

Task 3. Does size matter? Estimating a Gravity Model of Trade

Presentation: Week 8

For this task, you are required to conduct a gravity model analysis. You should pick a country for which the UN Comtrade database provides trade data. You will collect data on the required variables and use a statistical software to estimate the parameters of the model.

Your research questions are:

- What are major impediments to international trade?
- What is the effect of country size and distance on international trade?
- Is the gravity model of trade consistent with other theories of trade?

The objectives of this assignment are:

- to make you acquainted with some key data sources of international trade;
- to let you estimate the coefficients of the gravity model using statistical software
- to let you interpret the estimated coefficients of the gravity model;
- to let you practice writing a clear and concise report describing your sources, the data manipulations you performed, and the results you obtained.

Start by reading:

- chapter 2 from Krugman et al. (2015);
- read the research paper from week 1 (check qmplus);
- in that paper, you will have reference to other studies which helps you to write your brief literature review.

You can use other software to do the statistical work, but I strongly encourage you to give R a try.

Background

The gravity model of trade (Krugman et al., 2015, p. 45) predicts that trade increases with the partner country's GDP: the bigger your trade partner's GDP, the more you trade with that partner country. The model also predicts that trade decreases with the distance to the partner country: the further away your trade partner, the less you trade with that partner country. This is the gravity equation:

$$Trade_{ij} = A * \frac{GDP_i^\alpha GDP_j^\beta}{D_{ij}^\gamma}$$

The coefficients α and β are elasticities : α measures by how many percent trade increases when the partner country's GDP increases by one percent (keeping other things constant); the same interpretation applies to β . γ measures by how many percent

trade decreases when the distance to the partner country increases by one percent (keeping other things constant). This can be seen by taking the partial derivatives of the gravity equation with regard to GDP_i and D_{ij} and solve for α and β . If you are not familiar with calculus (and you needn't be for this course) the same argument can be made by using the following approximation. You will remember from your macroeconomics classes that when three variables are in a multiplicative relationship:

$$Z = X \times Y$$

the following approximation holds for small percentage changes:

$$\text{percentage change in } Z \approx \text{percentage change in } X + \text{percentage change in } Y$$

($\approx$ means: is approximately equal to). Use this property to express the gravity equation in terms of percentage changes:

$$\begin{aligned} \text{percentage change in } Trade_{ij} &\approx \text{percentage change in } A \\ &+ \alpha \times (\text{percentage change in } GDP_i) \\ &+ \beta \times (\text{percentage change in } GDP_j) \\ &- \gamma \times (\text{percentage change in } D_{ij}) \end{aligned}$$

As A is assumed to be a constant, the percentage change of A is zero. If we keep the home country's GDP and distance constant, we obtain:

$$\text{percentage change in } Trade_{ij} \approx \beta \times (\text{percentage change in } GDP_j)$$

Solving for β yields:

$$\beta = \frac{\text{percentage change in } Trade_{ij}}{\text{percentage change in } GDP_j}$$

That is, β is the elasticity of trade with regard to the partner country's GDP. Similarly, if we keep the home country's GDP and the partner country's GDP constant and solve for γ we obtain:

$$\gamma = \frac{\text{percentage change in } Trade_{ij}}{\text{percentage change in } D_{ij}}$$

which is the elasticity of trade with regard to the distance from the partner country. In the theory section of your coursework, it suffices to briefly introduce the gravity model (cite Krugman et al. (2015) as your source) and point out that the coefficients β and γ are elasticities of trade with regard to GDP of the partner country and distance to the partner country. Don't repeat the explanation about the elasticities above but do refer to this document as your source. If necessary, review the concept of elasticity in your introductory economics textbook.

Data

The trade and GDP data should be for the same year.

Bilateral trade data are available from the United Nations Comtrade database :

<https://comtrade.un.org/>

1. Follow the link "Get trade data" In the form, for "Period (year)" select the most recent year for which GDP data are available (GDP data is from the World Bank, see next paragraph).
2. For "Reporters" select the country you want to study, maybe the US? You decide.
3. For "Partners" select "All".
4. For "Type of product" select "Goods" (we'll ignore exports of services).
5. For "Trade flows" select "Export" (we'll use exports rather than total trade (exports + imports) to estimate the gravity model, and limit ourselves to the top-10 export destinations).
6. Click the "Preview" button. If nothing shows up, you probably picked a year for which no data are available yet; change "Period (year)" to the year before you originally picked.
7. A table will show up showing all trading partners in alphabetical order.
8. By clicking the down-arrow button to the right of "Trade value (US\$)," the trading partners are ranked from high to low exports.
9. Take only the top-ten countries ("World" is not a country).
10. Some countries list "Areas NES (not elsewhere specified)", "Bunkers" (code 837), or "Free Zone" (code 838) as trading partner entities. Ignore these entities and work with the top-10 countries.

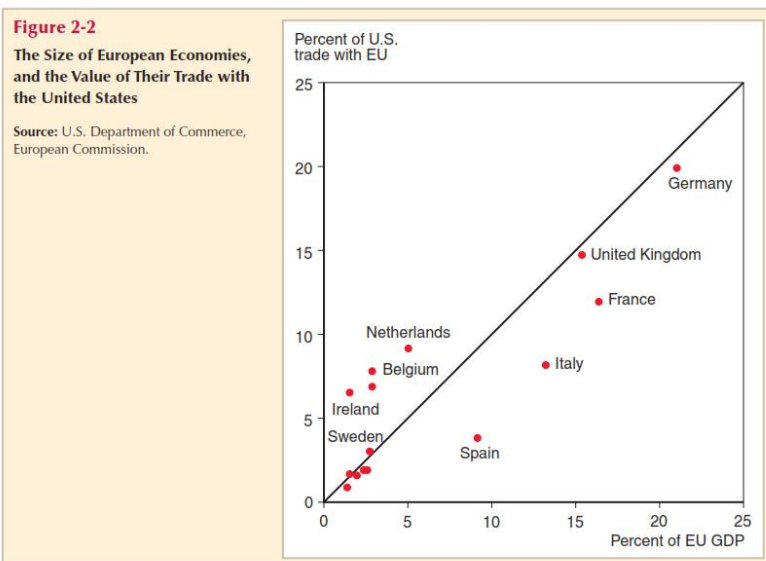
GDP data are available from World Bank: <https://data.worldbank.org/>

As trade is in current US\$, GDP has to be in current US\$ too. Go to the site and enter in the search window: **GDP (current US\$)**

Distance data are available from Centre d'Études Prospectives et d'Informations Internationales (CEPII) (2011) at:

http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=6

Register to get access to the data or drop me an email. The spreadsheet file `dist_cepii.xls` contains distances between countries. (Note that the downloaded file is a compressed file (`dist_cepii.zip`); double-click the zip file to expand it to `dist_cepii.xls`) The distances are in kilometers. The country codes are the three-letter ISO codes I just referred to. Pick one of the three distance measures provided. Report in the paper which one you picked and explain why.



Your paper should have a table with the relevant data for the top-10 export destinations: country, country ISO code, exports (millions of USD), GDP of destination country (millions of USD), and distance to destination country (km) (call it *table 1*, all numbers rounded).

- Put the table on a separate page at the end of the paper.
- Dumping some computer output on a separate page is not appropriate;
- The table should be numbered and have a title.
- The units of measurement should be mentioned, numbers should be rounded and right-aligned, and data sources should be documented in a note to the table.

What to do?

- Make a scatter plot (call it 'figure 1') that focuses on size (GDP) as a determinant of trade, like the one above from Krugman et al. (2015, figure 2-3 on p. 46). It is for EU.
- The scatter plot has GDP (as a percent of the GDP of all top-10 export destinations) on the horizontal axis, and exports (as a percent of the exports to all top-10 export destinations) on the vertical axis.
- Draw the scatter plot by hand on squared paper, or use a spreadsheet program or statistical software (like SPSS, Stata, Eviews, Gretl, or R).
- The axes should be labelled (variable and units of measurement) and each point should be labelled with the country code or country name. It's OK to handwrite the axis labels and point labels.
- Once done, discuss the scatter plot in the paper: Does trade increase with economic size, as the gravity model predicts? Are there anomalies? What are possible explanations of the anomalies?
- The scatter plot ignores the role of distance.

Multiple regression is a statistical technique that allows to measure the relationship of trade with multiple variables, for instance the partner countries' GDPs and distances (the gravity model). The gravity model is nonlinear in the parameters: α , β , and γ appear as powers of the variables. To estimate the coefficients using the Ordinary Least Squares method, we first have to make the model linear in the parameters. This can be done by taking logarithms:

$$\log(\text{Trade}_{ij}) = \log(A) + \alpha * \log(\text{In}Y_i) + \beta * \log(Y_j) - \gamma \log(D_{ij})$$

With the new variables $\ln T_{ij}$, $\ln Y_i$, and $\ln D_{ij}$ the model is linear in the parameters and the parameters can be estimated using the Ordinary Least Squares method.

Make sure that at least one of the students in your group understands multiple regression analysis and knows how to estimate regression coefficients. You can use a spreadsheet program (Excel) or a statistical software (like SPSS, Stata, Eviews, Gretl, or R).

In your essay, interpret the estimates of the slope coefficients β and γ : carefully explain what the numbers mean (hint: the coefficients are elasticities). Also report and discuss the *coefficient of determination*, a measure of the fraction of the variation in the left-hand side variable that is explained by the model.

Resources

Watch the following videos before or after you collected the data:

What is a Gravity Model?

<https://www.youtube.com/watch?v=e31BtKRjg1Y&list=PLTLaURNaDUAi74iK5Cz7brobWSTQjs0dw&index=1&t=117s>

How to run linear regression model in Excel

https://www.youtube.com/results?search_query=regression+in+excel

How to run linear regression model in R

https://www.youtube.com/results?search_query=regression+in+R