

## Empirical Exercise 4: Gravity for manufacturing vs. services

October 5, 2021

Instructor: Marc-Andreas Muendler [muendler@ucsd.edu](mailto:muendler@ucsd.edu)  
Teaching Assistant: Jacob Orchard [jdorchard@ucsd.edu](mailto:jdorchard@ucsd.edu)

**Due date and time: October 15, 5pm**

### Inputs and products

Please use Stata (any version) for your work. You may call any other software from within Stata (including Python, R, Perl, and system-level commands). Please base your analysis on the following files

ITPD-E by USITC      `itpd.dta`  
Gravity data by CEPII    `cepii-gravdata.dta`

in the online data folder at <https://econweb.ucsd.edu/muendler/teach/21f/435/gen>.

You may find the code from lecture a useful reference: `lec04.do` in the online lecture folder <https://econweb.ucsd.edu/muendler/teach/21f/435/lec04>.

Please submit three products to [canvas.ucsd.edu](https://canvas.ucsd.edu) by the due time: (i) a file with results titled *ee04.pdf*, (ii) a log file titled *ee04.log*, and (iii) a Stata code file titled *ee04.do* (which may call other software). **Your log file must exhaustively document the steps from the above input files to the output of results.**

### Tasks

#### 1. Preliminaries.

- (a) Construct two data files, one for trade in goods and one for trade in services. Use the ITPD-E data by USITC, do not remove self trade, keep only the years 2000 and 2015, and aggregate the trade flows to the source country (exporter), destination country (importer), and year level (over all industries).
- (b) Use the gravity data by CEPII at the source country (origin), destination country, and year level, keep only the years 2000 and 2015, and extract the following variables: population-weighted distance, contiguity (shared border), and common official language.
- (c) Combine (merge) the ITPD-E trade data for goods and services, each type of trade in a separate file, with CEPII gravity data at the source country, destination country, and year level. (Make sure your log file reports the merge results.)

#### 2. Regressions.

- (a) Run two regressions for trade in goods, on a sample that includes cross-border and self trade as well as zero trade (whenever the estimator allows).
  - i. Specify an ordinary least-squares regression for log trade on log distance in 2000 and 2015 (letting the coefficient estimate for trade flows vary by year), on an indicator for contiguity, an indicator for common official language, log internal distance for self trade, and an indicator variable for self trade. Note that the three regressors for log distance need to be mutually exclusive and set to zero whenever not applicable. Cluster the standard errors at the level of source-destination pairs. Condition on source-year and destination-year fixed effects.

- ii. Specify a Poisson pseudo maximum likelihood regression for trade (not log trade) on log distance in 2000 and 2015 (letting the coefficient estimate for trade flows vary by year), on an indicator for contiguity, and an indicator for common official language. Note that the two regressors for log distance need to be mutually exclusive and set them to zero whenever a flow is self trade (given self-trade country fixed effects to be included). Cluster the standard errors at the level of source-destination pairs. Condition on source-year, destination-year, and self-trade country fixed effects. Make sure your sample includes zero trade flows.
  - (b) Run the same two regression specifications for trade in services, on a sample that includes cross-border and self trade as well as zero trade (whenever the estimator allows).
3. Interpretation.
- In one sentence, highlight one key difference between your findings for goods and for services.