Replication archive for

Commuting, Migration and Local Employment Elasticities

Ferdinando Monte, Stephen J. Redding and Esteban Rossi-Hansberg

American Economic Review, forthcoming 2018

Table of contents

This pdf has a set of navigable side bookmarks.

Contents of the Archive

This archive has the following directories.

data

This folder and its subfolder contain the main datasets used to generate tables and figures in the data. What follows describes its content starting with the subfolders, and then describing .csv files and .dta files.

Subfolder "county shocks 5 pct with no comm\"

This subfolder contains 3111 .csv files, one per county, with the results of productivity shocks in a world with no commuting. These files are generated in the cmlee_main.nb file and used to produce figure C.7.

Subfolder "county shocks 5pct\"

This subfolder contains 3111 .csv files, one per county, with the results of productivity shocks in a world with no commuting. These files are generated in the cmlee_main.nb file and used to produce figure 2, 3, C.2, C.7, and Table 2.

Subfolder "county shocks 5pct endogDef\"

This subfolder contains 3111 .csv files, one per county, with the results of productivity shocks in a world with partially locally owned and partially nationally owned land. These files are generated in the cmlee_ED.nb file and used to produce figure B.8.

Subfolder "CZ calibration shocks 5pct\"

This subfolder contains 709 .csv files, one per commuting zone, with the results of productivity shocks for the model calibrated at commuting zone level. These files are generated in the cmlee_CZ.nb file and are used to produce figures C.11 and C.13 and table C.6.

Subfolder "Saiz price exf county shocks 5pct\"

This subfolder contains 3111 .csv files, one per county, with the results of productivity shocks in a model with positive elastic land supply. They are generated in the cmlee_main.nb file and are used to produce figure 3 and table C.2.

Subfolder "Spatially correlated\"

This subfolder contain 3 .csv files with the result of three separate counterfactual spatially correlated productivity shocks. They are generated in the cmlee_main.nb file and are used to produce figure C.4 (spCorr_mfg_BLS.csv), figure C.5 (spCorr_nonMfg_BLS.csv), and figure C.6 (spCorr_both_BLS.csv).

amenitiesBni.csv

This file contains a matrix of amenities that exactly rationalize commuting flows. The file is produced by cmlee_main.nb and used to generate results in footnote 25 in the main paper.

comparisonCFS_shares_rescaled_fullTradeDeficits.csv

This file contains model-generated equivalent CFS trade shares and actual CFS shares. The file is produced by cmlee_main.nb and used to generate figure B.2.

CZ_data_for_reduced_form_fullTD.csv

This file contains partial elasticities of employment holding fixed commuting, migration, and/or goods trade linkages and commuting zone calibrated productivities. The file is produced by cmlee_CZ.nb and used to generate figure B.7 and tables C.6 and C.7.

CZ_decreaseMedian.csv

This file contains the result of one counterfactual simulation reducing commuting costs by the median nationwide change in the model calibrated at commuting zone level. The file is produced by cmlee_CZ.nb and used to generate figures C.14 and C.15.

data_for_reduced_form_fullTD.csv

This file contains partial elasticities of employment holding fixed commuting, migration, and/or goods trade linkages and counties calibrated productivities. The file is produced by cmlee_main.nb and used to generate table B.2 and parts of other intermediate results.

decreaseMedian.csv

This file contains the result of one counterfactual simulation reducing commuting costs by the median nationwide change in the model calibrated at county level. The file is produced by cmlee_main.nb and used to generate figure C.9.

land_prices_fullTD.csv

This file contains model calibrated values of land. The file is produced by cmlee_main.nb and is used to generate figure C.1.

TradeLess20InitialEq.csv

This file contains the results of one counterfactual simulation reducing trade costs by 20% in the initial equilibrium. The file is produced by cmlee_main.nb and is used to generate figures C.10 and C.11.

TradeLess20Nocomm.csv

This file contains the results of one counterfactual simulation reducing trade costs by 20% in a counterfactual economy where commuting is not allowed. The file is produced by cmlee_main.nb and is used to generate figures C.10 and C.11.

bilateral distances.dta

This file contains bilateral distances between county centroids, for all county pairs. It is used to produce figure B.5 and parts of other intermediate results.

cfs2007.dta

This file contains bilateral trade flows at Commodity Flow Survey (CFS) area level for 2007. It is used to produce figures B.1 and B.2.

CMLEE.dta

This file contains for all county pairs data on commuting flows (after removing business trips), workplace and residence employment, average wage by place of work, fraction of commuters out of residence, fraction of commuters from outside the workplace, bilateral distance, and commuting zone codes. The original data sources are American Community Survey 2006-2010 and BEA county profile, as described in the paper. It is used to produce Table 1, B.1, B.3, and figures B.5, B.12, C.12 and parts of other intermediate results.

commuting_margins.dta

This file contains for all county pairs data on commuting flows (after removing business trips) and workplace and residence employment for 1990 and 2007. The original data sources are U.S. Census, 1990 and American Community Survey 2006-2010, as described in the main paper. It is used to produce tables C.4 and C.5.

commuting_zones_cbsa.dta

This file contains correspondences between county and year 2000 commuting zones from the U.S. Census. It is used to produce parts of intermediate results.

crime_county_DOJ.dta

This file contains violent crimes by county in 2007 from the Department of Justice, Federal Bureau of Investigation, as described in the paper. It is used to produce results discussed in footnote 25 in the main paper.

cz_90_00.dta

This file contains share various measures of commuting openness at commuting zone level for 1990 and 2000. The original data source is U.S. Census 1990 and 2000, as described in the paper. It is used to produce figure B.4.

cz_bilateral_distances.dta

This file contains bilateral distances between commuting zone centroids, for all pairs. It is used to produce figure B5 and parts of other intermediate results. It is used to produce figure B12 and parts of other intermediate results.

ghm_longform.dta

This file contains data for the Million-dollar plants exercise. The original data sources are Greenstone, Hornbeck and Moretti (2004), U.S. Census 1990, and relevant years of BEA county profiles, as discussed in the main paper. It produces tables 3, 4, C.3, figures 4, 5 and C.8, and results in footnote 36 in the main paper.

house_values.dta

This file contains data on median housing values from American Community Survey 2009-2013 as described in the paper. It is used to produce figure C.1.

hri.dta

This file contains values for the commuting-based Head and Ries (1990) index, as described in the paper. The original data sources are U.S. Census 1990 and American Community Survey 2006-2010. It is used to produce statistics reported in Section 5.

lambda-nn-n-all-yrs.dta

This file contains share of residents working in the same county for 1960-2007. The original data sources are various years of U.S. Census and American Community Survey, as described in the paper. It is used to produce figures 1 and B.3.

Rappaport Sachs BEA-level.dta

This file contains data on distance of a county from the ocean or a navigable river from Rappaport and Sachs (2003), where counties are aggregated at BEA level, as described in the paper. It is used to produce results for footnote 21 in the main paper.

saiz_pop_elasticities.dta

This file contains land supply elasticities for each county starting from data in Saiz (2010) as described in the paper. It is used to produce Figure 3 and Table C2.

do files

The folder contains only one file, “CMLEE Replication.do”. This do-file replicates the entire set of tables and figures in the paper and the web appendix. It draws on data files in the folder “data\” and its subfolders. All figures are saved in the “figures\” folder.

figures

This folder contains all figure files generated by CMLEE Replication.do. Figures are named as in the paper.

mathematica files

This folder contains Wolfram’s Mathematica notebooks necessary to generate counterfactual simulations. These files collectively draw on .csv files in the “simulation inputs\” directory below. These files collectively produce a set of .csv files in the “data\” folder and its subfolders, as detailed below:

cmlee_main.nb

This Mathematica notebook generates results for all the counterfactual simulations in the paper, excluding calibration at commuting zone (see cmlee_CZ.nb) and simulation with partially locally owned and partially nationally owned land (see cmlee_ED.nb). In particular, it generates the data for following counterfactual simulations (the letter sub-points here correspond to the letter sections under “Counterfactual Simulations” in the notebook):

a. 5% productivity shock to each county:
    - Data for section 3;
    - 3111 counterfactual simulation results, one per county, saved in “data\county shocks 5pct\”;
b. 5% productivity shock to each county in a world with no commuting:
    - data for figure C.7;
    - 3111 counterfactual simulation results, one per county, saved in “data\county shocks 5pct\”.
c. Spatially correlated shocks:
    - 3 files saved in “data\spatially correlated\”;
    - spCorr_mfg_BLS.csv: data for Figure C.4;
    - spCorr_nonMfg_BLS.csv: data for figure C.5;
    - spCorr_boht_BLS.csv: data for figure C.6.
d. 5% productivity shocks to each county, positive land supply elasticity:
    - data for figure 3 and table C.2;
    - 3111 counterfactual simulation results, one per county, saved in “data\Saiz price exf county shocks 5pct\”.
e. Reducing commuting costs using HRI changes 1990-2007:
    - data for figure C.9, welfare impact of varying changes in commuting costs;

- 1 file, “data\decreaseMedian.csv”; welfare results for the other changes are computed directly in the notebook.

f. Reducing trade costs with and without commuting:
   - 2 files saved in “data\”;
   - TradeLess20InitialEq.csv: data for figures C.10 and C.11, trade cost reductions in the initial equilibrium
   - TradeLess20Nocomm.csv: data for figures C.10 and C.11, trade cost reductions in a world with no commuting.

This Mathematica notebook also generates files to be used in further Stata processing. These files are all generated in the “data\” subfolder and described there. The files are:

- amenitiesBni.csv;
- comparisonCFS_shares_rescaled_fullTradeDeficits.csv;
- data_for_reduced_form_fullTD.csv;
- decreaseMedian.csv;
- land_prices_fullTD.csv.

cmlee_CZ.nb

This Mathematica notebook generates results for all the counterfactual simulations in the paper with the model calibrated at commuting zone level. In particular, it generates the data for following counterfactual simulations (the letter sub-points here correspond to the letter sections under “Counterfactual Simulations” in the notebook):

a. 5% productivity shock to each commuting zone:
   - data for section C.11, table C.6, figure C.13;
   - 709 counterfactual simulation results, one per commuting zone, saved in “data\CZ calibration shocks 5pct\”;

b. Reducing commuting costs using HRI changes 1990-2007
   - data for figure C.14 and C.15;
   - 1 file, “data\CZ_decreaseMedian.csv”.

This Mathematica notebook also generates the file “CZ_data_for_reduced_form_fullTD.csv” in the “data\” folder, and is described there.

cmlee_ED.nb

This Mathematica notebook generates results for the counterfactual simulations with partially locally owned and partially nationally owned land. It generates the data for following counterfactual simulations (the letter sub-points here correspond to the letter sections under “Counterfactual Simulations” in the notebook):

a. 5% productivity shock to each county:
   - Data for section B.17;
   - 3111 counterfactual simulation results, one per county, saved in “data\county shocks 5pct\endogDef\”.

simulation inputs

This folder contains csv data files used to produce the simulation results. Only the Mathematica codes in the "Mathematica files\" folder use files in this folder. The folder has the following files:

bea_cfs_list.csv

Correspondence between counties and cfs area. The first column is a CFS identifier, the second is a county identifier (different from the fips code) from 1 to 3,111. Counties are always ordered by fips code, so county number n in this data is the n-th county ordered by fips code.

bilateral_cfs_trade.csv

Sales from a CFS origin (row) to any CFS destination (column), Commodity Flow Survey 2007.

commuting_flows.csv

County-level commuters residing in a row location and working in a column location, in number of people

county_names.csv

Files with county fips codes, state fips codes, county names and state names.

distances_allpairs.csv

Matrix of bilateral distances between county centroids, all possible pairs.

distances_onlytrading.csv

Matrix of bilateral distances between county centroids; when two counties are in CFS that have no trade flows, the distance is set to 10^9 to shut down trade in the initial equilibrium (see paper for more details).

land_area.csv

Counties land area, Census Gazetteer 2000.

mfg_shares.csv

County-level share of employment in manufacturing, County Business Patterns 2007.

residentialEmp.csv

County-level number of residents in a residential location; obtained inverting commuting flows after correcting for business trips (see paper for more details).

residentialWage.csv

County-level average yearly wage of a job by residence, USD, consistent with commuting flows and workplace wages (see paper for more details).

resIncome.csv

County-level total residential income, in million, consistent with residential employment (see paper for more details).

saiz_price_elasticities.csv

County-level vector of land elasticities: the maximum in Saiz data for counties with no data, zero for internal counties in multi-counties MSAs and the estimated Saiz elasticity for single-counties MSA and outlying counties in multi-county MSA.

workplaceEmp.csv

County-level employment (number of jobs) from BEA county profiles, 2007.

workplaceWage.csv

Average yearly wage of a job by workplace, USD, from BEA county profiles, 2007.

temp

This directory gathers temporary files while "CMLEE Replication.do" is running. It is empty at the end of the code.

Replication Instructions

All figures and tables can be replicated by running "CMLEE_replication.do" file. The do-file is organized in sections. All figures will be saved in the "figures\" folder with the name they have in the main paper and in the web appendix. The current path is set at the beginning of the do file.

All counterfactual simulations can be replicated by running the relevant Mathematica notebook file, as described above. Each notebook file has four sections: Setup, Import Data; Functions; Data Preparation; Counterfactual Simulations. The first three sections have to be run before any of the subsections in the Counterfactual Simulations. Each subsection in the Counterfactual simulation can be run independently. The current path needs to be set in the Setup, Import Data >> Setup subsection.