

Empirical Exercise 5: The unitary distance coefficient

October 15, 2020

Instructor: Marc-Andreas Muendler muendler@ucsd.edu
Teaching Assistant: Jacob Orchard jdorchard@ucsd.edu

Due date and time: October 22, 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/20f/435/gen>.

You may find the code from the previous lecture 4 a useful reference: `lec04upd.do` in the online lecture folder <https://econweb.ucsd.edu/muendler/teach/20f/435/lec04>. **Note the update** from `lec04.do` to `lec04upd.do`.

Please submit three products to canvas.ucsd.edu by the due time: (i) a file with results titled *ee05.pdf*, (ii) a log file titled *ee05.log*, and (iii) a Stata code file titled *ee05.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) Follow Eaton and Kortum (2002, p. 1761-1762) and construct categories of (population weighted) distance in miles using the CEPII data (transforming the distance measure from kilometers into miles). Assign the distances into the following six categories: [0,375), [375,750), [750,1500), [1500,3000), [3000, 6000), and [6000,maximum) miles.
- (d) Combine (merge) the ITPD-E trade data for goods and services, each type of trade in separate a file, with the CEPII gravity data (including the distance categories) at the source country, destination country, and year level. (Make sure your log file reports the merge results.)

2. Regressions.

- (a) Just as in Empirical Exercise 4, 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). Decide which of the six distance categories you want to make your reference category (e.g. [0,375) miles) and omit it from all specifications.

- i. Specify an ordinary least-squares regression for log trade on five of the six constructed distance category indicators, 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.
Hint: In your regression, make sure you specify that there is no constant, otherwise one of the six distance categories would be omitted.
 - ii. Specify a Poisson pseudo maximum likelihood regression for trade (not log trade) on five of the six constructed distance category indicators, 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.
Hint: In your regression, make sure you specify that there is no constant, otherwise one of the six distance categories would be omitted.
- (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.

- Pick one of the two specifications for one of the two types of trade. For the five out of six coefficient estimates, use the formula $(\exp\{\beta_{\text{indic. coeff.}}\} - 1) \times 100\%$ to compute the percentage response of trade to the respective distance category.
- In one sentence, comment on whether you think the findings support the assumption of a constant elasticity of trade with respect to distance (unitary or not) in goods and services trade.

References

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica*, 70(5): 1741–79.