Empirical trade analysis

Gravity theory and estimation

Cosimo Beverelli

World Trade Organization

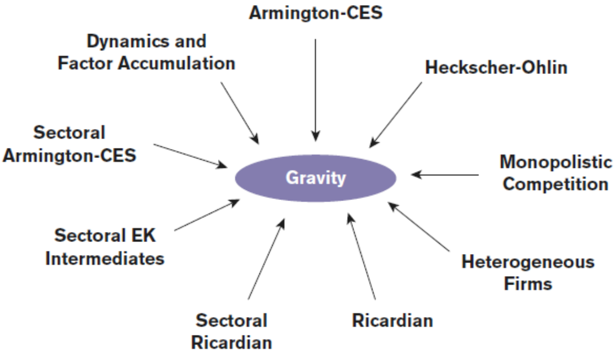
Outline

Introduction

- According to Newton's Law of Universal Gravitation, any particle in the universe attracts any other particle thanks to a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them
- Applied to international trade, Newton's Law of Gravity implies that, just as particles are mutually attracted in proportion to their sizes and proximity, countries trade in proportion to their respective market size (e.g. gross domestic products) and proximity
- This is the *a-theoretical* gravity model of international trade
- Today, we will derive a *theoretical* gravity model of international trade, based on the Armington-CES model of Anderson and van Wincoop (2003)
- We follow Chapter 1 of the Advanced Guide – where you will also find all references and extensions (e.g., derivations of Eaton-Kortum gravity and of gravity with tariffs in Appendices A and B to Chapter 1)
- Equation numbering is the same as in the Advanced Guide

All roads lead to gravity

Figure 1 Gravity model's strong theoretical foundations



Assumptions

- Consider a world that consists of N countries, where each economy produces a variety of goods (i.e. goods are differentiated by place of origin as in Armington, 1969) that is traded with the rest of the world
- The supply of each good is fixed to Q_i, and the factory-gate price for each variety is p_i
- Thus, the value of domestic production in a representative economy is defined as $Y_i = p_i Q_i$, where Y_i is also the nominal income in country i
- Country i's aggregate expenditure is denoted by E_i. Aggregate expenditure can also be expressed in terms of nominal income by $E_i = \phi_i Y_i$, where $\phi_i > 1$ shows that country i runs a trade deficit, while $0 < \phi_i < 1$ reflects a trade surplus (treated as exogenous)
- For brevity's sake, the time dimension t is omitted in the derivations

Consumer preferences

- On the demand side, consumer preferences are assumed to be homothetic, identical across countries, and given by a CES-utility function for country j:

$$\left\{ \sum_i \alpha_i^{\frac{1-\sigma}{\sigma}} c_{ij}^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{1-\sigma}{\sigma}} \tag{1}$$

where $\sigma > 1$ is the elasticity of substitution among different varieties, i.e. goods from different countries, $\alpha_i > 0$ is an exogenous CES preference parameter, and c_{ij} denotes consumption of varieties from country i in country j

Budget constraint

- Consumers maximize equation (1) subject to the following standard budget constraint:

$$\sum_i p_{ij} c_{ij} = E_j \tag{2}$$

- Equation (2) ensures that the total expenditure in country j, E_j, is equal to the total spending on varieties from all countries, including j, at delivered prices $p_{ij} = p_i t_{ij}$
- Prices are a function of factory-gate prices in the country of origin, p_i, marked up by bilateral trade costs, $t_{ij} \geq 1$, between trading partners i and j
- These bilateral trade costs are defined as iceberg costs, as is standard in the trade literature (Samuelson, 1952)
- In order to deliver one unit of its variety to country j, country i must ship $t_{ij} \geq 1$ units, since $1/t_{ij}$ of the initial shipment melts "en route"

Solution to consumer's optimization problem

- Solving the consumer's optimization problem (see Appendix 2.A in Baldwin et al., 2011) yields the expenditures on goods shipped from origin i to destination j as:

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

where X_{ij} denotes trade flows from exporter i to destination j and P_j can be interpreted as a CES consumer price index:

$$P_j = \left[\sum_i (\alpha_i p_i t_{ij}) \right]^{\frac{1}{1-\sigma}} \tag{4}$$

Observations

- From equation (4), expenditure in country j on goods from source i, X_{ij} , is:

1. Proportional to total expenditure, E_j, in destination j. The simple intuition is that, all else equal, larger/richer markets consume more of all varieties, including goods from i
2. Inversely related to the (delivered) prices of varieties from origin i to destination j, $p_{ij} = p_i t_{ij}$. This is a direct reflection of the law of demand, which depends not only on factory-gate price p_i but also on bilateral trade cost t_{ij} between partners i and j. The ideal combination that favours bilateral trade is an efficient producer, characterized by low factory-gate price, and low bilateral trade cost between countries i and j

Observations (ct'd)

- From equation (4), expenditure in country j on goods from source i, X_{ij}, is:

3. Directly related to the CES price aggregator P_j. This relationship reflects the substitution effects across varieties from different countries. All else equal, the relatively more expensive the rest of the varieties in the world are, the more consumers in country j will substitute away from them and toward the goods from country i
4. Contingent on the elasticity of substitution σ when factory-gate prices or the aggregate CES prices (or in the combination of those as a relative price) change. All else equal, a higher elasticity of substitution will magnify the trade diversion effects from the more expensive commodities to the cheaper ones

Market clearance

- Impose market clearance for goods from each origin, $Yi \equiv \sum_j X_{ij} \; \forall \; i$. Using equation (3), this yields:

$$Y_i = \sum_j \left(\frac{\alpha_i p_i t_{ij}}{P_j} \right)^{(1-\sigma)} E_j \tag{5}$$

Further manipulations

- Defining $Y \equiv \sum_i Yi$ and dividing equation (5) by Y, the terms can be rearranged to obtain:

$$(\alpha_i p_i)^{(1-\sigma)} = \frac{\frac{Y_i}{Y}}{\sum_j \left(\frac{t_{ij}}{P_j}\right)^{(1-\sigma)} \frac{E_j}{Y}} \tag{6}$$

- Following Anderson and van Wincoop (2003), the term in the denominator of equation (6) can be defined as $\Pi_i^{(1-\sigma)}$, allowing to rewrite equation (6) as:

$$(\alpha_i p_i)^{(1-\sigma)} = \frac{Y_i/Y}{\Pi_i^{(1-\sigma)}} \tag{7}$$

The structural gravity system

- Combining equation (7) with equations (3) and (4) yields the structural gravity system:

$$X_{ij} = \frac{Y_i E_j}{Y} \left(\frac{t_{ij}}{\Pi_i P_j} \right)^{(1-\sigma)} \tag{8}$$

$$\Pi_i^{(1-\sigma)} = \sum_j \left(\frac{t_{ij}}{P_j} \right)^{(1-\sigma)} \frac{E_j}{Y} \tag{9}$$

$$P_j^{(1-\sigma)} = \sum_i \left(\frac{t_{ij}}{\Pi_i} \right)^{(1-\sigma)} \frac{Y_i}{Y} \tag{10}$$

- Equation (10) (1-8) is the theoretical gravity equation that governs bilateral trade flows. It can be conveniently decomposed into two terms: (i) a size term, $Y_i E_j / Y$; and (ii) a trade cost term, $[t_{ij} / (\Pi_i P_j)]^{(1-\sigma)}$

Size vs. trade cost decomposition: size term

- The intuitive interpretation of the size term, $Y_i E_j / Y$, is as the hypothetical level of frictionless trade between partners i and j if there were no trade costs
- Mechanically, this can be shown by eliminating bilateral trade frictions (i.e. setting $t_{ij} = 1$), and re-deriving the gravity system
- Intuitively, a frictionless world implies that consumers will face the same price for a given variety regardless of their physical location and that their expenditure share on goods from a particular country will be equal to the share of production in the source country in the global economy (i.e. $X_{ij}/E_j = Y_i/Y$)
- Overall, the size term already carries some very useful information regarding the relationship between country size and bilateral trade flows: namely, large producers will export more to all destinations; big/rich markets will import more from all sources; and trade flows between countries i and j will be larger the more similar in size the trading partners are

Size vs. trade cost decomposition: trade cost term

- The natural interpretation of the trade cost term, $[t_{ij}/(\Pi_i P_j)]^{(1-\sigma)}$, is that it captures the total effects of trade costs that drive a wedge between realized and frictionless trade. The trade cost term consists of three components:

1. 1. Bilateral trade cost between partners i and j, t_{ij} , is typically approximated in the literature by various geographic and trade policy variables, such as bilateral distance, tariffs and the presence of regional trade agreements (RTAs) between partners i and j
2. The structural term P_j , coined by Anderson and van Wincoop (2003) as inward multilateral resistance represents importer j's ease of market access
3. The structural term Π_i , defined as outward multilateral resistances by Anderson and van Wincoop (2003), measures exporter i's ease of market access

- The multilateral resistances are general equilibrium trade costs that take into account the changes in prices, incomes and expenditures induced by bilateral trade cost changes

Analogy between the Newtonian theory of gravitation and the gravity trade model

- Newton's Law of Universal Gravitation:

$$F_{ij} = G\frac{M_i M_j}{D_{ij}^2}$$

  where F_{ij} is the gravitational force between objects i and j; G is a gravitational constant: M_i and M_j are object i and j's masses; and D_{ij} is distance between objects i and j

- Gravity Trade Model:

$$X_{ij} = \widetilde{G}\frac{Y_i E_j}{T_{ij}^{\theta}}$$

  where X_{ij} are exports from i to j; $\widetilde{G}$ is the inverse of world production, $\widetilde{G} \equiv 1/Y$; Y_i is country i's domestic production; E_j is country j's aggregate expenditure; and T_{ij} are total trade costs between countries i and j, $T_{ij}^{\theta} \equiv [t_{ij}/(\Pi_i P_j)]^{(\sigma-1)}$

The log-linear gravity equation

- Given the multiplicative nature of the structural gravity equation (8), and assuming that it holds in each period of time t, it is possible to log-linearize it and expand it with an additive error term, $\varepsilon_{ij,t}$:

$$\begin{aligned} \ln X_{ij,t} = & \ln E_{j,t} + \ln Y_{i,t} - \ln Y_t + \\ & + (1-\sigma) \ln t_{ij,t} - (1-\sigma) \ln P_{j,t} - (1-\sigma) \ln \Pi_{i,t} + \varepsilon_{ij,t} \end{aligned} \tag{11}$$

- Specification (11) is the most popular version of the empirical gravity equation, and it has been used routinely in the trade literature to study the effects of various determinants of bilateral trade:
- Geography, demographics, RTAs, tariffs, exports subsidies, embargoes, trade sanctions, the World Trade Organization membership, currency unions, foreign aid, immigration, foreign direct investment, cultural ties, trust, reputation, mega sporting events (Olympic Games and World Cup), melting ice caps, etc.

Eight challenges

1. Multilateral resistances
2. Zero trade flows
3. Heteroskedasticity of trade data
4. Bilateral trade costs
5. Endogeneity of trade policy
6. Non-discriminatory trade policy
7. Adjustment to trade policy changes
8. Gravity with disaggregated data

Eight challenges

1. Multilateral resistances

Challenge 1: Multilateral resistances

- The multilateral resistance terms $P_{j,t}$ and $\Pi_{i,t}$ are theoretical constructs and, as such, they are not directly observable
- If not appropriately controlled for, there will be an omitted variable bias (Baldwin and Taglioni's 2006 "Gold Medal Mistake")

Solutions to challenge 1

(i) Iterative NLLS procedure suggested by Anderson and van Wincoop (2003): first estimate the trade cost parameters without controlling for the multilateral resistances. Then, use the estimated trade costs to construct an initial set of multilateral resistances. re-estimate the gravity model obtaining new set of trade costs...until convergence

(ii) Approximation of the multilateral resistance terms by "remoteness" indexes constructed as functions of bilateral distance, and Gross Domestic Products (GDPs) (Wei, 1996; Baier and Bergstrand, 2009):

$$\ln REM\_EXP_{i,t} = \ln \left[\sum_j \frac{DIST_{ij}}{\left(\frac{E_{j,t}}{Y_t}\right)} \right]$$

$$\ln REM\_IMP_{j,t} = \ln \left[\sum_i \frac{DIST_{ij}}{\left(\frac{Y_{i,t}}{Y_t}\right)} \right]$$

Head and Mayer (2014) criticize such reduced-form approaches as they bear little resemblance to the theoretical counterpart of multilateral terms

Solutions to challenge 1 (ct'd)

(iii) Eliminate multilateral resistance terms by using appropriate ratios based on the structural gravity equation. Notable examples include Head and Ries (2001), Head et al. (2010), and Novy (2013)

(iv) Hummels (2001) and Feenstra (2016) suggest using directional (exporter and importer) fixed effects in cross-section estimations

- More recently, Olivero and Yotov (2012) demonstrate that the multilateral resistance terms should be accounted for by *exporter-time* and *importer-time* fixed effects in a dynamic gravity estimation framework with panel data
- In addition to accounting for the unobservable multilateral resistance terms, the exporter-time and importer-time fixed effects will also absorb the size variables ($E_{j,t}$ and $Y_{i,t}$) from the structural gravity model as well as all other observable and unobservable country-specific characteristics, which vary across these dimensions, including various national policies, institutions, and exchange rates

Eight challenges

2. Zero trade flows

Challenge 2: Zero trade flows

- The ordinary least-squares (OLS) estimator has been the most widely used technique to estimate various versions of the gravity equation (11)
- A clear drawback of the OLS approach, however, is that it cannot take into account the information contained in the zero trade flows, because these observations are simply dropped from the estimation sample when the value of trade is transformed into a logarithmic form
- The problem with the zeros becomes more pronounced the more disaggregated the trade data are
- It is especially severe for sectoral services trade due to the highly localized consumption and highly specialized production

Solutions to challenge 2

(i) One frequently applied and very convenient – but theoretically inconsistent – method is to just add a very small, and in fact completely arbitrary, value to replace the zero trade flows

- As noted in Head and Mayer (2014) this approach should be avoided because the results depend on the units of measurement and the interpretation of the gravity coefficients as elasticities is lost

(ii) Eaton and Tamura (1995) and Martin and Pham (2008) propose the use of the Tobit estimator as an econometric solution to the presence of zeroes

- However, gravity theory is silent about the determination of the Tobit thresholds, causing a disconnect between estimation and theory
- In practice, the Tobit model would apply to a situation where small values of trade are rounded to zero or actual zero trade might reflect desired negative trade

Solutions to challenge 2 (ct'd)

(iii) Helpman et al. (HMR, 2008) propose a theoretically-founded two-step selection process, where exporters must absorb some fixed costs to enter a market. Thus, fixed costs provide an intuitive economic explanation for the zero trade flows to bridge theory and empirics

- The HMR model is estimated in two stages: (i) a first-stage Probit estimation, which determines the probability to export, and (ii) a second-stage OLS estimation based on the positive sample of trade flows that also accounts for selection into exporting due to fixed costs of exporting
- Some challenges with the HMR estimation are that it is hard to find good exclusion restrictions for the first-stage Probit estimation and/or the need for custom programming when identification relies on functional form
- Additional difficulties with the HMR approach arise for panel data estimations and when dynamic considerations are taken into account

Solutions to challenge 2 (ct'd)

(iv) Egger et al. (2011) suggest a two-part gravity model that enables to decompose the effects of the explanatory variables on exports into an effect on the extensive country margin, i.e. the decision to export to a country at all, and on the intensive margin, i.e. the value of exports conditional on positive exports

- Additionally, and contrary to Helpman et al. (2008), their approach also takes care of potential endogenous regressors such as RTAs in the estimating equation for the extensive and intensive margin (see challenge 5)

(v) An easy and convenient solution to the presence of zero trade flows is to estimate the gravity model in multiplicative form instead of logarithmic form

- This approach, advocated by Santos Silva and Tenreyro (2006), consists in applying the Poisson Pseudo Maximum Likelihood (PPML) estimator to estimate the gravity model
- Monte Carlo simulations show that the PPML estimator performs very well even when the proportion of zeros is large

Eight challenges

3. Heteroskedasticity of trade data

Challenge 3: Heteroskedasticity of trade data

- Trade data are plagued by heteroskedasticity
- As pointed out by Santos Silva and Tenreyro (2006), in the presence of heteroskedasticity (and owing to Jensen's inequality), the estimates of the effects of trade costs and trade policy are not only biased but also inconsistent when the gravity model is estimated in log-linear form with the OLS estimator (or any other estimator that requires non-linear transformation)

Solutions to challenge 3

(i) Equation (11) can be estimated after transforming the dependent variable into size-adjusted trade, which is defined as the ratio between trade and the product of the sizes of the two markets, $X_{ij,t}/(E_{j,t}Y_{i,t})$ (Anderson and van Wincoop, 2003)

- The intuition behind this adjustment is that, arguably, the variance of the error term $\varepsilon_{ij,t}$ is proportional to the product of the sizes of the two markets
- A potential drawback of this approach is that it accounts for (the product of) country size as the only source of heteroskedasticity
- Furthermore, using the proposed size-adjusted trade as dependent variable would not eliminate the issue of zero trade flows challenge

(ii) A more comprehensive approach, proposed by Santos Silva and Tenreyro (2006), is to apply the PPML estimator

- In principle, the heteroskedasticity issue can be addressed with other estimators, such as the Gamma estimator
- As discussed above, the PPML estimator also effectively handles the presence of zero trade flows

Eight challenges

4. Bilateral trade costs

Challenge 4: Bilateral trade costs

- Proper specification of bilateral trade costs is crucial for partial equilibrium as well as for general equilibrium trade policy analysis

Solutions to challenge 4

(i) The standard practice suggested in the literature is to proxy for the bilateral trade cost term appearing in the structural gravity specification (11), $(1-\sigma)\ln t_{ij,t}$, by using a series of observable variables most of which have become standard covariates in empirical gravity specifications, namely:

$$(1-\sigma)\ln t_{ij,t} = \beta_1 lnDIST_{ij} + \beta_2 CNTG_{ij} + \beta_3 LANG_{ij} + \\ + \beta_4 CLNY_{ij} + \beta_5 RTA_{ij,t} + \beta_6 \widetilde{\tau}_{ij,t} \quad (12)$$

- $\ln DIST_{ij}$ is the logarithm of bilateral distance between trading partners i and j
- $CNTG_{ij}$ is an indicator variable that captures the presence of contiguous borders between countries i and j
- $LANG_{ij}$ and $CLNY_{ij}$ are dummy variables that take the value of one for common official language and for the presence of colonial ties, respectively
- $RTA_{ij,t}$ is a dummy variable that takes value one if an RTA between trading partners i and j is in place at time t, and zero otherwise

Solutions to challenge 4 (ct'd)

- The term $\widetilde{\tau}_{ij,t}$ accounts for bilateral tariffs and is defined as $\widetilde{\tau}_{ij,t} \equiv \ln\left(1 + \text{tariff}_{ij,t}\right)$, where $\text{tariff}_{ij,t}$ is the ad-valorem tariff that country j imposes on imports from country i at time t
- Importantly, since tariffs act as direct price shifters (See Appendix A.2 to Chapter 1 of the Advanced Guide for details), the coefficient on $\widetilde{\tau}_{ij,t}$ can be expressed only in terms of the trade elasticity of substitution $\beta_6 = -\sigma$
- This means that the trade elasticity itself can be recovered directly from the estimate on $\widetilde{\tau}_{ij,t}$, as $\widehat{\sigma} = -\widehat{\beta_6}$
- Note that this interpretation depends on trade flow data being expressed at cost, insurance and freight (c.i.f) prices, not inclusive of tariffs (see Appendix A.2 to Chapter 1 of the Advanced Guide for details)

Eight challenges

5. Endogeneity of trade policy

Challenge 5: Endogeneity of trade policy

- The trade policy variables $RTA_{ij,t}$ and $\widetilde{\tau}_{ij,t}$ are endogenous
- For instance, trade policy variables may suffer from "reverse causality", because, all else equal, a given country is more likely to liberalize its trade with another country that is already a significant trade partner

Solutions to challenge 5

(i) The issue of endogeneity of trade policy is well-known in the trade literature (Trefler, 1993)

(ii) However, primarily due to the lack of reliable instruments, early attempts to account for endogeneity of trade policy (Trefler, 1993) – in particular with reference to RTAs, Baier and Bergstrand (2007) propose applying the average treatment effect (ATE) methods described in Wooldridge (2010)

- In panel trade data, they suggest *first-differencing* bilateral trade flows or using *country-pair* fixed effects
- This eliminates or accounts for, respectively, the unobservable linkages between the endogenous trade policy covariate and the error term in gravity regressions
- Note that the set of pair fixed effects will absorb all bilateral time-invariant covariates (e.g. bilateral distance), but will have the plus of accounting for any unobservable time invariant trade cost components
- Egger and Nigai (2015) and Agnosteva et al. (2014) show that the pair-fixed effects are a better measure of bilateral trade costs than the standard set of gravity variables

Eight challenges

6. Non-discriminatory trade policy

Challenge 6: Non-discriminatory trade policy

- Researchers have struggled to estimate the effects of non-discriminatory trade policies, export subsidies or most-favoured-nation (MFN) tariffs, within the structural gravity model
- Non-discriminatory trade policy covariates are exporter- and/or importer-specific, and therefore they will be absorbed, respectively, by the exporter-time and by the importer-time fixed effects that need to be used in order to control for the multilateral resistances
- More generally, in the presence of importer and exporter fixed effects, the gravity model can no longer estimate the impact of any variable (i) affecting exporters' propensity to export to all destinations (e.g. being an island); (ii) affecting imports without regard to origin (e.g. country-level average applied tariff); and (iii) representing sums, averages, and differences of country-specific variables (Head and Mayer, 2014)

Solutions to challenge 6

(i) Approximate the multilateral resistances with the "remoteness" indexes rather than including directional (exporter and importer) fixed effects

- This approach is not recommended because it does not account properly for the multilateral resistance terms, and is therefore likely to produce biased gravity estimates (including the effects of trade policy)

(ii) Employ a two-stage estimation, where the estimates of the multilateral resistances from the first-stage gravity regression are explained in an auxiliary regression that includes the non-discriminatory covariate of interest (Anderson and Yotov, 2016; Head and Mayer, 2014)

Solutions to challenge 6 (ct'd)

(iii) As suggested by Heid et al. (2015), estimate the structural gravity model with international *and* intra-national trade flows by capitalizing on the fact that while non-discriminatory trade policies are country-specific, they do not apply to intra-national trade

- The estimates of non-discriminatory trade policies in the structural gravity model are less likely to be subject to endogeneity concerns as compared to their bilateral counterparts for two reasons:
  (a) It is unlikely that a non-discriminatory trade policy will be influenced by any bilateral trade flow
  (b) The directional fixed effects in the structural gravity model will absorb much of the unobserved correlation between the non-discriminatory trade policy covariates and the gravity error term

Eight challenges

7. Adjustment to trade policy changes

Challenge 7: Adjustment to trade policy changes

- The adjustment of trade flows in response to trade policy changes will not be instantaneous
- Trefler (2004) criticizes trade estimations pooled over consecutive years. The challenge of adjustment is even more pronounced in econometric specifications with fixed effects such as the ones we described, which allow identification through changes over years
- As noted in Cheng and Wall (2005), fixed-effects estimation applied to data pooled over consecutive years is sometimes criticized on the grounds that dependent and independent variables cannot fully adjust in a single year's time

Solutions to challenge 7

(i) Researchers have used panel data with intervals instead of data pooled over consecutive years

- For example, Trefler (2004) uses 3-year intervals, Anderson and Yotov (2016) use 4-year intervals, and Baier and Bergstrand (2007) use 5-year intervals
- Olivero and Yotov (2012) provide empirical evidence that gravity estimates obtained with 3-year and 5-year interval trade data are very similar, while estimations performed with panel samples pooled over consecutive years produce suspicious estimates of the trade cost elasticity parameters

Eight challenges

8. Gravity with disaggregated data

Challenge 8: Gravity with disaggregated data

- Many trade policies are negotiated and applied at the sectoral level, such as tariffs
- While it is in principle possible to aggregate trade policy and still use the aggregate gravity model, such aggregation practices should be avoided and, whenever possible, gravity should be estimated at the level of aggregation which is the target of the specific policy
- Furthermore, even for policies that are negotiated at the aggregate level (e.g. some RTAs), it may be desirable to also obtain sectoral effects because the effects of these non-discriminatory policies may actually be quite heterogeneous across sectors

Solutions to challenge 8

(i) Structural gravity is separable: bilateral expenditures across countries both at the aggregate and at the sectoral level are separable from output and expenditure at the country level (Larch and Yotov, 2016b)

- As demonstrated by Anderson and van Wincoop (2004), for a given set of country-level output ($Y^k_{i,t}$) and expenditure ($E^k_{j,t}$), where k denotes a class of goods/sector, theory delivers the familiar sectoral gravity equation:

$$X^k_{ij,t} = \frac{Y^k_{i,t} E^k_{j,t}}{Y^k_t} \left(\frac{t^k_{ij,t}}{\Pi^k_{i,t} P^k_{j,t}} \right)^{(1-\sigma_k)} \quad (13)$$

- Equation (13) can be estimated for each sector as if the data were aggregate
- Alternatively, the gravity model can be estimated with data pooled across sectors, in which case the proper treatment of the multilateral resistance requires *exporter-product-time* and *importer-product-time* fixed effects
- Depending on the question of interest, the estimates of the trade policy variables in gravity estimations that are pooled across sectors can be sector-specific or constrained to be common across sectors

Six recommendations

1. Use panel data
2. Use year intervals
3. Add intra-national trade flows
4. Include importer-time and exporter-time fixed effects
5. Include pair fixed effects
6. Estimate the model with PPML

Six recommendations

1. Use panel data

Recommendation 1: Use panel data

(i) Using panel data leads to improved estimation efficiency

(ii) The panel dimension enables to apply the pair-fixed-effects methods to address the issue of endogeneity of trade policy variables (Baier and Bergstrand, 2007)

(iii) The use of panel data allows for a flexible and comprehensive treatment and estimation of the effects of time-invariant bilateral trade costs with pair fixed effects

- What if panel data are not available? See next slide

In the absence of panel data

- In the absence of panel trade data, the gravity model can still be estimated with cross-section samples:

$$\ln X_{ij} = \ln X_{ij} + \ln Y_i - \ln Y + (1-\sigma)\ln t_{ij} - (1-\sigma)\ln P_j - (1-\sigma)\ln \Pi_i + \varepsilon_{ij}$$

- In a cross-section setting, recommendations 3, 4, and 6 continue to hold, namely gravity specification should include intra-national trade and directional (importer and exporter) fixed effects, and be estimated applying the PPML estimator
- However, the recommendations 2 and 5 to allow for adjustment in trade flows by using interval data and to include pair fixed effects are no longer applicable
- The gravity specification in cross-section should include the standard set of gravity variables (e.g., bilateral distance, contiguity...) instead of pair fixed effects, in order to proxy for bilateral trade costs
- That being said, the error term may capture systematic effects of unobserved trade costs
- In order to address the endogeneity of bilateral trade policy, IV treatment is highly recommended (Baier and Bergstrand, 2004; Egger et al., 2011)

Six recommendations

2. Use year intervals

Recommendation 2: Use year intervals

- Interval panel data should be employed in order to allow for adjustment in bilateral trade flows in response to trade policy or other changes in trade costs
- Olivero and Yotov (2012) find that gravity estimates obtained with 3-, 4-, and 5-year lags deliver similar results
- It is recommended to experiment with alternative intervals while keeping estimation efficiency in mind

Six recommendations

3. Add intra-national trade flows

Recommendation 3: Add intra-national trade flows

- The inclusion of intra-national trade data in structural gravity estimations is desirable for several reasons:
  (i) It ensures consistency with gravity theory, where consumers choose among and consume domestic as well as foreign varieties
  (ii) It leads to the theoretically consistent identification of the effects of bilateral trade policies (Dai et al., 2014)
  (iii) It enables to identify and estimate the effects of non-discriminatory trade policies (Heid et al., 2015)
  (iv) It resolves the "'distance puzzle" in trade, by measuring the effects of distance on international trade relative to the effects of distance on internal trade (Yotov, 2012)
  (v) It enables to capture the effects of globalization on international trade and to correct for biases in the estimation of the impact of RTAs on trade (Bergstrand et al., 2015)
- Intra-national trade data has to be constructed consistently as the difference between gross production value data and total exports

Six recommendations

4. Include importer-time and exporter-time fixed effects

Recommendation 4: Include importer-time and exporter-time fixed effects

- The use of exporter-time and importer-time fixed effects enables to control for the unobservable multilateral resistances, and potentially for any other observable and unobservable characteristics that vary over time for each exporter and importer, respectively (Anderson and van Wincoop, 2003)
- The estimates of the fixed effects of the gravity model can be used directly to recover the estimates of the general equilibrium effects of trade policy changes as well as to construct a series of useful general equilibrium indexes summarizing and aggregating consistently the effects of trade policy and trade costs (Anderson et al., 2015b; Larch and Yotov, 2016b)
- (See Chapter 2 of the Advanced Guide)

Six recommendations

5. Include pair fixed effects

Recommendation 5: Include pair fixed effects

- Two major benefits are associated with using pair fixed effects in gravity estimations:
  - (i) The pair fixed effects are able to account for the endogeneity of trade policy variables (Baier and Bergstrand, 2007)
  - (ii) The pair fixed effects provide a flexible and comprehensive account of the effects of all time-invariant bilateral trade costs, observable and unobservable (Egger and Nigai, 2015; Agnosteva et al., 2014).
- The downside is that one cannot identify the effects of any time-invariant bilateral determinants of trade flows, which will be absorbed by the pair fixed effects
- One way to address this issue is to apply a two-stage procedure, where the estimates of the pair fixed effects from the first-stage structural gravity equation are regressed on standard gravity variables in a second-stage estimation (Agnosteva et al., 2014)
- This two-step approach also enables to recover estimates of the pair fixed effects that cannot be identified directly in the first stage, due to missing or zero trade flows, allowing to construct the full matrix of bilateral trade costs and to perform counterfactual experiments (Anderson and Yotov, 2016)

Six recommendations

6. Estimate the model with PPML

Recommendation 6: Estimate the model with PPML

- The use of the PPML estimator is justified on various grounds:
  - (i) The PPML estimator, applied to the gravity model expressed in a multiplicative form, accounts for heteroskedasticity, which often plagues trade data (Santos Silva and Tenreyro, 2006)
  - (ii) The PPML estimator is able to take advantage of the information contained in the zero trade flows
  - (iii) The additive property of the PPML estimator ensures that the gravity fixed effects are identical to their corresponding structural terms (Arvis and Shepherd, 2013; Fally, 2015)
  - (iv) The PPML estimator can also be used to calculate theory-consistent general equilibrium effects of trade policies (Anderson et al., 2015b; Larch and Yotov, 2016b)
- As a robustness check, the gravity model can be estimated by applying the Gamma Pseudo Maximum Likelihood (GPML) and the OLS estimators (Head and Mayer, 2014)

Structural gravity estimating equation

- The best practices and recommendations proposed are reflected in the following generic and comprehensive econometric version of the structural gravity model, which can be modified and adjusted by researchers and policy makers depending on their specific needs:

$$X_{ij,t} = \exp[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \eta_1 BTP_{ij,t} + \eta_2 NES_{i,t} \times INTL_{ij} + \eta_3 NIP_{j,t} \times INTL_{ij}] \times \varepsilon_{ij,t} \tag{14}$$

where:

- $X_{ij,t}$ denotes nominal trade flows, which include international and intra-national trade, at non-consecutive year t
- The terms $\pi_{i,t}$, $\chi_{j,t}$ and μ_{ij} denote exporter-time, importer-time and pair fixed effects

Structural gravity estimating equation (ct'd)

$$X_{ij,t} = \exp[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \eta_1 BTP_{ij,t} + \eta_2 NES_{i,t} \times INTL_{ij} + \eta_3 NIP_{j,t} \times INTL_{ij}] \times \varepsilon_{ij,t}$$

- The term $BTP_{ij,t}$ represents the vector of any time-varying bilateral determinants of trade flows, such as RTAs, bilateral tariffs and currency unions
- In principle, the $BTP_{ij,t}$ vector may include any time-varying covariates

Structural gravity estimating equation (ct'd)

$$X_{ij,t} = \exp[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \eta_1 BTP_{ij,t} + \eta_2 NES_{i,t} \times INTL_{ij} + \eta_3 NIP_{j,t} \times INTL_{ij}] \times \varepsilon_{ij,t}$$

- The expression $NES_{i,t} \times INTL_{ij}$ is the product between $NES_{i,t}$ and $INTL_{ij}$
- $NES_{i,t}$ denotes the vector of any Non-discriminatory Export Support (NES) policies, such as export subsidies
- $INTL_{ij}$ is an "international trade" dummy variable. It takes a value of one for international trade between countries i and $j \neq i$, and zero otherwise (i.e. when the exporer and the importer are the same country, therefore trade is intra-national)
- The interaction term $NES_{i,t} \times INTL_{ij}$ varies along the ij, t dimension, allowing to identify the effects of any non-discriminatory export support policies, even in the presence of exporter-time fixed effects as required by gravity theory (Heid et al., 2015)
- With appropriate data on export support measures that act as direct price-shifters, the estimate of the coefficient(s) associated with the variable(s) $NES_{i,t} \times INTL_{ij}$ can be used to recover an estimate of the export supply elasticity, which has attracted little attention in the empirical trade literature

Structural gravity estimating equation (ct'd)

$$X_{ij,t} = \exp[\pi_{i,t} + \chi_{j,t} + \mu_{ij} + \eta_1 BTP_{ij,t} + \eta_2 NES_{i,t} \times INTL_{ij} + \eta_3 NIP_{j,t} \times INTL_{ij}] \times \varepsilon_{ij,t}$$

- The covariate $NIP_{j,t} \times INTL_{ij}$ is constructed as the product between the term $NIP_{j,t}$, which denotes the vector of any Non-discriminatory Import Protection (NIP) policies, such as MFN tariffs, and the dummy for bilateral international trade $INTL_{ij}$
- Given its bilateral nature, the interaction $NIP_{j,t} \times INTL_{ij}$ can be used to identify the effects of any non-discriminatory import protection policies

Interpretation of gravity estimates

- Two related methods are widely used to interpret the estimates from gravity regressions:
  1. Use the gravity estimates to construct *trade volume effects*
  2. Convert the estimates of various trade policies and other determinants of trade flows into *tariff equivalent effects*
- Consider the following simplified version of the empirical gravity model (14):

$$X_{ij,t} = \exp\left[\pi_{i,t} + \chi_{j,t} + \beta_{DIST} \ln DIST_{ij} + \beta_{RTA} RTA_{ij,t} + \beta_{TARIFF} \widetilde{\tau}_{ij,t}\right] \times \varepsilon_{ij,t} \quad (15)$$

where $DIST_{ij}$ and $RTA_{ij,t}$ were defined in slide 32 and $\widetilde{\tau}_{ij,t}$ was defined in slide 33

Trade volume effects

1. Trade volume effect of continuous variables

- The interpretation of the estimate of the coefficient on the logarithm of the continuous variable is simply the elasticity of (the value of trade flows) with respect to the continuous variable
- For example, the standard empirical value for the distance variable estimate in gravity regressions of $\widehat{\beta}_{DIST} = -1$ implies that a 10 percent increase in distance should be accompanied by a 10 percent decrease in trade flows (Disdier and Head, 2008; Head and Mayer, 2014)

Trade volume effects (ct'd)

2. Trade volume effect of dummy variables

- The volume effects triggered by a change in a dummy variable, such as the presence of RTAs, can be calculated in percentage terms as follows:

$$\left[e^{\widehat{\beta}_{DUMMY}} - 1\right] \times 100 \tag{16}$$

- To derive the formula in (17), note that:
  - $\ln X_{ij}(1)$ is the predicted value of trade when the dummy $= 1$, while $\ln X_{ij}(0)$ is the value of trade when dummy $= 0$
  - The difference $\ln X_{ij}(1) - \ln X_{ij}(0) = \widehat{\beta}_{DUMMY}$
  - Therefore, $X_{ij}(1)/X_{ij}(0) = e^{\widehat{\beta}_{DUMMY}}$
  - This in turn implies that the percentage change in trade value due to the dummy switching from 0 to 1 is:
    $100 \times [X_{ij}(1) - X_{ij}(0)]\,/X_{ij}(0) = \left[e^{\widehat{\beta}_{DUMMY}} - 1\right] \times 100$
- Alternatively, one can compute $\left\{\frac{e^{\widehat{\beta}_{DUMMY}}}{e^{\left[\frac{1}{2}\widehat{\text{var}}(\beta_{DUMMY})\right]}} - 1\right\} \times 100$, which is consistent and (almost) unbiased

Trade volume effects (ct'd)

$$\left[e^{\widehat{\beta}_{DUMMY}} - 1\right] \times 100 \quad (17)$$

- For example, the benchmark estimate of the effects of RTAs in gravity regressions found in the empirical literature of $\widehat{\beta}_{RTA} = 0.76$ implies that the RTAs that entered into force between 1960 and 2000 on average have increased trade by $\left[e^{0.76} - 1\right] \times 100 = 114$ percent (Baier and Bergstrand, 2007)

A useful decomposition

- With the exception of the direct price shifters, such as tariffs, the estimates of the remaining gravity covariates consist of two components: (i) a structural component and (ii) a trade cost component
- For example, the structural interpretation of the estimate of the coefficient of distance is $\widehat{\beta}_{DIST} = (1 - \widehat{\sigma})\,\rho$, where ρ is the elasticity of trade costs with respect to distance
- This decomposition enables to recover the distance elasticity of trade costs as: $\rho = \widehat{\beta}_{DIST} / (1 - \widehat{\sigma})$

Tariff equivalent effects

- Tariff equivalent effect of quota dummy (ad-valorem tariff equivalent whose removal would have generated the same impact as the removal of the quota):

$$\left[e\left(\frac{\widehat{\beta}_{QUOTA}}{\widehat{\beta}_{TARIFF}}\right) - 1 \right] \times 100$$

- Tariff equivalent effect of RTAs (average tariff equivalent fall of the introduction of RTAs):

$$\left[e\left(\frac{\widehat{\beta}_{RTA}}{-\widehat{\beta}_{TARIFF}}\right) - 1 \right] \times 100 \qquad (18)$$

- (See pp. 128-130 of the Practical Guide) and slides on Tariff Equivalents

Tariff equivalent effects (ct'd)

- In principle, since $\beta_{TARIFF} = -\sigma$ (see slide 33), no data on tariffs are needed in order to obtain tariff equivalent effects of other gravity covariates as long as reliable estimates of the trade elasticity of substitution are available from outside studies
- Returning to the example of the effects of RTAs from Baier and Bergstrand (2007), and taking a representative value for the elasticity of substitution from the literature, $\sigma = 5$, the average tariff-equivalent fall of the introduction of RTAs would amount to $\left[e^{\left(\frac{0.76}{5}\right)} - 1\right] \times 100 = 16.4$ percent

Bilateral aggregate trade flows data

- IMF Direction of Trade Statistics (DOTS)
    - The database covers 184 countries
    - Annual data are available from 1947
    - Data are reported in US dollars

Merchandise trade flows data

- UN COMTRADE
  - The database covers more than 160 countries
  - Data are reported in US dollars
  - HS (up to 6 digits) and SITC classifications for disaggregate data
- CEPII's BACI
  - "Cleaned" version of COMTRADE (mirror data used when informative, re-exports excluded)
  - The database covers more than 200 countries
  - Time lag of 1-2 years w.r.t. COMTRADE
- Feenstra and Romalis (2014) World Trade Flows database
  - Omits observations with cif/fob ratio $\in (0.1, 10)$ and when cif $<$ 50'000 USD
  - The database covers 185 countries for the period 1984-2014

Services trade flows data

- OECD Trade in Services database
  - Bilateral services trade for 12 main services sectors and several sub-sectors according to the Extended Balance of Payments (EBOPS) 2010 classification
  - The database covers 35 countries including 32 OECD member countries as well as the Russian Federation, Colombia and Latvia from 1999 onwards
- UN Service Trade Database
  - Covers 46 economies from 2000 onwards and follows the EBOPS 2002 classification
- WTO-UNCTAD-ITC database
  - Bilateral annual service flows data for 36 countries at the same level of disaggregation as the OECD data from 2005 onwards according to the EBOPS 2010 classification
  - These bilateral data can be retrieved from the ITC TradeMap
- Database by Francois and Pindyuk (2013)
  - Based on adjusted data from the OECD, Eurostat, UN and IMF
  - Bilateral service flows data classified according to the EBOPS 2002 classification and covering 248 countries, on average, for the period 1981-2010

Agriculture and resource sectors data

- FAOSTAT database
  - Information on an annual basis in more than 100 countries
  - The Detailed Trade Matrix reports information on agricultural bilateral trade flows for over 600 food and commodities per year
  - Data for both quantities (in tons) and values (in thousands of US dollars) of agricultural imports and exports
  - Data are available for the period 1986-2013 and are gathered from national sources

Intra-national trade flows data

- Big problem: one cannot use apparent consumption (defined as the difference between production and total exports)...rather, gross production should be used
- This is because aggregate production data are usually measured and reported as value added (e.g. GDP), while total exports are reported as gross value

Intra-national trade flows data

- World Bank's Trade, Production and Protection (TPP) database
  - Covers approximately 100 countries for the period 1976-2004
  - Information according to the Industrial Standard Industrial Classification (ISIC) Rev. 3 classification at the 3-digit level
- CEPII's Trade, Production and Bilateral Protection (TradeProd) database
  - Covers over 150 countries for the period 1980âĂŞ2006
  - Data in ISIC Rev. 2 at the 3-digit level
- CEPII's Trade, Production and Bilateral Protection (TradeProd) database
  - Covers over 150 countries for the period 1980âĂŞ2006
  - Data in ISIC Rev. 2 at the 3-digit level
- UNIDO Industrial Statistics (INDSTAT)
  - Covers 166 countries
  - INDSTAT2: Data from 1962 onwards at the 2-digit level of ISIC Rev. 3
  - INDSTAT4: Data from 1990 onwards at the 4-digit level of ISIC Rev. 3

Databases used in this structural gravity course

1. The database "gravity_dataset_AG_TPA_applications.dta" is a balanced panel covering the aggregate manufacturing sector of 69 countries over the period 1986-2006

- Sources:
  - Consistently constructed international and intra-national trade flows data, which were assembled and provided by Thomas Zylkin
  - The original sources for the international trade data are the UN COMTRADE database and the CEPII TradeProd database
  - Intra-national trade for each country is constructed as the difference between total manufacturing production and total manufacturing exports
  - Both of these variables are reported on a gross basis, which ensures consistency between intra-national and international trade
  - Three sources are used to construct the production data: the UN UNIDO INDSTAT database, the CEPII TradeProd database, and the World Bank's TPP database

Databases used in this structural gravity course (ct'd)

1. Database "gravity_dataset_AG_TPA_applications.dta" (ct'd)

- Sources (ct'd):
    - The data on RTAs were taken from Mario Larch's Regional Trade Agreements Database
    - All standard gravity variables including distance, contiguous borders, common language, and colonial ties are from the CEPII GeoDist database
    - An important advantage of the GeoDist database is that the weighted-average methods used to construct distance ensure consistency between the measures of intra-national and international distance

Databases used in this structural gravity course (ct'd)

2. Database "gravity_dataset_AG_TPA_exercise1.dta'

- Derived from "gravity_dataset_AG_TPA_applications.dta"
- 65 countries, 6 years (1986, 1990, 1994, 1998, 2002, 2006)
- Includes WTO membership data: $WTO_{ij,t}$ is a dummy variable taking value one if both the importer and the exporter are WTO members in year t (zero otherwise and in case of intra-national trade)

3. Database "gravity_dataset_AG_TPA_exercise2.dta'

- Derived from "gravity_dataset_AG_TPA_applications.dta"
- 52 countries, 7 years (1988, 1991, 1994, 1997, 2000, 2003, 2006)
- Includes tariff data

Databases used in this structural gravity course (ct'd)

4. Database "exporter_WB_DoingBusiness_time_docs.dta"
   - From World Bank Doing Business indicators
   - Data on number of documents to export and time to export (in days)
   - Available for 175 countries
5. Database "importer_WB_DoingBusiness_time_docs.dta"
   - From World Bank Doing Business indicators
   - Data on number of documents to import and time to import (in days)
   - Available for 175 countries