemployment/Pop." Produced by the code "3-coefs_graphs_decadal.do"
4. Figure_2A.png: Panel (a) of Figure 2, titled "China Shock and unemployment in CZs with high vs. low DNWR" and subtitle "Estimated β3,h". Produced by the code "5-dnwr_figures.do"
5. Figure_2B.png: Panel (b) of Figure 2, titled "China Shock and unemployment in CZs with high vs. low DNWR" and subtitle "β1,h vs. β1,h + β3,h". Produced by the code "5-dnwr_figures.do"
6. Figure_6A.png: Panel (a) of Figure 6, titled "Effects of the China Shock on employment and non-employment" and subtitle "Manuf. employment/Pop.". Produced by the code "3-coefs_graphs_decadal.do"
7. Figure_6B.png: Panel (b) of Figure 6, titled "Effects of the China Shock on employment and non-employment" and subtitle "Non-manuf. employment/Pop.". Produced by the code "3-coefs_graphs_decadal.do"
8. Figure_6C.png: Panel (c) of Figure 6, titled "Effects of the China Shock on employment and non-employment" and subtitle "NILF/Pop.". Produced by the code "3-coefs_graphs_decadal.do"
9. Figure_6D.png: Panel (d) of Figure 6, titled "Effects of the China Shock on employment and non-employment" and subtitle "Unemployment/Pop.". Produced by the code "3-coefs_graphs_decadal.do"
10. Figure_A1.png: Appendix A Figure A1, titled "Baseline empirical analysis using an alternative construction of the unemployment-to-population ratio". Produced by the code "3-coefs_graphs_decadal.do"
11. Figure_A2A.png: Panel (a) of Appendix Figure A2, titled "China Shock on unemployment in CZs with high vs low DNWR. Share of individuals with negative wage changes in state s is below the median across all states" and subtitle "Estimated β3,h". Produced by the code "5-dnwr_figures.do"
12. Figure_A2B.png: Panel (b) of Appendix Figure A2, titled "China Shock on unemployment in CZs with high vs low DNWR. Share of individuals with negative wage changes in state s is below the median across all states" and subtitle "β1,h vs. β1,h + β3,h". Produced by the code "5-dnwr_figures.do"
13. Figure_A3A.png: Panel (a) of Appendix Figure A3, titled "China Shock on unemployment in CZs with high vs low DNWR. Share of individuals with negative wage changes among non-zero wage changes in state s is below the mean across all states" and subtitle "Estimated β3,h". Produced by the code "5-dnwr_figures.do"
14. Figure_A3B.png: Panel (b) of Appendix Figure A3, titled "China Shock on unemployment in CZs with high vs low DNWR. Share of individuals with negative wage changes among non-zero wage changes in state s is below the mean across all states" and subtitle "β1,h vs. β1,h + β3,h". Produced by the code "5-dnwr_figures.do"

15. Figure_A4A.png: Panel (a) of Appendix Figure A4, titled "China Shock on unemployment in CZs with high vs low DNWR. Share of individuals with negative wage changes among non-zero wage changes in state s is below the median across all states" and subtitle "Estimated β3,h". Produced by the code "5-dnwr_figures.do"
16. Figure_A4B.png: Panel (b) of Appendix Figure A4, titled "China Shock on unemployment in CZs with high vs low DNWR. Share of individuals with negative wage changes among non-zero wage changes in state s is below the median across all states" and subtitle "β1,h vs. β1,h + β3,h". Produced by the code "5-dnwr_figures.do"
17. Figure_A5A.png: Panel (a) of Appendix Figure A5, titled "China Shock on unemployment in States with high vs low DNWR. Share of individuals with negative wage changes in state s is below the mean across all states" and subtitle "Estimated β3,h". Produced by the code "5-dnwr_figures.do"
18. Figure_A5B.png: Panel (b) of Appendix Figure A5, titled "China Shock on unemployment in States with high vs low DNWR. Share of individuals with negative wage changes in state s is below the mean across all states" and subtitle "β1,h vs. β1,h + β3,h". Produced by the code "5-dnwr_figures.do"
19. Figure_A6A.png: Panel (a) of Appendix Figure A6, titled "China Shock on unemployment in States with high vs low DNWR. Share of individuals with negative wage changes in state s is below the median across all states" and subtitle "Estimated β3,h". Produced by the code "5-dnwr_figures.do"
20. Figure_A6B.png: Panel (b) of Appendix Figure A6, titled "China Shock on unemployment in States with high vs low DNWR. Share of individuals with negative wage changes in state s is below the median across all states" and subtitle "β1,h vs. β1,h + β3,h". Produced by the code "5-dnwr_figures.do"
21. Figure_A7A.png: Panel (a) of Appendix Figure A7 titled "China Shock on unemployment in States with high vs low DNWR. Share of individuals with negative wage changes among non-zero wage changes in state s is below the mean across all states" and subtitle "Estimated β3,h". Produced by the code "5-dnwr_figures.do"
22. Figure_A7B.png: Panel (b) of Appendix Figure A7, titled "China Shock on unemployment in States with high vs low DNWR. Share of individuals with negative wage changes among non-zero wage changes in state s is below the mean across all states" and subtitle "β1,h vs. β1,h + β3,h". Produced by the code "5-dnwr_figures.do"
23. Figure_A8A.png: Panel (a) of Appendix Figure A8, titled "China Shock on unemployment in States with high vs low DNWR. Share of individuals with negative wage changes among non-zero wage changes in state s is below the median across all states" and subtitle "Estimated β3,h". Produced by the code "5-dnwr_figures.do"
24. Figure_A8B.png: Panel (b) of Appendix Figure A8, titled "China Shock on unemployment in States with high vs low DNWR. Share of individuals with negative wage changes among non-zero wage changes in state s is below the median across all states" and subtitle "β1,h vs. β1,h + β3,h". Produced by the code "5-dnwr_figures.do"

3. Data for the Calibration Exercise ("DataCalibration" Folder)

3.1. R and RStudio

All the codes in this folder are written and implemented using R and RStudio. R is a free, open-source programming language and software environment for statistical computing and graphics, available for download from the following webpage:

https://www.r-project.org/

RStudio is a free integrated development environment (IDE) for R that facilitates coding, debugging, and visualization. It can be accessed at:

https://www.rstudio.com/

We ran our codes using R 4.1.2, but older or newer software versions should work similarly.

Note #1: We use the command "setwd()" at the start of the main code "0-MASTER" .Rmd, which moves the current working directory up one level from the file's location. This ensures that relative file paths in the code work as intended regardless of where the entire replication folder is located in a given computer. For consistent replication, please place the files in a directory structure that matches the project's organization (i.e., one main folder which contains three subfolders labeled: "codes", "0-Raw_Data", "1-Intermediate_Processed_Data", "2-Final_Data").

Note #2: Some Rmd scripts also call Stata (see section 2.1 for a software description). To implement this, one needs to modify line 3 of the "master file" to specify the path where the Stata executable file is located.

3.2. Codes Folder

The "codes" folder contains ten scripts. The file "0-MASTER.Rmd" (located outside of the "codes " folder) is a central script that runs all the other codes and generates all the exhibits. These scripts do not produce any final output for the paper; they only generate intermediate inputs for the model. Here is a description of the ten scripts:

- 1-WIOD_VA_shares_io_shares_distances.Rmd: Calculates bilateral trade flows between countries for all sectors. Calculates calibration targets. Calculates the share of value-added for each sector in each country. Calculates the input-output matrix shares for each country. Calculates the distances between all regions. Calculates the distance elasticity and own dummy coefficients for trade flows in services and agriculture, which correspond to the output in Table C1.
- 2-State_country.Rmd: Calculates state-country bilateral flows for all sectors except services.
- 3-State_state.Rmd: Calculates state-state bilateral flows for all sectors except services and agriculture.
- 4-Gravity_services.Rmd: Calculates state-state and state-country bilateral service flows using a gravity system approach.
- 5-Gravity_agriculture.Rmd: Calculates state-state bilateral flows for agriculture using a gravity system approach.
- 6-Matrices_VA_shares_states.Rmd: Combines the results of the previous five scripts to obtain the final matrix of bilateral flows (for all sectors and regions). Calculates the share of value added in gross output for each US state and sector.
- 7-Exposure.Rmd: Calculates employment levels by state-sector for the year 2000 using BLS and CENSUS data. Calculates the change in imports from China to the US and other advanced economies to run a linear regression using the change of sector US imports from China as the dependent variable and the change of sector advanced economies' imports from China as the independent variable. Computes the exposure measures of equation (20) from the manuscript.
- 8-Employment_1999_2000.Rmd: Computes the employment level for each state and sector in 2000 using CBS data. Computes the employment level for each state and sector in 2000 and 1999 using BLS data. Computes the employment level for each country and sector in 1999 and 2000 using ILO and SEA data. Combines the previous Output and apply proportionality to ensure consistency with WIOD SEA.

- 9-Migration_matrix.Rmd**:** Computes ACS total state-to-state migration movements (knowing the sector of destination, but not the sector of origin) by year for 1999-2001. Computes IRS total state-to-state migration movements (not knowing the sector of origin nor the sector of destination) for 1999. Computes CPS total state-to-state migration movements (knowing both the origin and destination sector) by year for 1999-2001. Combines the data from the previous three steps to compute our final sector-state-to-sector-state mobility flows. Calculates the corresponding migration shares and a case without migration (that is, only using sector reallocation).
- 10-Joining_all.Rmd: Joins all the final data (except for the calibration targets) in a single file.

The following table provides a concordance between the codes ran by the "master" file, their output, and the corresponding paper exhibit:

Program	Output	Paper Exhibit #

1-WIOD_VA_shares_io_shares_distances.Rmd:	`Intermediate_Processed_Data/WiodFixAgg, i,".csv`
	`1-Intermediate_Processed_Data/wiot_full.csv`
	`1-Intermediate_Processed_Data//WIOD_countries.csv`
	`1-Intermediate_Processed_Data/country_country_step_.csv`
	`1-Intermediate_Processed_Data/labor_shares_countries.csv`
	`1-Intermediate_Processed_Data/value_added_countries", yr, ".csv`
	`1-Intermediate_Processed_Data/io_shares.csv`
	`1-Intermediate_Processed_Data/io_allyears.xlsx`
	`1-Intermediate_Processed_Data/country_coordinates.csv`
	`1-Intermediate_Processed_Data/distances.csv`
	`2-Final_Data/calibration_targets.xlsx`	

2-State_country.Rmd	`1-Intermediate_Processed_Data/census_exports.csv`
	`1-Intermediate_Processed_Data/census_imports.csv`
	`1-Intermediate_Processed_Data/state_imports_exports.csv`
	`1-Intermediate_Processed_Data/state_country_step_y_.csv`
	`1-Intermediate_Processed_Data/state_country_step_e_.csv`	

3-State_state.Rmd	`1-Intermediate_Processed_Data/CFSapportionment.csv`
	`1-Intermediate_Processed_Data/CFS_Xijk.csv`
	`1-Intermediate_Processed_Data/state_cfs_step_.csv`	

4-Gravity_services.Rmd	`1-Intermediate_Processed_Data/state_exp_rev_services.csv`
	`1-Intermediate_Processed_Data/data_services.csv`
	`1-Intermediate_Processed_Data/vector_lambda_.csv`
	`1-Intermediate_Processed_Data/matrix_B_.csv`
	`1-Intermediate_Processed_Data/vector_solution_.csv`
	`1-Intermediate_Processed_Data/Xij_matrix_services_.csv`	

5-Gravity_agriculture.Rmd	`1-Intermediate_Processed_Data/data_agriculture.csv`
	`1-Intermediate_Processed_Data/agric_mat_B_.csv`
	`1-Intermediate_Processed_Data/agric_vec_lambda_.csv`
	`1-Intermediate_Processed_Data/vec_agric_solution_.csv`
	`1-Intermediate_Processed_Data/Xij_matrix_agric_.csv`	

6-Matrices_VA_shares_states.Rmd	`1-Intermediate_Processed_Data/final_matrix_.csv`
	`1-Intermediate_Processed_Data/bilat_matrix_allyears.xlsx`
	`1-Intermediate_Processed_Data/labor_shares_states.csv`
	`1-Intermediate_Processed_Data/va_shares_allyears.xlsx`	

7-Exposure.Rmd	`1-Intermediate_Processed_Data/state_emp_share_2000.dta` `1-Intermediate_Processed_Data/exposures.xlsx`	

8-Employment_1999_2000.Rmd	`1-Intermediate_Processed_Data//L_`2000CENSUS`.csv`
	`1-Intermediate_Processed_Data//L_`1999BLS`.csv`
	`1-Intermediate_Processed_Data//L_`2000BLS`.csv`	

9-Migration_matrix.Rmd	`1-Intermediate_Processed_Data//acs_temp.csv`
	`1-Intermediate_Processed_Data/acs.csv`
	`1-Intermediate_Processed_Data/irs.csv`
	`1-Intermediate_Processed_Data/cps_nber_2_yearly.csv`
	`1-Intermediate_Processed_Data/mu_1999.xlsx`
	`1-Intermediate_Processed_Data/mu_1999_no_migration.xlsx`	

10-Joining_all.Rmd	`2-Final_Data/InputData.xlsx`	

3.3. Raw Data Folder

The “0-Raw_Data” folder stores our main datasets. These files are publicly available and can be downloaded following the instructions in this Readme.

3.3.1. Classification tables

Files: 0-Raw_Data/regions.csv and 0-Raw_Data/sectors.csv

All the scripts use `0-Raw_Data/regions.csv` and `0-Raw_Data/sectors.csv`: the first .csv file is simply a table that lists all of our regions of interest, assigns them a number, and classifies them into two categories ("country" and "state"); the second .csv file is a table that creates a correspondence between WIOD sectors, NAICS sectors, and our final sectors.

3.3.2. Country-country bilateral flows

Files: 0-Raw_Data/WIOD/wiot_*.xlsx

The first script uses the collection of 0-Raw_Data/WIOD/wiot_*.xlsx files to construct bilateral flows between countries for all sectors (WIOD, 2025). These files are the 2013 world input-output tables by year for 2000-2011. They can be downloaded directly from WIOD's page: select *WIOT Tables Excel*.

3.3.3. Data for distances

Files: 0-Raw_Data\Fips*, 0-Raw_Data\Population_geography*

To construct trade distances, the first script needs coordinates and populations for the regions of interest (to weigh the distances derived solely from coordinates).

For states, that information is stored in the `0-Raw_Data\Fips` folder, which files are taken from here and here.

For countries, that information is stored in the `0-Raw_Data\Population_geography` folder, which files come from:

a) United Nations: Population Division, *Population of Urban Agglomerations with 300,000 Inhabitants or More in 2018, by country, 1950-2035 (thousands)*. Which can be downloaded here.

b) United Nations: Population Division, *Total Population - Both Sexes. De facto population in a country, area or region as of 1 July of the year indicated. Figures are presented in thousands, w*hich can be downloaded here.

c) Republic of Cyprus: Statistical Service, *CENSUS OF POPULATION, HOUSING UNITS AND POPULATION ENUMERATED BY DISTRICT AND URBAN/RURAL AREA (1.10.2021) PRELIMINARY RESULTS, COMPARED TO THE CENSUS OF POPULATION (1.10.2011)*. Which can be downloaded here.

The final outputs are primarily based on input a), supplemented by some information from b) and c). Some countries did not fulfil the requirement of providing information for at least two cities. For all these special cases (with the exception of Cyprus), b) provides the necessary information, as it includes the population of cities by sex (we only take the aggregation of male and female). From b), we use the urban agglomeration category, when possible, to compare it with the primary UN base we are utilizing. In the specific case of

Cyprus, the city's information is sourced from the country's Statistical Service; since the available population statistics refer to the provinces (named after the main city within them), rather than the city itself, we use the province's urban population instead of the total province population.

Finally, all population statistics must be expressed in thousands of people.

3.3.4. US state-country trade database

Files: 0-Raw_Data/State-Exports-Imports/State Exports by NAICS Commodities-i.csv and 0-Raw_Data/State-Exports-Imports/State Imports by NAICS Commodities-i.csv.

To calculate state-country bilateral flows for all sectors except services, the second script uses a US state-country trade database for 2002-2019 (imports as of 2008) at the 3-digit NAICS. The files are: 0-Raw_Data/State-Exports-Imports/State Exports by NAICS Commodities-i.csv and 0-Raw_Data/State-Exports-Imports/State Imports by NAICS Commodities-i.csv. To download them, follow these steps:

- Enter the USA Trade Online (USA Trade Statistics, 2025) website. Register, log in, and then click *ACCESS DATA*. At this point, the main page for extracting data will appear with the data selection options.
- Choose *State Export Data (Origin of Movement)*: *NAICS*. The main sections to complete will appear on the left side of the webpage, specifically measures, state, commodity, country, and time.
- The measure of interest is *Total Exports Value ($US)*. Select the first eleven states ranging from *All states* to Florida. Then, in the commodity section, select the check in the second column on the upper part: this will select all 3-digit NAICS categories but not their subsections. Do not select the **All commodities** option. Within the country section, select the check next to **World Total** and then unselect **World Total itself**. Finally, for the time section, select the check that belongs to the first column of the upper section; this will select all years but not their month subdivision.
- Generate the report through the option REPORT, which is located on the left side of the **WELCOME (NAME)** sign.
- Download the data using the down green arrow option, and choose to do it using a comma separated csv file.
- Repeat the past steps but selecting the following 11 states on the list: Georgia-Maryland, Massachusetts-New Mexico, New York-South Dakota, and Tennessee-unknown. There is no need to exit the webpage to gather the information of the next 11 states; just modify it in the State section.
- Rename each downloaded file as State Exports by NAICS Commodities-i.csv, $i \in {1,2,3,4,5}$
- To proceed with the import data download, choose the *Data Source Selection* option on the upper side.
- Choose *State Import Data (State of Destination)*: *NAICS*. The process is fairly similar with some exceptions. In measures, choose *Customs Value (Gen) ($US)*.
- Rename each downloaded file as State Imports by NAICS Commodities-i.csv, $i \in {1,2,3,4,5}$

3.3.5. State-state trade database

Files: 0-Raw_Data/CFS/CFS2002mine.xlsx, 0-Raw_Data/CFS/CFS2007.xlsx, 0-Raw_Data/CFS/CF1200A24.csv, 0-Raw_Data/CFS/NAICS to CFS Crosswalk.xlsx and 0-Raw_Data/CFS/CFS_2012_00A18_with_ann.csv.

The third script uses the Census Bureau's Commodity Flow Survey (CFS) for 2002, 2007, and 2012 to construct bilateral flows between states for all sectors except services and agriculture.

- CFS data for 2002 can be found here. Specifically, Table 14. Shipment Characteristics by Destination and Two-Digit Commodity for State of Origin: 2002. We simply converted it to .xlsx format (which is the origin of 0-Raw_Data/CFS/CFS2002mine.xlsx).
- CFS data for 2007 can be downloaded from here. Specifically, *CF0700A22.zip*. We simply converted the .dta file (inside the zip folder) to .xlsx format (which is the origin of 0-Raw_Data/CFS/CFS2007.xlsx).
- CFS for 2012 can downloaded from here. Specifically, *CF1200A24.zip*. We simply converted the .dta files (inside the zip folders) to .csv format (and that is the origin of 0-Raw_Data/CFS/CF1200A24.csv).
- Those data sets give state-state bilateral trade flows in commodities. It is important to note that each commodity can be associated with more than one NAICS sector, depending on the industry that that commodity was produced in. So, to distribute the bilateral flows from the previous three data sets to our 12 manufacturing sectors, script 6 calculates first the proportion of the amount of a commodity associated with a NAICS sector, with respect to the total amount of that commodity. To get those total amounts associated with a NAICS sector, one needs to download *CF0700A18.zip* from here to distribute for 2002 and 2007, and *CF1200A18.zip* from here to distribute for 2012. That is the origin of 0-Raw_Data/CFS/NAICS to CFS Crosswalk.xlsx and 0-Raw_Data/CFS/CFS_2012_00A18_with_ann.csv, respectively.

3.3.6. Revenue by state industry and expenditure by state

Files: 0-Raw_Data\Expenditure\SAEXP1_1997_2017_ALL_AREAS_.csv, and 0-Raw_Data\SAGDP\SAGDP2N_ .csv.

To construct the gravity system for services, script 4 uses BEA's expenditure by state (SAEXP1) and GDP by state and industry (SAGDP2). Due to new estimations, the versions of these data sets that we have used are not available in the main BEA's Regional Accounts page anymore. However, they can still be found as archived releases. To get 0-Raw_Data\Expenditure\SAEXP1_1997_2017_ALL_AREAS_.csv, download it from here; and to get the collection of 0-Raw_Data\SAGDP\SAGDP2N_ .csv files, download them from here.

3.3.7. Agriculture and fishing revenue

Files: 0-Raw_Data\Agriculture_Census\data_agriculture.csv and 0-Raw_Data\Agriculture_Census\fish_raw.xlsx.

To construct the gravity system for sector 14, script 5 uses USDA's Agriculture Census data; specifically, *Table 2. Market Value of Agricultural Products Sold Including Landlord's Share and Direct Sales: 2007 and 2002*, which can be downloaded from here; we simply put the data in .csv format to create 0-Raw_Data\Agriculture_Census\data_agriculture.csv.

For the same purpose, script 5 also uses NOAA's data. To create 0-Raw_Data\Agriculture_Census\fish_raw.xlsx, we join the *U.S domestic landings, by region and by state* tables of each year's (from 2000 to 2007) *Fisheries of the United States Full Report*, which can be found here.

3.3.8. Data for value-added shares

Files: 0-Raw_Data\Labor_shares\gdp_state_*.csv, 0-Raw_Data\Labor_shares\subsidies_state_.*csv files, and 0-Raw_Data\Labor_shares\taxes_state_*.csv.

To construct value added shares, script 6 needs BEA's GDP (SAGDP2), subsidies (SAGDP5), and taxes (SAGDP6) by state and industry. The versions of these datasets we have used are no longer available on the main BEA's Regional Accounts page. However, they can still be found as archived releases. To get the information contained in the collection of 0-Raw_Data\Labor_shares\gdp_state_.csv files, 0-Raw_Data\Labor_shares\gdp_state_.csv files, and 0-Raw_Data\Labor_shares\gdp_state_.csv files, download the .zip folder from here, and select, respectively: SAGDP2N__ALL_AREAS_1997_2018.csv, SAGDP5N__ALL_AREAS_1997_2017.csv, and SAGDP6N__ALL_AREAS_1997_2017.csv.

3.3.9. Data for labor shares by sectors

Files: 0-Raw_Data\CENSUS_2000\PUMS5*, 0-Raw_Data\ADH_employment*, 0-Raw_Data/emp_SAEMP25S_BLS.csv and 0-Raw_Data\ADH_employment\state_emp_2000.dta.

To calculate our exposure measures, script 7 needs labor shares by state and sector. We calculate one version of labor shares using Census Bureau's data without person weights. We use the 5 % sample PUMS files of the 2000 Census. These files are available here. One needs to open the folder *All additional files for the PUMS 5-Percent Dataset* and download the file *PUMS5.txt* within each state's folder. These files are the content of our 0-Raw_Data\CENSUS_2000\PUMS5 folder. However, for Alabama the PUMS5 file (PUMS5_01.txt) was incomplete. So, we replace it with the corresponding REVISEDPUMS5 file (we rename it as PUMS5_01.txt to write an easier loop to import the files in script 7).

We also use other two versions of labor shares using BLS-BEA data and ADH data. To calculate BLS-BEA shares, we use 0-Raw_Data/emp_SAEMP25S_BLS.csv, which can be accessed from here (together with this link. We take the ADH files from their replication package (also on Prof. David Dorn's website). To obtain "0-Raw_Data/ADH_employment/state_emp_2000.dta", one must run first run the cbp2000_imputations.do, cbp2000_to_cz.do, and cz_to_states.do codes (in that order) from ADH's replication file.

3.3.10. Employment allocation in each country and sector for 1999 and 2000

Files: 0-Raw_Data\CENSUS_2000\P_World unemployment rates.xlsx, 0-Raw_Data\CENSUS_2000\EAP_2WAP_SEX_AGE_RT_A-filtered-2023-11-02.csv, and 0-Raw_Data\CENSUS_2000\L_1999_2000_countries.csv.

Script 8 computes the employment level for each country and sector for year 1999 and year 2000 using ILO and SEA data. To get the necessary raw data, the steps to follow are:

1. Download ILO estimate country unemployment rates u from here, selecting only our years and countries of interest; that is how we get 0-Raw_Data\CENSUS_2000\P_World unemployment rates.xlsx.
2. Download ILO estimates for country labor force participation rates lfp from here, selecting only "Total" rows and only for our years and countries of interest; that is how we get 0-Raw_Data\CENSUS_2000\EAP_2WAP_SEX_AGE_RT_A-filtered-2023-11-02.csv.
3. Employment for sectors 1-14 for each country are taken from WIOD's Socio Economic Accounts (SEA). We download SEA's release of July 2014 in Excel format from here. The only variable of interest is EMP (number of persons engaged in thousands); after reclassifying it into our 1-14 sectors, we get our 0-Raw_Data\CENSUS_2000\L_1999_2000_countries.csv file.

3.3.11. Employment allocation in each state and sector using CBS data

File: 0-Raw_Data\CENSUS_2000\employment_2000.xls.

Script 8 calculates the employment allocation in each state and sector for 2000 using CBS data. To get the necessary raw data, the steps to follow are:

1. Download the employment, unemployment and non-participation levels for each state in year 2000 that CBS published based on the 2000 Decennial Census. The file is entitled *Table 1. Employment Status of the Population 16 Years Old and Over in Households for the United States, States, Counties, and for Puerto Rico: 2000*. Download the revised Excel file; it can be accessed here. That is our 0-Raw_Data\CENSUS_2000\employment_2000.xls.
2. However, the data from the previous step does not include distribution of workers by sectors. To gather such disaggregation, we use the 5 % sample REVISEDPUMS files of the 2000 Census. These files are available here. One needs to open the folder *All additional files for the PUMS 5-Percent Dataset* and download the file REVISEDPUMS5.txt within each state's folder. Also, note that the variable dictionary for the census data is located in the same link and is entitled 5%_PUMS_record_layout.xls. Note that each observation is compressed in a single line to save space. Therefore, the variables of interest have to be created using start and stop character positions. For more information regarding the NAICS industry codes used in the census sample files, see p. 547 here. The most updated version of the previous NAICS table can be found here. Whenever available, we use NAICS codes instead of Census Industry codes, and then recode them to final sectors. In the 2000 census sample we only keep observations type "P" (persons) with age in between 25 and 65, and that are employed. Non-participation will be considered an extra sector: sector 0. The "employment" for sector 0 is taken directly as the reported value of non-participants for each state in 0-Raw_Data\CENSUS_2000\employment_2000.xls; to construct the other 14 sectors, NAICS sectors are recoded into our final sectors and we apply a proportionality rule (using person weights instead of respondents) with respect to the population employment levels by state.

3.3.12. Employment allocation in each state and sector using BLS data

Files: 0-Raw_Data\staadata.xlsx, 0-Raw_Data\L_1999\Sectors\sector_, and Raw_Data\L_2000\Sectors\sector_ .

Script 8 also computes the employment level for each state and sector for 1999 and 2000 using BLS data. To get the necessary raw data, the steps to follow are:

1. Download BLS' *annual average industry-specific NAICS-Based Quarterly Census of Employment and Wages by state* (for 1999 and 2000) or, from now on, QCEW, here. QCEW datasets take values at the state-level and at the county-level; for each NAICS (sub-)sector we take the values at the state level. That is the content of the 0-Raw_Data\L_1999\Sectors\sector_ and 0-Raw_Data\L_2000\Sectors\sector_ folders.
2. To calculate sector 00, download BLS' *Employment status of the civilian noninstitutional population, Annual Average Series* (for 1999 and 2000) here. The result is our 0-Raw_Data\staadata.xlsx file.

3.3.13. Migration flows using ACS, CPS, and IRS data.

Files: 0-Raw_Data\ACS\"year"*, 0-Raw_Data/CPS/CPS_NBER/Inputs/cps_panel.dta, and Raw_Data\IRS\1999to2000CountyMigration\1999to2000flow.dta.

Script 9 calculates migration flows between sectors and states using ACS, CPS, and IRS data.

The American Community Survey (ACS) Public Use Microdata Sample (PUMS) files provide details of workers' current employment status, sector, and state. It also asks the state in which respondents lived the prior year. However, this survey does not provide information regarding people's employment status and sector in the previous year. Therefore, as a second component of our calculations, we use the Current Population Survey (CPS), which provides details of people's employment status, industry sector and state for each month from 1999 to 2007. However, this survey does not provide information regarding movements across states. Thus, we combine both the ACS and the CPS surveys to compute the labor transitions across states and sectors every year.

Steps to follow:

1. The ACS PUMS csv files can be downloaded from the FTP server in here. Within the FTP server, open the (1 year ACS) folders for the years 2000-2007 and download the files titled *csv_p{state}.zip*. That is the content of our `0-Raw_Data\ACS\year` folders. There is no need to change the name of the downloaded csv files after decompressing the zip files. The survey's variable dictionary can be accessed through this link and the questionnaire through this other link. It is important to note that the state of origin variable changed its categories in 2003 (this issue is corrected in code).
2. We only keep observations with age in between 25 and 65, and that are either employed, unemployed or not in the labor force; each observation has a person weight that comes with the survey. Unemployment plus not in the labor force will be considered an extra sector: sector 0. We convert NAICS sectors into the final sectors of our paper. Following CDP (2019) advice, in order to estimate the transitions more precisely, we use not only the annual changes of each year of interest but also the ones from the two following years. Hence, three lists of changes are seen now as annual changes for the year of interest. The current sector, current state, and previous year state are used to compute partial transition matrices with state of origin and destination and destination sector (as we said earlier, there is no information about the sector of origin in the ACS data).
3. CPS data (NBER CPS, 2025) is stored in `0-Raw_Data/CPS/CPS_NBER/Inputs/cps_panel.dta`. To obtain it, the next steps must be followed: a) Login to IPUMS CPS site b) Create an account at the *Register* option. c) Go to *Data->Browse and select data* d) Select the samples. It is important to select the cross-sectional format and unselect the ASEC supplement. Go to the monthly based section and select the desired months of each year. e) Click *Submit sample selections*. f) Select the variables (for Household and Person): The core variables selected for the creation of the base were: year, serial, month, hwtfinl, cpsid, asecflag, mish, numprec, statefip, nfams, hrhhid, hrhhid2, hhrespln, pernum, wtfinl, cpsidv, cpsidp, age, sex, race, marst, popstat, bpl, nativity, empstat, labforce, occ, and ind. g) Go to view cart and create the data extract. The variable *mish* is important because it identifies the month in a survey at the household level. h) Download the data extract in a ".dta" format.
4. The CPS surveys households in a 4-8-4 format; that is, it interviews the household for 4 consecutive months, gives them an 8-month break and interviews them again for 4 consecutive months. Also, a new household can start the 4-8-4 sequence any month of the year. Since the CPS data comes in a monthly frequency and we are interested in workers' movements across time, we have to match households and individuals across time. Each household ID is a concatenation of 4 variables (see code). However, this ID suffers a simplification in May-2004, which makes automatic matches harder around that period. The necessary adjustment is detailed in the new dictionary and implemented in the code; basically, the new household ID is the previous one with some specific digits taken out. Persons within a household do not have a unique ID, therefore, one is created for them using the household ID and the person's age and gender.
5. We only keep observations with age in between 25 and 65, and that are either employed, unemployed or out of the labor force; each observation has a person weight that comes with the survey (specifically, the one named: *Weight-composited final weight: Person's final composited*

weight. Used to tabulate BLS's official published labor force statistics). Again, unemployment plus not in labor force will be considered sector 0 (again, not being consistent with script 8). CPS uses Census industry codes instead of NAICS, so, a different recoding is applied to convert them to the final sectors of interest. Also, the Census industry codes were updated in 2003 and changed positions in the variable dictionary. For more information regarding the Census industry codes changes see link; the old and new codes can be accessed here and here.

6. We match CPS observations (individuals) across time using the interview number. Each of the first 4 monthly interviews are 12 months apart from the final 4 interviews, and the first four and final four are consecutive in months. We are interested in recording annual changes for each year. That is why worker movements' information is gathered using the following interview matching system: (1,5), (2,6), (3,7) and (4,8). The latter is equivalent to following each individual for twelve months. To avoid double counting, we use only the match of the first and fifth interview: (1,5). The annual change observation is assigned to the respective year of the base month of the change. Following CDP (2019) advice, in order to estimate the transitions more precisely, we use not only the annual changes of each year of interest, but also the ones from the two previous and following years. Five lists of changes are seen now as annual changes for the year of interest. These matches, together with each person's information about past and current state and sector make it possible to compute changes in the labor allocation across time. The CPS survey does not account for movements across states and that is why we still need the ACS values. We follow CDP's assumption that interstates movements (i to j) across sectors follow the same pattern that intrastate movements in state j across sectors.
7. Finally, we use a proportionality rule to make our migration flows sum up (over sectors of origin and destination) to the state-state flows of IRS, which can be downloaded from here (specifically, the .zip folder for County-to-County Migration Data from 1999 to 2000). We use this data in script 9 to produce our `0-Raw_Data\IRS\1999to2000CountyMigration\1999to2000flow.dta`.

3.3.14. Note: Large files

Four types of raw data files are larger than 150 MB. We already mentioned them above in their relevant subsection, but we explained them here again. The files are as follows:

1. For the code *codes/3-State_state.Rmd*, the file `0-Raw_Data/CFS/CF1200A24.csv`.
2. For the codes codes/7-Exposure.Rmd and *codes/8-Employment_1999_2000.Rmd*: the collection of `0-Raw_Data/CENSUS_2000/PUMS5/PUMS5_` .txt files for each US state. The same comment applies to the collection of `0-Raw_Data/CENSUS_2000/REVISEDPUMS5/REVISEDPUMS5_ .txt` files.
3. For the code *codes/9-Migration_matrix.Rmd*: the collection of `0-Raw_Data/ACS/20*/c2ssp**.csv` .txt files for each US state.
4. For the code *codes/9-Migration_matrix.Rmd*: `0-Raw_Data/CPS/CPS_NBER/Input/cps_panel.dta.`

3.4. "2-Final_Data" Folder

The "2-Final_Data" folder contains two files. Both files are intermediate inputs used in the model. These are the files:

1. calibration_targets.xlsx: This file includes the changes in U.S. imports from China predicted by the changes in imports from China to other high-income countries. We use these as targets to calibrate the productivity increase in China. Our procedure is explained in Section 5 of the paper.

2. InputData.xlsx: This file contains nine sheets, each serving as an input for the model. These sheets correspond to the data for the quantitative exercise described in Section 4 of the paper.

4. Model Output ("Model" Folder)

The "Model" folder contains 21 Matlab code files (i.e., ".m" files), a subfolder with inputs required to run the programs (titled "Inputs"), and 9 Matlab result files (i.e., ".mat" files). The file titled "NEST_Main1" contains a primary code that calls other functions (namely, all the ones that start with "NEBA" or "NECO") to produce the files "Results1"-"Results4Flex". The file titled "NEST_Main2" contains another primary code that calls the function "NEST_OuterAlgoCF_FUNC" to produce file "ResultsDeltaTable". Finally, the files that start with "NEST_Produce", contain codes that use the stored results to produce exhibits and inline numbers in the paper. These codes were run with Matlab R2023b. If the codes are run with a different version of Matlab, everything should still go through, but there might be slight compatibility issues.

4.1. Matlab

Matlab R2023b can be purchased for a fee from the company MathWorks using the following link:

https://www.mathworks.com/products/get-matlab.html?s_tid=gn_getml

Note on user-contributed packages: The folder contains a Matlab file named "parsave", which is a user-contributed command that provides some flexibility while saving results. Namely, it works inside "parfor" Matlab loops (which do not accept a "regular" save command), and it allows to save variables without having to refer to their name strings, one can simply refer to their names directly. This command was created by Joost H. Weijs in 2016 and he holds the copyright.

Note on directories: The main Matlab files do not set any special directory, so the user has to point Matlab to the folder where the files are located. The user should also keep the same structure as the one in the replication file, with a main folder that contains all codes and results and a subfolder named "Inputs" that contains the required files.

4.2. Inputs Subfolder

The "Inputs" folder contains two excel files and one Matlab code:

1. InputData.xlsx: This is an Excel file that contains all the raw data required to run the model that comes from the "DataCalibration" part of the replication. Specifically, it contains nine sheets, as follows:
    1. BILATRAW: Contains 14 matrices of size 87*87 containing bilateral trade flows between any pair of regions in our sample for each of the 14 market sectors. The 14 matrices are stacked vertically, so the total matrix is of size 1218*87.
    2. IO: Contains a 14*14 matrix of input-output coefficients for each of the 38 countries in the sample. The matrices are stacked vertically, so the total matrix is of size 532*14.
    3. VA: Contains an 87*14 matrix of value-added shares for each region-sector.
    4. EXP: Contains a measure of exposure to the China shock for each of the 50 U.S. states.
    5. L1999BLS: Contains an 87*15 matrix of labor participation for each region-sector (here there are 15 sectors, instead of 14, because this includes home production), coming from the BLS for the year 1999.
    6. L2000BLS: Same as the previous bullet point but for the year 2000.

7. L2000Census: Same as the previous bullet point but coming from the Census instead of the BLS.
8. MURAW: A 750*750 matrix of mobility between any state-sector pair and any other state-sector pair in the U.S.
9. MURAWNM: Same as the previous bullet point but restricted so people do not move across states and only move across sectors within a state.

2. DataCleaner.m: A Matlab file that uses the data in "InputData.xlsx" and processes it in order to obtain the cleaned data which is written to the file "ProcessedData.xlsx". Specifically, there are three sections:
    1. Section 1: It processes the data in BILATRAW in order to avoid negative Cobb-Douglas shares, as described in Section C.4 of the online appendix to our paper.
    2. Section 2: It processes MURAW in order to make it compatible with the labor vectors in 1999 and 2000, as described in Section C.3 of the online appendix to our paper.
    3. Section 3: It processes MURAWNM in order to make it compatible with the labor vectors in 1999 and 2000, as described in Section C.3 of the online appendix to our paper.
3. ProcessedData.xlsx: It contains the processed data produced by the "DataCleaner.m" file. Specifically, it contains the following sheets:
    1. BIFIXED: Has 14 matrices of size 87*87 containing bilateral trade flows between any pair of regions in our sample for each of the 14 market sectors, but adjusted to not imply any Cobb-Douglas shares that take negative values.
    2. L1999MIX: It contains the 1999 "CBLS" data described in Section C.3 of our online appendix.
    3. MUFIXED: A 750*750 matrix of mobility between any state-sector pair and any other state-sector pair in the U.S., but adjusted to be compatible with the labor vectors in 1999 and 2000, as described in Section C.3 of the online appendix to our paper.
    4. MUFIXEDNM: Same as the previous bullet point but restricted so people do not move across states and only move across sectors within a state.

4.3. Codes

The "Model" folder contains 8 Matlab code files. Specifically:

1. NEST_Main1.m: This is the central code file that runs the model in the paper. It takes as input the files in the folder "Inputs", and also uses Matlab functions that start with "NEBA" and "NECO". It runs the algorithms to solve for the baseline economy and the counterfactual economy described in online appendix Sections B.4 and B.6. It runs the model and saves the results for four different calibrations (the four calibrations in columns 2 to 5 of Table 1 in the paper), both for the case of flexibility (where the DNWR parameter is set to 0.01, implying no DNWR), and for the case of rigidity (where delta is set to its calibrated value). Therefore, this file produces eight result files, titled ResultsY or ResultsYFlex, where "Y" stands for the calibration number, between 1 and 4. This file could be parallelized (with a Matlab "parfor" loop) for more efficient computation, but the replication code does not give a parallelized version.
2. NEST_Main2.m: This does something similar to the file "NEST_Main1.m" but for many values of the DNWR parameter delta. Specifically, it runs the model for 44 values of delta. In order to make this efficiently, the file uses parallel computing, using a Matlab "parfor" loop to distribute the calculations across 44 cores. However, the number of cores can be adjusted down (to 20, 8, 4, or even 1), if the user does not have 44 cores at their disposal. This can be achieved by changing the value of "K" from 44 to the desired number of available cores. The file uses the function "NEST_OuterAlgoCF_FUNC.m" which is a function version of the code in the file "NEST_Main1.m" which only stores the most important outcome variables. It saves the results in a file called: "ResultsDeltaTable.mat".

3. NEST_ProduceDispersionResults.m: Uses the "Results1" baseline results to produce the dispersion numbers mentioned in the second-to-last paragraph of Section 6.1.
4. NEST_ProduceFigure6ModelPart.m: Uses the "Results1" and "Results2" files to produce the green and black lines of Figure 6.
5. NEST_ProduceFigure8.m: Uses the "ResultsDeltaTable" file to produce Figure 8.
6. NEST_ProduceFigure3457.m: Uses the "Results1" and "Results2" files to produce Figures 3, 4, 5, and 7 in the paper.
7. NEST_ProduceTable1.m: Uses the "Results1"-"Results4Flex" files to produce all the numbers in Table 1.
8. NEST_ProduceTableD1.m: Uses the "Results1" and "Results1Flex" files to produce all the numbers in appendix Table D1.

The "Model" folder also contains 13 Matlab function files (these are still labelled ".m", but they have a different structure than the 8 code files described above). Specifically:

1. NEBA_Block1.m: This is a Matlab function to run the first block of the "baseline" dot algorithm described in online appendix Section B.4.
2. NEBA_Block2.m: This is a Matlab function to run the second block of the "baseline" dot algorithm described in online appendix Section B.4.
3. NEBA_Block3.m: This is a Matlab function to run the third block of the "baseline" dot algorithm described in online appendix Section B.4.
4. NEBA_PeriodTatoSolver.m: This function solves for dot wages and labors (and also prices dot, trade shares new levels and new revenue levels) of the one-period temporary equilibrium in the RUV model using an extended Alvarez-&-Lucas-type tatonnement algorithm that applies in the case of DNWR. This algorithm is described in online appendix Section B.7.
5. NEBA_PriceSolver.m: This function solves for sectoral prices (hats or dots) and level trade shares using a contraction mapping algorithm. This algorithm is described around the first equation of online appendix Section B.7.
6. NEBA_RevenueSolver.m: This function solves for new revenues (hat or dots) using a linear system of equations. This algorithm is described around the third equation of online appendix Section B.7.
7. NEBA_TatoNewGuess.m: This function computes a new guess for wages and labor (hats or dots) given an initial guess. This guess is described in detail in online appendix Section B.7.
8. NECO_Block1.m: This is a Matlab function to run the first block of the "counterfactual" hat algorithm described in online appendix Section B.6.
9. NECO_Block2.m: This is a Matlab function to run the second block of the "counterfactual" hat algorithm described in online appendix Section B.6.
10. NECO_Block3.m: This is a Matlab function to run the third block of the "counterfactual" hat algorithm described in online appendix Section B.6.
11. NECO_PeriodTatoSolver.m: This function solves for hat wages and labors (and also prices hat, trade shares new levels and new revenue levels) of the one period temporary equilibrium in the RUV model using an extended Alvarez-&-Lucas-type tatonnement algorithm that applies in the case of DNWR. This algorithm is described in online appendix Section B.7.
12. NEST_OuterAlgoCF_FUNC.m: A function version of the code in the file "NEST_Main1.m" which only stores the most important outcome variables.
13. parsave.m: A user-contributed function that provides some flexibility while saving results. Namely, it works inside "parfor" Matlab loops (which do not accept a "regular" save command), and it allows to save variables without having to refer to their name strings, one can simply refer to their names directly. It was created by Joost H. Weijs in 2016 and he holds the copyright.

As noted above, of the 21 Matlab codes, only 8 are codes that have to be ran by the user, the remaining 13 are functions that are used by the other codes. The following table provides a concordance between the 8 codes that would be run by the user, their output, and the corresponding paper exhibit:

Program	Output	Paper Figure/Table #
NEST_Main1.m	Files “Results1” through “Results4Flex”	
NEST_Main2.m	File “ResultsDeltaTable”	
NEST_ProduceDispersionResults.m		Inline numbers in Section 6.1
NEST_ProduceFigure6ModelPart.m		Figure 6, green and black lines
NEST_ProduceFigure8.m		Figure 8
NEST_ProduceFigure3457.m		Figures 3, 4, 5, and 7
NEST_ProduceTable1.m		Table1
NEST_ProduceTableD1.m		Appendix Table D1

4.4. Output Files

The following result files are produced by the codes:

1. Results1.mat: Matlab file containing the results of the baseline calibration.
2. Results1Flex.mat: Matlab file containing the results of the baseline calibration but with wage flexibility (i.e., the DNWR parameter delta has been set to 0.01).
3. Results2.mat: Matlab file containing the results of the longer-shock calibration.
4. Results2Flex.mat: Matlab file containing the results of the longer-shock calibration but with wage flexibility (i.e., the DNWR parameter delta has been set to 0.01).
5. Results3.mat: Matlab file containing the results of the no-migration calibration.
6. Results3Flex.mat: Matlab file containing the results of the no-migration calibration but with wage flexibility (i.e., the DNWR parameter delta has been set to 0.01).
7. Results4.mat: Matlab file containing the results of the calibration imposing that the elasticity across sectors be the same as the elasticity across regions.
8. Results4Flex.mat: Matlab file containing the results of the calibration imposing that the elasticity across sectors be the same as the elasticity across regions but with wage flexibility (i.e., the DNWR parameter delta has been set to 0.01).
9. ResultsDeltaTable.mat: Matlab file containing the results of the baseline calibration across many possible values of the DNWR parameter delta.

Note that the Matlab files starting “NEST_Produce” do not really save their output (say a figure or a table) automatically, they simply display the output and the user can save the figure or table how they see fit.

References

Autor, David H.; Dorn, David; Hanson, Gordon H., 2013. “The China Syndrome: Local Labor Market Effects of Import Competition in the United States.” American Economic Review, 103(6), 2121-2168.

IPUMS, 2025. URL: https://usa.ipums.org/usa/

Jo, Yoon J., 2024. "Beyond the Spike at Zero: Understanding Nominal Wage Rigidity through Empirical and Model-Based Approaches," forthcoming in the Journal of International Economics.

Local Area Unemployment Statistics (LAUS), 2025. URL: https://data.bls.gov/PDQWeb/la

NBER CPS, 2025. URL: https://www.nber.org/research/data/current-population-survey-cps-merged-outgoing-rotation-group-earnings-data

Rodriguez-Clare, Andres; Ulate, Mauricio; Vasquez, Jose P., 2025, "Replication Data for: "Trade with Nominal Rigidities: Understanding the Unemployment and Welfare Effects of the China Shock"," Harvard Dataverse, https://doi.org/10.7910/DVN/R96X0Z.

Surveillance, Epidemiology, and End Results (SEER) Program, 2025. URL: https://seer.cancer.gov/popdata/download.html

USA Trade Statistics, 2025. URL: http://usatrade.census.gov/

WIOD, 2025. URL: https://www.rug.nl/ggdc/valuechain/wiod/wiod-2013-release