Gravity Models: Theoretical Foundations and estimation issues

6th ARTNet Capacity Building Workshop for Trade Research

Vientiane, 21 June 2010

Outline

- Theoretical foundations
  - Adding up and equilibrium conditions
  - Deriving gravity from new trade theory
- Empirical equations and some related estimation issues
  - Pitfalls to avoid
  - Using country and time-varying fixed effects

The Gravity Model: what it is?

- Econometric model (ex-post analysis)
- Workhorse in a number of fields. It has been used to analyze the impact on trade of GATT/WTO membership, RTAs, currency unions, migration flows, FDI between countries, disasters ...
- Initially, not based on a theoretical model

What explains its popularity?

- High explanatory power
- Data easily available
- There are established standard practices that facilitate the work of researchers

The gravity model: what is it? (2)

- Some intuition
- **Newton's Law** of Universal Gravitation:

$$F_{ij} = \frac{GM_iM_j}{D_{ij}^2}$$

F= attractive force; M= mass; D=distance; G = gravitational constant

- Usual **Gravity Model** specification similar to Newton's Law:

$$x_{ij} = \frac{KY_i^{\alpha}Y_j^{\beta}}{T_{ij}^{\chi}}$$

x_{ij}= exports from i to j;
Y= economic size (GDP, POP)
T =Trade costs

The gravity model: what is it? (3)

- Standard applications test the significance of additional variables. Some examples:

RTAs:

$$x_{ij} = \frac{KY_i^{\alpha} Y_j^{\beta} RTA_{ij}^{\delta}}{T_{ij}^{\chi}}$$

Currency union:

$$x_{ij} = \frac{KY_i^{\alpha} Y_j^{\beta} CU_{ij}^{\varepsilon}}{T_{ij}^{\chi}}$$

New barrier (SPS):

$$x_{ij} = \frac{KY_i^{\alpha} Y_j^{\beta}}{T_{ij}^{\chi} SPS_{ij}^{\eta}}$$

Theoretical foundation of the gravity equation

Consider an importing country j and exporting country i. We can derive the gravity equation in several basic steps:

Step 1: The expenditure (budget) allocation identity for j is given by:

$$x_{ij} = s_{ij} E_j \qquad \sum_i s_{ij} = 1$$

- where x_{ij} are spending of j on i's products, E_j is total expenditure of j and s_{ij} is the share of j's spending on goods from i
- x_{ij} can be interpreted as j's imports from I
- The shares sum to 1 if spending by j on j's products are included

...Theoretical foundation (2)

- The "trick" is to show that the share can be written in the form:

Step 2: Share can be expressed as a function:

$$s_{ij} = \frac{A_i \phi_{ij}}{\Phi_j}$$

where A_i is some exporter specific variable (e.g. exporter's productivity), φ_{ij} is an importer-exporter specific variable (e.g. accessibility of the destination market to the exporter) and Φ_j is an importer-specific item (e.g. the degree of competition in the importing country).

...Theoretical foundation (3)

- Now, switching to the exporter (i), we have:

Step 3: General equilibrium condition (for i)

$$Q_i = \sum_j x_{ij} = A_i \sum_j \frac{\phi_{ij} E_j}{\Phi_j}$$

This states that total exports by i, including to itself, must sum up to the value of its production (Q_i). This implies:

$$A_i = \frac{Q_i}{\sum_j \frac{\phi_{ij} E_j}{\Phi_j}} \quad \Rightarrow \quad x_{ij} = \left(\frac{Q_i}{\sum_j \frac{\phi_{ij} E_j}{\Phi_j}}\right) \phi_{ij} \left(\frac{1}{\Phi_j}\right) E_j$$

Deriving gravity from different models

Krugman model:

- Each country (i) has N_i firms producing a variety each.
- Utility is Dixit-Stiglitz type which exhibits love of variety (δ is elasticity of substitution).
- There are "iceberg" trade costs
- Solving max problem, share takes the form:

$$s_{ij} = N_i (p_{ij} / P_j)^{1-\sigma} \quad \text{where} \qquad P_j = \left(\sum_{i=1}^{K} N_i p_{ij}^{1-\sigma}\right)^{1/(1-\sigma)}$$

Here p_{ij} is price of export of I in j's market and P_j is the CES price index

Deriving gravity from different models (2)

- Trade costs imply:

$$p_{ij} = \tau_{ij} p_i$$

Here p_i is the producer price in the exporting country and τ_{ij} represent trade costs, both natural (e.g. distance) and man-made (e.g. tariffs). Note that $\tau_{ij} \geq 1$. Expanding we have:

$$x_{ij} = N_i \left(\frac{p_i \tau_{ij}}{P_j}\right)^{1-\sigma} E_j$$

Deriving gravity from different models (3)

Applying general equilibrium condition gives us:

$$Q_i = \sum_j x_{ij} = N_i p_i^{1-\sigma} \sum_j \left(\frac{\tau_{ij}}{P_j}\right)^{1-\sigma} E_j$$

$$\Rightarrow \quad x_{ij} = \tau_{ij}^{1-\sigma} \left(\frac{Q_i}{\Omega_i P_j^{1-\sigma}}\right) E_j$$

Where:

$$s_{ij} = \tau_{ij}^{1-\sigma} \left(\frac{Q_i}{\Omega_i P_j^{1-\sigma}}\right) \qquad \Omega_i = \sum_j \left(\frac{\tau_{ij}}{P_j}\right)^{1-\sigma} E_j$$

Estimation of gravity equations

Caution

- Three things to be careful about (Baldwin and Taglione, 2006):

1. Use country fixed effects
   - Unobserved variables are correlated with error term
2. Do not average trade flows (use either imports or exports)
   - Gravity is an expenditure equation explaining the value of spending by one nation on the goods produced by another nation
3. Do not deflate trade flows or GDP
   - Gravity is an expenditure function allocating nominal GDP into nominal imports.

Why country fixed effects?

- Ω_i and $P_j^{1-\delta}$ are unobservable and correlated with the explanatory variables (trade costs). In typical regressions, the variables will get lumped together with the error term. Hence, trade costs will be correlated with the error term resulting in biased and inconsistent estimates.

$$\ln x_{ij} = \beta_0 + \beta_1 \ln GDP_i + \beta_2 \ln GDP_j + \beta_3 \ln \tau_{ij} + \beta_4 others$$

$$+ unobserved(\Omega_i, P_j^{1-\sigma}) + \mu$$

v=Error term in regression which is correlated with τ_{ij}

Solution: country-specific fixed effects

- Importer (exporter) dummy= it is a 0,1 dummy that denotes the importer (exporter)
  - They are used to proxy for:

    Ω_i and $P_j^{1-\sigma}$

  - They do not control for unobserved characteristics of pair of countries e.g. they have a RTA in place (need country pair fixed effects for this)

Gravity equation ...using fixed effects: cross section analysis

- In cross section analysis, using country fixed effects yields consistent estimates

$$\ln x_{ij} = \beta_0 + \beta_1 D_i + \beta_2 D_j + \beta_3 \ln \tau_{ij} + \beta_4 others + \mu$$

where D_i, D_j=country specific dummies

- There are 2m dummies, where m=number of countries. Total observations= m(m-1)

Note: It is impossible to estimate the coefficient for GDP and other country-specific variables. To remedy that:

$$\ln x_{ij} = \beta_0 + \beta_1 D_i + \beta_2 D_j + \beta_3 \ln \tau_{ij} + \beta_4 \ln(GDP_i * GDP_j) + \beta_5 others + \mu$$

Use product of GDPs as variable but now we assume common coefficient.

Example of importance of fixed effects

- An example where country fixed effects mattered is Rose (2004) which looked at the effect of WTO membership on trade

TABLE 1—BENCHMARK RESULTS

	Default	No industrial countries	Post 1970	With country effects
Both in GATT/WTO	−0.04	−0.21	−0.08	0.15
	(0.05)	(0.07)	(0.07)	(0.05)
One in GATT/WTO	−0.06	−0.20	−0.09	0.05
	(0.05)	(0.06)	(0.07)	(0.04)
GSP	0.86	0.04	0.84	0.70
	(0.03)	(0.10)	(0.03)	(0.03)
Log distance	−1.12	−1.23	−1.22	−1.31
	(0.02)	(0.03)	(0.02)	(0.02)
Log product real GDP	0.92	0.96	0.95	0.16
	(0.01)	(0.02)	(0.01)	(0.05)
Log product real GDP p/c	0.32	0.20	0.32	0.54
	(0.01)	(0.02)	(0.02)	(0.05)
Regional FTA	1.20	1.50	1.10	0.94
	(0.11)	(0.15)	(0.12)	(0.13)
Currency union	1.12	1.00	1.23	1.19
	(0.12)	(0.15)	(0.15)	(0.12)
Common language	0.31	0.10	0.35	0.27
	(0.04)	(0.06)	(0.04)	(0.04)
Land border	0.53	0.72	0.69	0.28
	(0.11)	(0.12)	(0.12)	(0.11)
Number landlocked	−0.27	−0.28	−0.31	−1.54
	(0.03)	(0.05)	(0.03)	(0.32)
Number islands	0.04	−0.14	0.03	−0.87
	(0.04)	(0.06)	(0.04)	(0.19)
Log product land area	−0.10	−0.17	−0.10	0.38
	(0.01)	(0.01)	(0.01)	(0.03)
Common colonizer	0.58	0.73	0.52	0.60
	(0.07)	(0.07)	(0.07)	(0.06)
Currently colonized	1.08	—	1.12	0.72
	(0.23)		(0.41)	(0.26)
Ever colony	1.16	−0.42	1.28	1.27
	(0.12)	(0.57)	(0.12)	(0.11)
Common country	−0.02	—	−0.32	0.31
	(1.08)		(1.04)	(0.58)
Observations	234,597	114,615	183,328	234,597
R^2	0.65	0.47	0.65	0.70
RMSE	1.98	2.36	2.10	1.82

Notes: Regressand: log real trade. OLS with year effects (intercepts not reported). Robust standard errors (clustering by country-pairs) are in parentheses.

from industrial countries.[16] The second uses only data after 1970. Finally, I add country-specific fixed effects to the benchmark equation at the extreme left of the table. The key result—that membership in the GATT/WTO is associated with an economically and statistically insignificant increase in trade—seems robust. Indeed, six of the eight coefficients are actually negative (though usually insignificantly so). The *largest* coefficient in Table 1 indicates that a pair of countries both in the GATT traded only ($\exp(0.15) - 1 \approx$) 16 percent more than a pair of countries outside the GATT. This is small compared to other effects (e.g., regional trade associations), the long-term growth of trade, intuition, and the hype surrounding the GATT/WTO.

To summarize, I have been unable to find evidence that membership in the GATT/WTO has had a strong positive effect on international trade. But since the GSP *is* associated with an

[16] I follow the IMF in defining countries as "industrial" if they have an IFS country code less than 200. No, the GSP coefficient is not a mistake; some (nonindustrial) Eastern European countries extended GSP preferences.

Gravity equation ...using fixed effects: Panel Data (1)

- *Suppose one has panel data. It is now possible to estimate the coefficient for GDP and other country-specific variables*

$$\ln x_{ijt} = \beta_0 + \beta_1 \ln(GDP_{it}) + \beta_2 \ln(GDP_{jt}) + \beta_3 \ln \tau_{ijt} + \beta_4 others + \beta_5 D_i + \beta_6 D_j + u_{ij}$$

Where D_i, D_j=country specific dummies

There are 2m such dummies

Total observations= m(m-1)T

- It is still not possible to estimate time-invariant country specific characteristics (eg. Island state, landlockedness)

Gravity equation ... using time-varying fixed effects: Panel Data (2)

- Potential problem: There may be variation over time of Ω_i and P_j (i.e. Ω_{it} and P_{jt})
- To address this, use time varying country specific dummies:

$$\ln x_{ijt} = \beta_0 + \beta_1 D_{it} + \beta_2 D_{jt} + \beta_3 \tau_{ijt} + \beta_4 others + u_{ij}$$

where D_{it}, D_{jt} =time-varying country specific dummies. There are 2mT such dummies.

Note: it is impossible to estimate the coefficient of GDP

Why not average bilateral trade as dependent variable?

- The model gives us:

$$\frac{x_{ij}+x_{ji}}{2}=(\frac{1}{2})(\tau_{ij}^{1-\sigma}(\frac{Q_i}{\Omega_i}\frac{E_j}{P_j^{1-\sigma}})+\tau_{ji}^{1-\sigma}(\frac{Q_j}{\Omega_j}\frac{E_i}{P_i^{1-\sigma}}))$$

- Taking the logs of both sides will not give us the usual form of the gravity equation.

Why not average bilateral trade as dependent variable?(2)

- If one wants to use averages, better take the logs first before averaging (one gets log of geometric mean)

$$\frac{\ln(x_{ij})+\ln(x_{ji})}{2}=\ln\sqrt{x_{ij}*x_{ji}}=\ln\sqrt{\tau_{ij}^{1-\sigma}\left(\frac{Q_i}{\Omega_i}\frac{E_j}{P_j^{1-\sigma}}\right)\tau_{ji}^{1-\sigma}\left(\frac{Q_j}{\Omega_j}\frac{E_i}{P_i^{1-\sigma}}\right)}$$

$$=\frac{\ln(\tau_{ij})+\ln(\tau_{ji})}{2}+\ln(Q_i)+\ln(E_j)+\frac{\ln(\Omega_i)+\ln(\Omega_j)}{2}+\ldots$$

- Still not easy to interpret the coefficients of a regression (based on averages) so preferable to stick to unidirectional flows

Problem with deflation

- Suppose we use real GDP and deflate trade flows:

$$(\frac{x_{ij}}{\Pi_{US}}) = \tau_{ij}^{1-\sigma}(\frac{Q_i}{\Pi_i}\frac{E_j}{\Pi_j})*(\frac{\Pi_i\Pi_j}{\Pi_{US}\Omega_i P_j^{1-\sigma}})$$

- In a gravity equation, the last term gets thrown into the disturbance term. Even if the econometrician corrects for the unobserved variables with country-specific dummies, there will be a danger of spurious correlation if the price deflators are correlated because of international inflation.

References

- Anderson J.E. (1979) “A Theoretical Foundation for the Gravity Equation”, American Economic Review, 69(1):106-116
- Anderson, J.E. and E. van Wincoop, 2003, “Gravity With Gravitas: A Solution to the Border Puzzle,” American Economic Review, 93(1), 170–192.
- Baldwin R. and D. Taglioni, 2006, “Gravity for Dummies and Dummies for Gravity Equations”, NBER Working Paper No. W12516.
- Feenstra, R.(2004) Advanced International Trade, MIT Press.
- Krugman, P (1980) “Scale Economies, Product Differentiation, and the Pattern of Trade”, American Economic Review, 70(5):950-959.
- Rose, A. K. (2004) “Do We Really Know that the WTO increases Trade?” American Economic Review, 94(1): 98-114.

Data sources

- Feenstra bilateral trade data (1962-2000) from UN Comtrade http://cid.econ.ucdavis.edu/data/undata/undata.html
- CEPII Distance and other geography variables http://www.cepii.fr/anglaisgraph/bdd/distances.html