Gravity Models and the Armington Assumption

Background

Economists love the elegance and completeness of physics, and what could be more elegant than Newton's Law of Universal Gravity? To recap: The gravitational force between two objects is proportional to the product of their masses and inversely proportional to the square of the distance that separates them,

$$F_g = G \frac{m_1 m_2}{r^2}. \tag{1}$$

The gravity "theory" of trade replaces objects with countries, masses with GDP, and distance with, well, distance. Note that Newton's theory is linear in logs, a property conducive to OLS.

Border effects

One of the most famous applications of gravity is McCallum (1995), which uses an atheoretic gravity equation to compare trade between Canadian provinces with trade between Canadian provinces and U.S. states,

$$x_{ij} = e^{\alpha_1} \frac{y_i^{\alpha_2} y_j^{\alpha_3}}{d_{ij}^{\alpha_4} e^{\alpha_5 \delta_{ij}}}, \tag{2}$$

where x_{ij} are exports (by value) from region i to j, y_i is GDP in region i, d_{ij} is the distance between regions i and j, and δ_{ij} is a dummy variable equal to 1 if i and j are both provinces. This last variable captures the *border effect*. In logs, and with an ad-hoc error term, we can estimate

$$\log x_{ij} = \alpha_1 + \alpha_2 \log y_i + \alpha_3 \log y_j + \alpha_4 \log d_{ij} + \alpha_5 \delta_{ij} + \epsilon_{ij}. \tag{3}$$

Estimating this equation typically finds income elasticities (α_2 and α_3) near 1, a negative coefficient on distance, and a coefficient on the border dummy near 3.0. The border effect is $\exp(\alpha_5) = 20.1$, which implies that, after controlling for size and distance, Canadian provinces trade about 20 times more with each other than they do with the United States. This statistic is largely considered puzzling given the relatively low level of formal barriers to trade between the U.S. and Canada.

McCallum's gravity regressions also highlight the strength of the gravity model: The adjusted R^2 of the regressions are in the range of 0.80-0.90! Gravity in trade may not be a law, but it is pretty close. That such a simple empirical model fits the data this well explains the persistence of this idea in the profession.

In the next few classes we will study formal models that emit a gravity equation in equilibrium. The ability of a model to account for gravity is a good first check: If a model can generate gravity, we know it will be able to account for observed aggregate trade flows.

Gravity with Gravitas

Anderson and van Wincoop (2003) show how a gravity-like equation can be derived from a model with constant elasticity of substitution preferences and the *Armington* assumption. The Armington assumption is that goods are nationally differentiated: The good produced in Canada is different from the good produced in Mexico by the very fact that it is made in Canada. Molson is different from Corona not because it is made with a different recipe, but because it is brewed in Canada.

One advantage of this assumption (this is pre-differentiated good, monopolistic competition models) is that it generates intraindustry trade.

The model

There are J regions indexed by j, and preferences are

$$U_j = \left(\sum_{i=1}^{J} \beta_i^{(1-\sigma)/\sigma} c_{ij}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)} \tag{4}$$

and the budget constraint is $\sum_i p_{ij} c_{ij} = y_j$. Notice the preference parameters β_i are country-specific (in this case that implies: good-specific). In computable general equilibrium models, these parameters would be set to pin down base year trade flows. These parameters are sometimes called *home bias* parameters because countries' consumption baskets are skewed towards domestic goods. The elasticity of substitution between goods is σ.

Trade costs are denoted t_{ij} and the exporter price is denoted p_i. The price paid by the agent in j is $p_{ij} = t_{ij} p_i$. Trade costs are not observed by the econometrician.

The CES preferences yields the usual demand functions (NB: in expenditure form! $x = p \cdot q$),

$$x_{ij} = \left(\frac{\beta_i p_i t_{ij}}{P_j} \right)^{1-\sigma} y_j, \tag{5}$$

with the usual aggregate price index,

$$P_j = \left(\sum_{i=1}^{J} (\beta_i p_i t_{ij})^{(1-\sigma)} \right)^{1/(1-\sigma)}. \tag{6}$$

The aggregate value of production is $y_i = \sum_{j=1}^{J} x_{ij}$, which, after substituting (5) implies

$$y_i = (\beta_i p_i)^{1-\sigma} \sum_{j=1}^{J} (t_{ij}/P_j)^{1-\sigma} y_j. \tag{7}$$

Solving (7) for $\beta_i p_i$ (what the authors call *scaled prices*) and substituting into (5) gets us

$$x_{ij} = \frac{y_i y_j}{y_w} \left(\frac{t_{ij}}{\Pi_i P_j} \right)^{1-\sigma} \tag{8}$$

$$\Pi_i \equiv \left(\sum_{j=1}^{J} (t_{ij}/P_j)^{1-\sigma} \theta_j \right)^{1/(1-\sigma)} \tag{9}$$

with $\theta_j = y_j/y_w$. Using (7) in (6) — substituting scaled prices as before — gets us

$$P_j = \left(\sum_{i=1}^{J} (t_{ij}/\Pi_i)^{(1-\sigma)} \theta_i \right)^{1/(1-\sigma)}. \tag{10}$$

That's it: (9) and (10) form a system of $J \times 2$ equations in $J \times 2$ unknowns, $\{\Pi_i, P_i\}$. If we make the assumption $t_{ij} = t_{ji}$ (a reasonable assumption if these are transportation costs, probably not if these are policy parameters) then the solution to the system is trivial: $\Pi_i = P_i$, and gravity follows

$$x_{ij} = \frac{y_i y_j}{y_w} \left(\frac{t_{ij}}{P_i P_j} \right)^{1-\sigma} \tag{11}$$

$$P_j = \left(\sum_{i=1}^{J} (t_{ij}/P_i)^{(1-\sigma)} \theta_i \right)^{1/(1-\sigma)} \tag{12}$$

Anderson and Van Wincoop refer to the P_j as *multilateral resistance* terms, since they contain the costs (and sizes) of all the regions. P_j is a size-weighted index of the costs of region j trading with all other regions, adjusting both for trade costs, and the average price of goods being consumed in the other regions. How would you expect P_j to differ between New Zealand and Portugal, two countries that are of almost the same size?

Notice that the multilateral resistance terms enter the gravity equation (11) with positive coefficients $(1 - \sigma > 0)$. What matters for bilateral trade is the bilateral trade cost t_{ij} **relative** to the multilateral resistance of the pair. If i and j are relatively "remote" so that P_i and P_j are large, than the bilateral trade cost matters less for bilateral trade. Bilateral trade costs alone are not enough to understand bilateral trade patterns.

These terms show up in most models with CES preferences, and these terms feature prominently in Eaton-Kortum type Ricardian models. (more on that in future classes). Our structural gravity equation has a lot of intuitive features

- Bilateral trade flows decrease with trade costs between i and j, ceteris paribus.

- Bilateral trade flows increase with the size of i or j, ceteris paribus.

- The effects of trade costs and multilateral resistance are amplified by $\sigma > 1$.

- Bilateral trade flows are homogenous of degree 0 in t, where t is the entire vector of trade costs. This follows from P_j being homogeneous of degree 1/2.

Verify: $P_j(t\lambda) = P_j(t)\lambda^{1/2}$

$$P_j(\lambda t) = \left(\sum_{i=1}^{J} P_i(\lambda t)^{\sigma-1} \theta_i (\lambda t_{ij})^{1-\sigma} \right)^{1/(1-\sigma)} \tag{13}$$

$$P_j(\lambda t) = \left(\sum_{i=1}^{J} P_i(t)^{\sigma-1} \lambda^{-(1-\sigma)/2} \theta_i t_{ij}^{1-\sigma} \lambda^{1-\sigma} \right)^{1/(1-\sigma)} \tag{14}$$

$$P_j(\lambda t) = \lambda^{1/2} \left(\sum_{i=1}^{J} P_i(t)^{\sigma-1} \theta_i t_{ij}^{1-\sigma} \right)^{1/(1-\sigma)} \tag{15}$$

Country size and trade barriers

Completely differentiating (12) and doing a bit of algebra yields

$$dP_j = \left(0.5 - \theta_j + 0.5 \sum_{k=1}^{J} \theta_k^2 \right) dt \tag{16}$$

Trade barriers matter more for smaller countries. The intuition is that smaller countries rely more on trade with the rest of the world: New Yorkers can trade with Canadians, but they can also trade with Floridians, Californians, and Texans. None of the latter 3 count as import and export partners although they are very large regions. Ontarians can trade with the United States, and also with Saskatchewan or Alberta, much smaller provinces.

This leads us to a big-picture question: Why do we care about trade between countries, per se? One answer is that countries are the units in which most data are collected. Another reason might be that we expect policy to apply homogenously within a country and heterogeneously across countries, so it provides us with some (non-exogenous) variation. Whatever your reason may be, it is worth keeping in mind these existential questions when thinking about your own research.

Back to (16): A two country world consisting of the US and Canada has roughly $\theta_{can} = 0.1$ and $\theta_{us} = 0.9$. This implies that an increase in trade barriers should affect Canadian multilateral resistance about 80 times more than that of the US ($dP_{CAN} = 0.81dt$, $dP_{USA} = 0.01dt$). This feature of the model is consistent with the data: From table 1 in Anderson and van Wincoop (2003), we can see that the border effect is larger for Canada than the United States (16.4 for Canada, 0.4 for the United States).

Empirics

Trade costs are unobservable, so make them a function of observable variables. To stay close to McCallum,

$$t_{ij} = b^{(1-\delta_{ij})} d_{ij}^{\rho} \tag{17}$$

where d_{ij} is distance, b is equal to 1 if the two regions are in the same country and equal to $1 + \tau$, the tariff equivalent of the border, if in different countries, and δ_{ij} is the indicator variable equal to 1 if the two regions are in the same country. Given this specification the empirical gravity equation is

$$\log \frac{x_{ij}}{y_i y_j} = a_0 + a_1 \log d_{ij} + a_2 (1 - \delta_{ij}) - \log P_i^{1-\sigma} - \log P_j^{1-\sigma} + \epsilon_{ij} \tag{18}$$

along with

$$P_j^{1-\sigma} = \sum_{i=1}^{J} P_i^{(\sigma-1)} \theta_i \exp \left(a_1 \log d_{ij} + a_2 \left(1 - \delta_{ij} \right) \right) \quad j = 1 \ldots J. \tag{19}$$

The parameter vector is $\gamma = [a_0, a_1, a_2]'$ and the data are vectors of trade flows, GDP, distances, and indicator variables, δ.

Compare (18) to McCallum's version in (3): there are 2 differences. The first is that (18) restricts the income elasticities to be 1, as in the theory. In (3) the income elasticities are estimated, and are usually found to be close to 1.

The second, and more substantive difference, is the inclusion of the two P terms. The P terms

include **all of the bilateral trade barriers** as defined in (19). This implies that McCallum's gravity model suffers from omitted variable bias.

For a given parameter vector, the system defined in (19) can be solved for $\{P_i^{\sigma-1}\}$ and then the sum of squared errors of (18) can be computed. Adjust the parameter vector to minimize this sum of squared errors. An alternative method, which is commonly used, is to not solve for the multilateral resistance terms separately, but to use country-level fixed effects — but then we can't do counterfactuals.

In the two country model $a_1 = (1 - \sigma)\rho \approx -0.80$ and $a_2 = (1 - \sigma)\log b \approx -1.60$. σ cannot be identified here, but a common value might be 5. In that case, b would be approximately 1.5, implying provinces trade 1.5 times more with each other. Remember McCallum's number was 20!

In table 3 we can see the estimates of $P^{1-\sigma}$, reported as the average over all the regions in each country. The average value for the United States is 0.77 while the average value for Canada is 2.45: The average Canadian province is much more "remote" than an average U.S. state.

Understanding McCallum's border effect

The idea: With estimated parameters in hand, we solve the model with border effects and without, which will tell us how important borders are in restricting trade. Counterfactuals like this are why we have models!

Keeping σ and ρ unchanged, solve the model once with borders, $b = 1.5$, and once without, $b = 1$. To solve the model without the border effect, we need to choose a σ because the prices are no longer normalized to 1; we'll stick with $\sigma = 5$. Table 3 in the paper shows the (geometric) average $P_j^{1-\sigma}$. The ratio of $P_j^{1-\sigma}$ with borders effects to $\tilde{P}_j^{1-\sigma}$ without border effects is 1.02 for the U.S. and 2.08 for Canada. Borders matter a lot more for small countries, as implied by theory.

Define the size-adjusted trade flow as

$$z_{ij} = x_{ij}\frac{y_w}{y_i y_j} = b^{(1-\sigma)(1-\delta_{ij})}d_{ij}^{\rho(1-\sigma)} P_u^{\sigma-1}P_c^{\sigma-1}. \tag{20}$$

The ratio of average (w.r.t. to states or provinces) cross-country trade in a world with borders to trade in a world without borders (the tilde variables):

$$\frac{z_{uc}}{\tilde{z}_{uc}} = b^{1-\sigma}\left(\frac{P_u^{\sigma-1}}{\tilde{P}_u^{\sigma-1}}\right)\left(\frac{P_c^{\sigma-1}}{\tilde{P}_c^{\sigma-1}}\right) \tag{21}$$

$$\frac{z_{uc}}{\tilde{z}_{uc}} = 1.5^{-4} \times 1.02 \times 2.08 = 0.41 \tag{22}$$

Trade with borders is about 41 percent of trade without borders, and most of it is driven by multilateral effects, not bilateral. That is, the trade dampening effect of $b^{1-\sigma}$ is partially offset by the increase in multilateral resistance in Canada, $P_c^{\sigma-1}$, that arises from the border. (Note: trade is balanced in the model, so $z_{uc} = z_{cu}$.)

Within country trade ($b = 1$):

$$\frac{z_{cc}}{\tilde{z}_{cc}} = 1.0^{-4} \times 2.08 \times 2.08 = 4.32 \tag{23}$$

$$\frac{z_{uu}}{\tilde{z}_{uu}} = 1.0^{-4} \times 1.02 \times 1.02 = 1.04 \tag{24}$$

We can see directly from these tables the McCallum type calculations: Borders increase intraprovincial trade by a factor of 4.32/0.41=10.5 and intrastate trade by a factor of 1.04/0.41=2.6.

More on Armington's Assumption

Armington-style national product differentiation is commonly used in the computable, or applied general equilibrium (AGE or CGE) models that were the workhorse models among academics for policy analysis in the 1980s and 1990s, and are still very popular in policy circles. (e.g., the GTAP project at Purdue University)

The basic structure is similar to that we have laid out above, but with much more detail. A typical model structure:

- J countries

- I industries

- N factors of production, although often just K and L

Production of good i in country j uses capital, labor, and intermediate inputs of each good $k = 1, ..., I$

$$y_i^j = \min \left\{ A_j (K_i^j)^{\alpha_i^j} (L_i^j)^{1-\alpha_i^j}, a_{i1}^j x_{i1}^j, ..., a_{iI}^j x_{iI}^j \right\} \qquad i = 1, ..., I \ \ j = 1, ..., J \tag{25}$$

where intermediate good x_k^j is made up of good i from all countries: an Armington Aggregate. Feasibility has consumption and intermediate good use of good k equal to the amount available

$$c_k^j + \sum_i x_{ik}^j = \left(\sum_n b_k^{nj} (m_k^{nj})^\rho \right)^{1/\rho}, \tag{26}$$

where m_k^{nj} are imports of good k from country n to j. Notice the b parameters play a similar role as in our model above. The consumers problem is

$$U_j = \max \left(\sum_k \beta_k^j (c_k^j)^\theta \right)^{1/\theta} \tag{27}$$

$$\text{s.t.} \sum_k p_k^j c_k^j = wL + rK + T. \tag{28}$$

Notice that there is no law of motion for the capital stock: the older versions of these models usually did not include dynamics, and capital was thought of as an endowment.

There are a lot of parameters floating around this model, but almost all of them can be calibrated from an input-output matrix. In this sense, AGE style models take the industrial structure of

a country very seriously, which allows the researcher to make predictions about how particular industries will change as policy parameters are changed. Even if you do not plan to use these kinds of models, understanding an input-output table is still useful: it forces you to confront the national income and product accounts in a consistent way.

A simple introduction to AGE models can be found in Kehoe and Kehoe (1994).

——————

Anderson, James E. and Eric van Wincoop (2003). "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93 (1), pp. 170–192.

Kehoe, Patrick J. and Timothy J. Kehoe (1994). "A Primer on Static Applied General Equilibrium Models." *Quarterly Review* 18 (2), pp. 2–16.

McCallum, John (1995). "National Borders Matter: Canada-U.S. Regional Trade Patterns." *American Economic Review* 85 (3), pp. 615–23.