International Trade and Globalization

Solutions To Problem Set Three

MATTEO BERTASIO [1]

[1]In collaboration with Filippo Boggetti, Luca Bottigelli.

Exercise 1-Estimating gravity equations with reghdfe

Download the gravity dataset from CEPII as provided by Head and Mayer

(see httpa://sites.google.com/site/hiegravity/data-sources, choose the "Lighter dataset").

1. **What percentage of trade flow observations are equal to zero? Is this consistent with the Eaton-Kortum model?**

In the EK model we know that the probability that country j export to country n is:

$$\pi_{nj} = \frac{T_j \left(w_j d_{nj}\right)^{-\theta}}{\Phi_n}$$

So as long as we are not in limit cases (like iceberg costs going up to ∞) the value is always positive (since T_j the location of the Frechet distribution is positive). Hence zero flows are **NOT consistent** with EK unless we introduce fixed costs (and not only icebergs).

We have 495098 observations equal to 0 out of 1204671 obs.

2. **Estimate a naive gravity model that only includes year fixed effects. That is, estimate**

$$\ln Y_{nj,t} = \alpha_t + \beta_1 \ln Y_n + \beta_2 \ln Y_j + \beta_3 \ln d_{nj} + \varepsilon_{nj,t}$$

where α_t are year fixed effects. Comment your results. How do they compare to Tinbergen? Note: I recommend you download the reghdfe package on Stata. You will not need it for this regression, but later regressions have around 20'000 fixed effects, so it will be useful.

We don't use STATA and instead use R as software. We find coefficients to bee $\beta_1 = 1.03$, $\beta_2 = 0.86$, $\beta_3 = -1.22$.

3. **When you took bogs to estimate the previous equation, what happened with zero trade flows? Do you think this is a problem?**

Note that by using R we had to manually drop zero obs before applying logs. This is not the best method because a large part of obs (almost 41%) are 0. We could find more complicated estimators to make sure of not loosing obs.

4. **Now, estimate a gravity model with multiateral resistance terms (if., year-origin and year-destination fixed effects). Is your estimate for the elasticity of tade with respect to distance different than in Question 2?**

Now distance have an higher absolute value, becoming -1.47

5. **Introduce some additional variables in to the gravity equation of Question 4: a dummy for a shared border, a dummy for a regional trade agreement, and a dummy for a shared language. How does the distance elasticity change?**

Now the distance coefficient is -1.267. But this is **NOT** longer the elasticity. Recall in fact that to consider a coefficient a partial derivative we need to have that all others regressor are orthogonal to distance. And now shared border clearly is correlated with distance.

6. **Finally, reduce the dataset to include only members of the OECD. Which share of observations are zeros now? Re-estimate the equation in Question 5, and report the distance elasticity.**

Now we have a coefficient of log distance equal to -1.1 and zero values are 3% of obs.
IMPORTANT NOTE: Chile (CHL) and Colombia (COL) despite being OECD today, they were not in the highest year recorded ion the dataset 2006 and hence have been excluded.

Exercise 2 - Gravity and the ACR formula in an Armington model

Consider again the Armington model of PS 1. The world has I countries and I goods. Each country can only produce one good, which has the same index as the country (e.g., country 1 can only produce good 1 , but not good $2, 3, \ldots$ and I). One unit of the country's good can be produced with one unit of the country's labour.

In each country, consumers have identical CES preferences, given by

$$U\left(c_1, c_2, \ldots, c_1\right)-\left(\sum_{i=1}^{I}\left(c_i\right)^{\frac{\varepsilon-1}{\tau}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$$

and supply inelastically 1 unit of labour. The total population of country n is denoted L_n. Furthermore, there are iceberg trade costs: for delivering one unit of a good from country j to country n, it is necessary to ship d_{nj} units of the good, with $d_{nj} \geq 1$. As usual, $d_{nn}=1$.

1. **Due to the Armington structure, country n must buy good i from country i (and good i is also the only thing it buys from country i). Using this observation, derive an expression for the total spending of country n on goods from country i, $Y_{ni}=p_n(i)c_n(i)$, as a function of $p_n(i), P_n, w_n$ and L_n.**

We know that price of good j is simply given by:

$$p_j=\frac{w_j}{z_j} \rightarrow p_j=w_j$$

Where the first equality comes from the optimality conditions of firms and the fact that each country produce only one good. The arrow comes from the fact that we have assumed that technology is constant and equal to 1 across countries. From consumer optimization (recall that all consumers have the same preferences) we know that the consumption of good j in country n is:

$$c_{n,j}^{*}=\arg\max_{c_j}\left(\sum_{i=1}^{I}\left(c_i\right)^{\frac{\varepsilon-1}{\tau}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$$

subject to

$$\sum_{i=1}^{I} p_i d_{n,i} c_{n,i} \geq w_n L_n$$

FOCs read:

$$c_{n,i} = \left(\frac{d_{n,i}p_i}{d_{n,j}p_j}\right)^{-\varepsilon} c_{n,j} \ \forall i \in I$$

Using the budget constraint (which is binding by Walras' law) we get that:

$$c_{n,j} = \frac{w_n L_n}{\sum_{i=1}^{I}(d_{n,i}w_i)^{1-\varepsilon}}(d_{n,j}w_j)^{-\varepsilon}$$

where we used the price determination we explained at the beginning of the section. Now it becomes trivial to find the total expenditure as:

$$p_j c_{n,j} = \frac{w_n L_n}{\sum_{i=1}^{I}(d_{n,i}w_i)^{1-\varepsilon}}(d_{n,j}w_j)^{1-\varepsilon} \tag{1}$$

which is what we needed to find.

2. **Use this expression to show that the Armington model generates a gravity equation.**

Take equation (1) and sum over the countries n to get that[1]:

$$\sum_{n=1}^{N} p_j c_{n,j} = X_j = w_j^{1-\varepsilon} \sum_{n=1}^{N} \frac{w_n L_n}{\sum_{i=1}^{I}(d_{n,i}w_i)^{1-\varepsilon}}(d_{n,j})^{1-\varepsilon}$$

From which we immediately get that (plugging this expression in (1)):

$$X_{n,j} = \frac{X_n X_j}{\sum_{n=1}^{N} \frac{w_n L_n}{\sum_{i=1}^{I}(d_{n,i}w_i)^{1-\varepsilon}}(d_{n,j})^{1-\varepsilon}} \frac{d_{n,j}^{1-\varepsilon}}{\sum_{i=1}^{I}(d_{n,i}w_i)^{1-\varepsilon}}$$

which is a gravity equation. De facto we can take logs to obtain an estimable equation:

$$\ln X_{n,j} = \ln X_n + \ln X_j - (\varepsilon - 1)\ln d_{n,j} + \Lambda_j + \Gamma_n$$

N.B. This formulation includes the multilateral resistance terms. Namely we have to account for:

[1] Please note that given the structure of the Argminton model we have that $w_n L_n = X_n$ to be consistent with the LN

1. $\Lambda_j = \ln \sum_{n=1}^{N} \frac{w_n L_n}{\sum_{i=1}^{I}(d_{n,i}w_i)^{1-\varepsilon}} \left(d_{n,j}\right)^{1-\varepsilon}$ This expression captures the real-resource cost of global market access for exporter j it scales how much real income is reachable by j, factoring in iceberg trade frictions. In other words, it tells us: "How much of world real income would be absorbed by trade costs if j exported to everyone at equilibrium prices."

2. $\Gamma_n = \ln \frac{1}{\sum_{i=1}^{I}(d_{n,i}w_i)^{1-\varepsilon}}$ which instead computes the attractiveness of country n as an importer. De facto we are computing a price index at equilibrium prices with iceberg costs from all the countries to country n.

Estimation: this structure force us, if we have data to add at least origin and destination fixed effects. If we have panel data than we can add time fixed effects but recall that adding them directly creates a problem of multicollinearity. Hence the clever way is to add year destination and year origin fixed effect if we want to estimate the gravity equation.

A NOTE ON MULT. REGRESSIONS: Recall that the partial regression theorem tells us that $\varepsilon - 1$ is the "effect of trade distance on trade flows net of variance explained from other regressor" so when we want to interpret coefficient of a multivariate regression as partial derivatives we need to be careful. In this case thanks also to the inelastic labor supply we have that the trade distance is orthogonal to other regressors so no problems, but usually wee need to be careful.

3. **Deduce an expression for the share of country n 's national income spent on goods from country i,π_{ni}**

$$\pi_{n,i} = \frac{p_i c_{n,i}}{w_n L_n} = \frac{\left(d_{n,i}w_i\right)^{1-\varepsilon}}{\sum_{i=1}^{I}\left(d_{n,i}w_i\right)^{1-\varepsilon}}$$

4. **From your answers to the previous questions, deduce the Armington version of the ACR formula,**

$$\boldsymbol{\frac{w_n}{P_n} = \pi_{nn}^{-\frac{1}{\varepsilon-1}}}$$

Note that the share of goods produced and consumed in the country is simply:

$$\pi_{n,n} = \frac{\left(w_n\right)^{1-\varepsilon}}{\sum_{i=1}^{I}\left(d_{n,i}w_i\right)^{1-\varepsilon}}$$

And recalling the optimal price we get that:

$$\frac{w_n}{\mathbb{P}_n} = \pi_{nn}^{-\frac{1}{\varepsilon-1}}$$

The ARC formula is hence:

$$\frac{(w_n/\mathbb{P}_n)_{\text{Trade}}}{(w_n/\mathbb{P}_n)_{\text{Autarky}}} = \left(\frac{\pi_{nn}^{\text{Trade}}}{\pi_{nn}^{\text{Autarky}}}\right)^{-\frac{1}{\varepsilon-1}}$$

5. **Are there any differences between this gravity equation or ACR formula and the ones in the lecture notes? Interpret.**

The exponent of the local consumption ratio, in the classical ACR is $-\frac{1}{\theta}$ where θ measure the dispersion of productivity. Here we don't have this parameter and the exponent is proportional to the elasticity of substitution across goods *i.e.* if $\varepsilon \to \infty$ goods are perfect substitute and there is no advantage of trade (you can think at the **perceived** production possibility frontier to be the same under autarky or under trade).

6. **Suppose you want to calculate the gains from trade for a particular country. You use a gravity equation to estimate the trade elasticity, and then the ACR formula to calculate gains from trade. Does it make any difference for this whether you believe that the underlying model of the world is Armington or Eaton-Kortum?**

The result are invariant to the type of model we are considering. This is true since there is an internal coherence. If we believe the world is similar to Argminton, then when we estimate the gravity equation:

$$\ln X_{n,j} = \ln X_n + \ln X_j - (\varepsilon - 1)\ln d_{n,j} + \Lambda_j + \Gamma_n$$

we get that $(\varepsilon - 1) = k$ while wheen using EK we would estimate:

$$\ln X_{jn} = \ln X_n + \ln X_j - \theta \ln d_{jn} - \ln \Phi_n - \ln \chi_j.$$

finding that $\theta = k$. Now the two ARC formulas would be equal since in EK:

$$\frac{(w_n/P_n)_{\text{Trade}}}{(w_n/P_n)_{\text{Autarky}}} = (\pi_{nn})^{-\frac{1}{k}}$$

since we have already estimated by the gravity eq. the productivity dispersion. Similarly in the Argminton model:

$$\frac{(w_n/\mathbb{P}_n)_{\text{Trade}}}{(w_n/\mathbb{P}_n)_{\text{Autarky}}} = \left(\frac{\pi_{nn}^{\text{Trade}}}{\pi_{nn}^{\text{Autarky}}}\right)^{-\frac{1}{k}}$$

and hence we have full equivalency.

Exercise 3-Simulating Brexit

Download the code and dat a for Costinot and Rodriguez-Clare's "Trade Theory with Numbers" paper, at https://economics.mit.edu/files/9216. This folder contains some data files, and a series of Matlab files which can be used to calculate the consequences of tariff and trade cost changes. Costinot does this for different models, including the basic Eaton-Kortum model, and models with multiple sectors and intermediate inputs similar to those we have seen.

Basic instructions Download and save all files in the same folder. Then, open MATLAB run the $Step_{01}$, $Step_{02}$, $Step_{03}$ and $Step_{04}$ (in this order). This reads in the data and does some preliminary work needed to solve the model.

We focus on the first part of Table 4.2 of the Costinot pape (reproduced as Table 4 in the LN), where we assume that initially, tariffs are 0 between all countries of the world, and calculate the welfare changes which would occur if the United States imposed a unilateral import tariff of 40% on every other country. Column 1 of the table shows the consequences of this change in the basic EatonKortum model, and Column 4 the consequences in a model with multiple sectors and intermediate inputs. Download the $Step\_10\_PS3Version$ file available on Blackboard (my simplified version of Costinot's Step 10), and save it in the same folder as the other files. This file produces as final output Columns 1 and 4 of Table 4.2 .

In the program, tariffs are stored in two tensors (three-dimensional matrices). "tij_s3D" contains all tariffs before the trade policy change, and $tijs\_p3D$ contains all tariffs after the trade policy change. These tensors are build such that their (i, j, s) element gives you the tariff that country i faces when exporting to country j in sector s. Countries are ordered alphabetically just as in Table 4.2 of the paper. So, in the first row, third column, you will find the tariff that Belgium imposes on Australian imports, while in the third row, first column, you will find the tariff that Australia imposes on Belgian imports. For the work you will do, it is useful to note that Great Britain is Number 13 , and the United States Number 33.

Your task in this code is only to change the $tijs_p3D$ tensors, according to the trade policy scenarios we want to consider: you should leave all other aspects of the code unchanged.

Reproducing Costinot's results

1. **Change the code in order to reproduce Costinot's results in Columns 1 and 4 of Table 4.2. That is, change$tijs_p3D$ such that the US charges 40% tariffs on all other countries of the world (except, of course, on itself), and check that you are getting the same results as in the paper.**

We can reproduce the table in the LN computing the % change.

Table 1: Replication

Country	(1)	(4)	Country	(1)	(4)
AUS	-0.10%	-0.28%	IRL	-0.91%	-1.58%
AUT	-0.09%	-0.13%	ITA	-0.07%	-0.07%
BEL	-0.16%	-0.26%	JPN	-0.10%	-0.11%
BRA	-0.10%	-0.16%	KOR	-0.22%	-0.34%
CAN	-1.20%	-2.28%	MEX	-1.08%	-1.67%
CHN	-0.22%	-0.46%	NLD	-0.22%	-0.34%
CZE	-0.05%	-0.08%	POL	-0.04%	-0.08%
DEU	-0.16%	-0.20%	PRT	-0.06%	-0.09%
DNK	-0.20%	-0.26%	ROM	-0.03%	0.01%
ESP	-0.06%	-0.07%	RUS	-0.03%	-0.13%
FIN	-0.09%	-0.10%	SVK	-0.05%	-0.04%
FRA	-0.09%	-0.13%	SVN	-0.05%	-0.10%
GBR	-0.16%	-0.31%	SWE	-0.15%	-0.19%
GRC	-0.08%	-0.06%	TUR	-0.03%	-0.02%
HUN	-0.13%	-0.17%	TWN	-0.46%	-0.76%
IDN	-0.09%	-0.14%	USA	0.21%	0.63%
IND	-0.16%	-0.25%	RoW	-0.49%	-0.97%

Brexit

2. **Next, use the code to simulate the following (very stylized) version of Brexit: change tijs _p3D such that after Brexit, tariffs between the United King dom and all EU member states go up to 10% in all sectors. Generate a table with your results. Which countries win and which countries lose from Brexit? Does taking into account intermediate inputs make a large difference?**

The following table summarize the effects. Note that we highlighted the countries that gain. Yes, changing makes a huge difference, as we can see the only country gaining if no intermediaries are added is "rest of the world" (RoW), and they become more in the case of intermediaries.

Table 2: Welfare Changes from Brexit

Country	Single Sec.	Intermediate	Country	Single Sec.	Intermediate
AUS	0.00%	0.01%	IRL	-0.06%	0.21%
AUT	-0.01%	-0.00%	ITA	-0.02%	-0.00%
BEL	-0.08%	-0.05%	JPN	0.00%	0.00%
BRA	0.00%	0.01%	KOR	0.00%	0.01%
CAN	0.00%	0.01%	MEX	0.00%	0.00%
CHN	0.00%	0.02%	NLD	-0.07%	-0.02%
CZE	-0.04%	-0.06%	POL	-0.02%	-0.01%
DEU	-0.02%	0.01%	PRT	-0.05%	-0.07%
DNK	-0.02%	0.02%	ROM	-0.02%	-0.02%
ESP	-0.02%	-0.01%	RUS	0.00%	0.02%
FIN	-0.02%	-0.01%	SVK	-0.04%	-0.06%
FRA	-0.02%	-0.01%	SVN	-0.01%	0.00%
GBR	-0.18%	-0.53%	SWE	-0.02%	-0.01%
GRC	-0.01%	-0.00%	TUR	0.00%	0.02%
HUN	-0.04%	-0.02%	TWN	0.00%	0.03%
IDN	0.00%	0.00%	USA	0.00%	0.01%
IND	0.00%	0.01%	RoW	0.01%	0.04%

3. **Have a look at the paper of Dhingra et al. Compare their results with the ones you are getting. Which aspects of the analysis could explain the differences between the two?**

In the paper they also add non tariffs barrier and hence when UK exit also this non tariffs barriers rise, making a huge difference with our estimates (in terms of welfare).

1 Exercise 2

Consider a version of the Spence signalling model seen in class where education can improve the productivity of the workers. There are two types of workers, the low type θ_l and the high type θ_h. The fraction of high-type workers in the population is λ. The low type worker is unable to take any education, while the high type worker can take education $e \in \{0,1\}$. The productivities for the two types of workers are given by: $Y_l(e) = \alpha_l$ and $Y_h(e) = \alpha_h + ke$, with $\alpha_l, \alpha_h, k > 0$. The high type's cost of education is $v > 0$.

1. **Characterize the pooling equilibrium. Under what condition, the pooling equilibrium exists?**

We denote by $\mu(0) = \mathbb{I}(\theta = \theta_h \mid e = 0)$ and $\mu(1) = \mathbb{I}(\theta = \theta_h \mid e = 1)$ the subjective beliefs of the employers. For any pooling equilibrium where $e^*(\theta_l) = e^*(\theta_h) = 0$, it must be the case that

$$\begin{aligned}&\{(e^*(\theta_l), e^*(\theta_h), w^*(0), w^*(1), \mu^*(0)) : \\ &e^*(\theta_l) = e^*(\theta_h) = 0, w^*(0) = w^*(1) = \lambda\alpha_h + (1-\lambda)\alpha_l, 0 \leq \mu^*(1) \leq \tfrac{\lambda(\alpha_h - \alpha_l)+v}{\alpha_h + k - \alpha_l}, \mu^*(0) = \lambda\Big\}\end{aligned}$$

Note that in this pooling equilibrium IR for both agents is trivially satisfied. Moreover since the low type cannot be educated it must be that his IC is trivially satisfied. We only need to check that the IC of the high type is satisfied, hence we need to prove that he is better off not pursuing an education:

$$\mu^*(1)(\alpha_h + k - v)(1 - \mu^*(1))(\alpha_l - v) \geq \lambda\alpha_h + (1-\lambda)\alpha_l$$

Which implies that the set of belies on the educated supporting a polling are:

$$0 \leq \mu^*(1) \leq \frac{\lambda(\alpha_h - \alpha_l) + v}{\alpha_h + k - \alpha_l}$$

2. **Characterize the separating equilibrium. Under what condition, the separating equilibrium exists?**

The separating is described by:

$$\begin{aligned}&\{(e^*(\theta_l), e^*(\theta_h), w^*(0), w^*(1), \mu^*(0)) : \\ &e^*(\theta_l) = 0, e^*(\theta_h) = 1, w^*(0) = \alpha_l, w^*(1) = \alpha_h + k, \mu^*(1) = 1, \mu^*(0) = 0\}\end{aligned}$$

Again in this models IR are satisfied for both agents and IC oof the low guy does not even exists since he cannot get education. Hence we need to check that there is no incentive for the high type to deviate. This is simply to check that:

$$\alpha_h + k - v \geq \alpha_l$$

3. **Illustrate the pooling and the separating equilibrium in the (e, w)-space.**