The Gravity Model

James E. Anderson

Boston College and NBER

March 2011

Outline

Motivation

- Striking facts: enormous variation in economic interaction across space.
- Gravity $\Rightarrow$ flows increasing in size of markets, decreasing in distance, very good fit, stable coefficient estimates.
- Implication: unobservable trade costs are BIG.
- What makes a model successful or the facts it addresses interesting (to economists)? Gravity example.
  - Intellectual orphan long ignored by mainstream economists — e.g. Leamer and Levinsohn (1995) *Handbook*.
  - Progress with theoretical foundations $\Rightarrow$ adoption by the family (Feenstra, 2004) and continuing refinements.
  - Intuitive appeal empirically and theoretically $\Rightarrow$ popularity.

History

Analogy with Newton's Law of Gravity.

$$X_{ij} = Y_i E_j / d_{ij}^2$$

gives the predicted movement of goods or labor between i and j, X_{ij} as product of origin mass Y_i, destination mass E_j, divided by distance d_{ij}^2.

Looser analogy with mass and distance exponents estimated to be around 1 for each. Better fit with more proxies for resistance to trade such as common language, borders, etc.

First applied to migration in UK by Ravenstein (1884)

First applied to trade by Tinbergen (1962).

More Recent History

Bilateral frictions alone seem inadequate to explain X_{ij}; flow from i to j is influenced by

- resistance to i's shipments on its other possible destinations,
- resistance to shipments to j from j's other possible sources of supply;
- analogue to Newtonian gravity N-body problem.

Remoteness: $\sum_i d_{ij}/Y_i$ captures intuition that each country j has distance from all others that matters. Index is atheoretic and fails to deal with simultaneity of N-body analogy.

Multilateral resistance of structural gravity model is the solution.

Aggregate vs. Disaggregate

Gravity applied mostly to aggregate flows of goods or populations. But theory applies to disaggregated goods (e.g., by sectors) and factors (e.g., by skill level of migrants). Serious downward aggregation bias (Anderson and Yotov, 2010).

Gravity applied mostly to highly aggregated regions (e.g., nations). But theory applies to disaggregated regions.

Which aggregates? Usual gravity mass variables are GDP's for Ys and national incomes for Es. But gravity logic applies to gross flows, not value added or expenditure on final goods.

Frictionless Gravity

Action via mass variables: clear intuition, some not quite obvious results.

Implications from frictionless gravity point toward a structural theory. *Benchmark trade pattern of frictionless world helps draw inference of trade frictions.*

Smooth world: agents purchase goods in same proportions everywhere. Then

$$\frac{X_{ij}}{E_j} = \frac{Y_i}{\sum_i Y_i} = \frac{Y_i}{Y}.$$

Multiply both sides by E_j, yielding frictionless gravity:

$$X_{ij} = \frac{Y_i E_j}{Y} = s_i b_j Y, \tag{1}$$

where $b_j = E_j / Y$ and $s_i = Y_i / Y$.

Size Effects in Frictionless Gravity

1. Big producers have big market shares everywhere,
2. small sellers are more open in the sense of trading more with the rest of the world,
3. the world is more open the more similar in size and the more specialized the countries are,
4. the world is more open the greater the number of countries, and
5. world openness rises with convergence under the simplifying assumption of balanced trade.

Size Effects Formally

Implication 1: big producers have big market shares everywhere because the frictionless gravity prediction is that :

$$X_{ij}/E_j = s_i.$$

Implication 2, small sellers are more open in the sense of trading more with the rest of the world follows from

$$\sum_{i \neq j} X_{ij}/E_j = 1 - Y_j/Y = 1 - s_j$$

using $\sum_j E_j = \sum_i Y_i$, balanced trade for the world.

World Openness

Define world openness as the ratio of international shipments to total shipments, $\sum_j \sum_{i \neq j} X_{ij}/Y$. Dividing (1) through by Y , world openness is given by

$$\sum_j \sum_{i \neq j} X_{ij}/Y = \sum_j b_j(1 - s_j) = 1 - \sum_j b_j s_j.$$

Using standard properties of covariance and $\sum_j s_j = \sum_j b_j = 1$:

$$\sum_j \sum_{i \neq j} X_{ij}/Y = 1 - 1/N - N r_{bs} \sqrt{Var(s) Var(b)} \qquad (2)$$

Variance $Var(s)$, $Var(b)$ measures size dis-similarity while the correlation of s and b, r_{bs}, is an inverse measure of specialization. Implication 3 follows from equation (2).

More on Size Effects

More novel *implication 4, world openness is ordinarily increasing in the number of countries*: smaller countries are more open; division makes for more and smaller countries.

Differentiate $\sum_j \sum_{i \neq j} X_{ij}/Y = 1 - \sum_j b_j s_j$, yielding

$$-\sum_j (b_j ds_j + s_j db_j).$$

The differential expression above should ordinarily be positive.

More on Size Effects

On aggregate trade data, gravity $\Rightarrow$ *implication 5, world openness rises with convergence* under balanced trade, $b_j = s_j, \forall j$. The right hand side of equation (2)

$$\to NVar(s) + 1/N$$

while per capita income convergence lowers $Var(s)$ toward $Var(population)$.

Baier and Bergstrand (2001) find relatively little action from convergence in postwar growth.

Recent rise of China and India might give more action.

But size interacts with frictions and their incidence ...

Basic Elements

Detail, high variation of bilateral shipments $\Rightarrow$ infer trade costs.

Economize on other details of economic interaction to preserve bilateral detail.

Solution: Modularity, consistent with many g, e. superstructures.

- Focus on distribution for given levels of supply at each origin and given demand at each destination.
- Requires separability restrictions on preferences and/or technology
- Requires separability of distribution costs. Iceberg melting $\iff$ distribution uses resources in same proportion as production.

Three Structural Frameworks of Distribution

The same structural gravity model $\Leftarrow$ 3 frameworks:

- Differentiated Products in Demand, given supply and expenditure
- Differentiated Productivity in Supply, homogeneous demand
- Discrete Choice Aggregation (3rd lecture)

General equilibrium of distribution (conditional on upper level supply and demand variables).

Differentiated Demand

CES demand structure (final or intermediate)

$$X_{ij}^k = \left(\frac{\beta_i^k p_i^k t_{ij}^k}{P_j^k} \right)^{1-\sigma_k} E_j^k \tag{3}$$

where

$$P_j^k = \left(\sum_i (\beta_i^k p_i^k t_{ij}^k)^{1-\sigma_k} \right)^{1/(1-\sigma_k)} \tag{4}$$

Market clearance at end user prices: $Y_i^k = \sum_j X_{ij}^k$.

Use (3) in market clearance eqn., then factor out $(\beta_i^k p_i^k)^{1-\sigma_k}$.

Distribution Equilibrium

Substitute for $(\beta_i^k p_i^k)^{1-\sigma_k}$:

$$X_{ij}^k = \frac{E_j^k Y_i^k}{Y^k} \left(\frac{t_{ij}^k}{P_j^k \Pi_i^k} \right)^{1-\sigma_k} \tag{5}$$

$$(\Pi_i^k)^{1-\sigma_k} = \sum_j \left(\frac{t_{ij}^k}{P_j^k} \right)^{1-\sigma_k} \frac{E_j^k}{Y^k} \tag{6}$$

$$(P_j^k)^{1-\sigma_k} = \sum_i \left(\frac{t_{ij}^k}{\Pi_i^k} \right)^{1-\sigma_k} \frac{Y_i^k}{Y^k}. \tag{7}$$

(5)-(7) is the structural gravity model. (5) is the trade flow equation. Π_i^k denotes outward multilateral resistance (OMR), while P_j^k denotes inward multilateral resistance (IMR).

Empirical Orientation

YX_{ij}/Y_iE_j is ratio of actual trade to frictionless trade; related to proxies for unobservable trade frictions:

$$t_{ij} = \exp(\gamma_0 + \gamma_1 \ln d_{ij} + \gamma_z z + \phi_i + \mu_j + \ln \epsilon_{ij}) \tag{8}$$

where d_{ij} is bilateral distance, z is other controls such as common language, ϕ_i is an exporter fixed effect and μ_j is an importer fixed effect.

Could also use full information methods to replace ϕs and μs. More risky estimation (but Anderson-Yotov 2010b suggested structural gravity comes very close).

My take: gravity is about estimating (8). Agnosticism about upper level linkage. Size effects controlled for by origin and destination fixed effects.

Disaggregation Again

Appropriate disaggregation should be used to estimate (8).

Trade costs probably differ by firm. If data permits ...

Distinction between arms-length and various degrees of affiliation probably matters.

Appropriate proxies, direct data where possible (but endogeneity).

Normalization

(8) can only identify *relative* trade costs from perturbations of relative trade flows from frictionless trade flows. Levels of t_{ij} implied by regression are due to implicit normalization.

In practice, normalize $\{t_{ij}\}$ by $\min_i t_{ii} = 1$.

Given ts, (6)-(7) can be solved for Πs and Ps up to a normalization. Full g.e. consistency restricts the normalization in each sector.

Incidence

Πs and *P*s are sellers' and buyers' incidence respectively.

- Replace actual t_{ij}^ks by $\Pi_i^k P_j^k$ in equations (6)-(7): they continue to hold given *E*s and *Y*s.
- Structural gravity also implies

$$(\beta_i^k p_i^{*k} \Pi_i^k)^{1-\sigma_k} = Y_i^k / Y^k, \quad \forall i, k \tag{9}$$

where βs are CES parameters and p^*s are 'factory gate' prices.

- Because $\sum_i (\beta_i^k p_i^{*k} \Pi_i^k)^{1-\sigma_k} = 1$, the left hand side of (9) is a CES expenditure share on 'world' market. Effectively good *k* from *i* is sold at uniform sellers' incidence Π_i^k to a 'world' market.

Incidence and TFP

Multilateral resistance is interpreted as the incidence of TFP frictions in distribution.

Sectoral TFP friction in distribution from i in good k:
$\bar{t}_i^k = \sum_j t_{ij}^k y_{ij}^k / \sum_j y_{ij}^k$. Laspeyres index of trade frictions.
Compare to Π_i^k.

$\bar{t}_i^k$ gives the sellers' incidence Π_i^k only under the p.e. and inconsistent assumption that all incidence falls on the seller i.

Laspeyres TFP and incidence of TFP differ in magnitude and for IMR the correlation between them is low (Anderson and Yotov, 2010a,b).

Useful Structural Gravity Indexes

$$CHB_i \equiv \left(\frac{t_{ii}}{\Pi_i P_i} \right)^{1-\sigma}.$$

CHB varies substantially by country, product and time despite constant gravity coefficients (Anderson and Yotov; 2010a,b).

Border Effects: inter-regional/international trade costs. For British Columbia's exports to adjacent Alberta and across the US border to adjacent Washington

$$\frac{X_{BC,AB}}{X_{BC,WA}} = \left(\frac{t_{BC,WA}}{t_{BC,AB}} \frac{P_{AB}}{P_{WA}} \right)^{\sigma-1}.$$

Anderson and Yotov (2010a) decompose sellers' incidence into domestic and international components for Canada's provinces.

Comparative Statics

Structural gravity $\Rightarrow$ comparative statics of incidence.

- For given E and Y shares, calculate changes in incidence Π's and P's due to changes in t's.

  Trade flows in (5) are invariant to a uniform rise in trade costs (including costs of internal shipment). Π's and P's are HD(1/2) in t's.
- For given t's calculate changes in incidence due to change E, Y's.
  The empirical literature tends to indicate little change in gravity coefficients, but big changes in CHB (Anderson-Yotov 2010a,b).

Full general equilibrium structure $\Rightarrow$ changes in Y, E's induced by change in ts or other exogenous variables.

Differential Productivity

(5)-(7) also $\Leftarrow$ by Ricardian technology drawn from a Frechet probability distribution (Eaton and Kortum, 2002). CES demand for intermediates homogeneous across origins; substitution on intensive margin disappears in equilibrium.

Productivity draws plus trade frictions $\Rightarrow$ equilibrium wages that assign to countries effectively CES proportions of the continuum of goods.

Frechet dispersion parameter $\rightarrow$ comparative advantage, substitutability on extensive margin like σ on intensive margin. Location parameter differs nationally, represents absolute advantage, acts like β_i^k's taste advantage on intensive margin.

Absolute advantage and country size explain the Ys and Es

Discrete Choice Gravity

Three aspects of discrete choice over origin-destination flows

1. zeroes in trade flows
   - some potential flows are equal to zero.
   - may imply non-CES (with $\sigma > 1$), hence choke price. Translog promise here.
   - but may imply fixed export cost, discrete choice by firms to enter.
   - fixed export cost suggests a volume effect, selection of number of active firms.
2. migrant choice of destination
3. firm's choice over destinations

Zeroes

HMR: CES/Armington preferences; zeroes $\Leftarrow$ fixed costs of export facing firms. High productivity firms choose to pay the fixed cost of exporting; if none do then zero trade from i.
The gravity model becomes

$$
\begin{aligned}
X_{ij}^k &= \frac{E_j^k Y_i^k}{Y^k} V_{ij}^k \left(\frac{t_{ij}^k}{P_j^k \Pi_i^k} \right)^{1-\sigma_k} \\
(\Pi_i^k)^{1-\sigma_k} &= \sum_j \left(\frac{t_{ij}^k}{P_j^k} \right)^{1-\sigma_k} \frac{V_{ij}^k E_j^k}{Y^k} \\
(P_j^k)^{1-\sigma_k} &= \sum_i \left(\frac{t_{ij}^k}{\Pi_i^k}. \right)^{1-\sigma_k} \frac{V_{ij}^k Y_i^k}{Y^k}.
\end{aligned}
$$

V_{ij}^k generated by selection equation separately estimated.

Econometrics of Zeroes

HMR results suggest selection is potent: reduces variable trade cost coefficients.

OLS alternative: drop the zeroes; disallow selection. In principle this is OK if no selection. Tobit is not OK.

Santos-Silva and Tenreyro (2006) suggest that estimation the log of (8) is biased due to non-normal heteroskedastic ϵ_{ij}. Alternative estimation is Poisson-Pseudo-Maximum-Likelihood (PPML). No treatment of selection.

Anderson-Yotov (2010b) find that OLS, PPML and HMR give equivalent results on normalized *ts*, due to near perfect collinearity of gravity coefficients in the three methods.

Selection in HMR

Selection mechanism: non-negative profits requirement selects upper tail of Pareto productivity distribution of potential trading firms. Pareto is consistent with observation that the largest and most productive firms export the most. Pareto allows estimation with industry trade data.

Notice that zeroes do not arise if the support of the productivity distribution is not bounded above. Artificial restriction?

Besedes and Prusa (2006 JIE, CJE): (US 10 digit HS bilateral) trade flickers on and off, tending to contradict fixed costs.

Translog $\Rightarrow$ choke prices, combine with fixed costs: possibly discriminate between two explanations of zeroes.

Migration and Multinomial Logit

Let w^i denote the wage at location $i, \forall i$. Migrant from j to i with iceberg cost factor $\delta^{ji} > 1$ receives net wage (w^i/δ^{ji}).

Assume logarithmic utility. Migrant's utility of migration is $u^{ji} = \ln w^i - \ln \delta^{ji} - \ln w^j + \ln \epsilon^{jih}$, where the idiosyncratic utility of migration $\ln \epsilon^{jih}$ is not observable by the econometrician. Worker chooses the destination with the largest surplus.

McFadden (1973) showed that if $\ln \epsilon$ had the type-1 extreme value distribution, the probability that a randomly drawn individual would pick any particular migration destination has the multinomial logit form.

Migrant proportions thus have the multinomial logit form, w^i and w^j are fixed effects and δ^{ji} is proxied by gravity variables.

Structural Gravity Setup

Make use of market clearance in the same way as for goods (thus model distribution of stocks of labor).

With logarithmic utility, the migration equation is analogous to the CES demand

$$M^{ji} = \frac{w^i/\delta^{ji}}{\sum_k w^k/\delta^{jk}} N^j. \tag{10}$$

Solve the market clearance equations for each destination of labor and substitute out in the bilateral flow equations.

Structural Gravity Model of Migration

$$M^{ji} = \frac{L^i N^j}{N} \frac{1/\delta^{ji}}{\Omega^i W^j}. \tag{11}$$

$$\Omega^i = \sum_j \frac{1/\delta^{ji}}{W^j} \frac{N^j}{N}. \tag{12}$$

$$W^j = \sum_k \frac{1/\delta^{jk}}{\Omega^k} \frac{L^k}{N}. \tag{13}$$

Can get the CES form if utility is the log of a CRRA function.

Ω and W have buyers and sellers incidence interpretation.

Other Potential Applications

Services trade (Head and Mayer)

FDI

Portfolio investment not very amenable to finding structural gravity so far.

Multinomial Logit and Trade

One approach to trading firms' choice of destinations follows the migration model and applies multinomial logit.

In principle this could look exactly like (8), but attention to the detailed implications might suggest useful alternative trade proxies or direct evidence.

An important objection to this simple procedure is that firms do not choose *one* destination out of many independent choices. Instead, the firm may realize that each choice affects the cost of reaching other destinations. Multinomial logit is no longer appropriate.

Sequencing Export Entry

Morales (Harvard, 2010) has applied new techniques from applied IO to estimate how firms choose multiple markets over time.

Cannot infer exact size of sunk costs from choices of firms, but the observed choices (coupled with weak assumptions on firm behavior) can be used to rule out certain regions of the parameter space. Averaging over different firms' decisions in different markets yields a set (or interval) of parameter values consistent with firm behavior.

Moment Inequality Estimation, first application in trade. No need to impose that firms have perfect foresight nor that they have rational expectations.