The gravity models for trade research

ARTNeT Capacity Building Workshop on

“the Use of Gravity Modelling”

19-22 March 2013
ESCAP

Bangkok, Thailand

Dr. Witada Anukoonwattaka
Trade and Investment Division, ESCAP
anukoonwattaka@un.org

UNITED NATIONS ESCAP
Economic and Social Commission for Asia and the Pacific

ARTNeT

Outline

- The intuitive gravity model and its problems
- Theoretical Gravity models
- Estimating theoretical gravity models
  - FE approach
  - Baier-Bergstrand approach
- Often-made mistakes

The intuitive gravity model

What is the gravity model?

- Gravity model is a very popular econometric model in international trade
- The name came from its utilizing the gravitational force concept as an analogy to explain the volume of bilateral trade flows
  – Proposed by Tinbergen (1962)
- Initially, it was not based on theoretical model, but just intuition only
- Later on, a range of rigorous theoretical foundation has been given.
  – The most well-known benchmark so far is Anderson and van Wincoop (2003).

Gravity Analogy

Gravity force equation

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

Gravity force between two objects depends on their masses and inversely proportional to the square of distance between them.

Intuitive gravity for trade

$$X_{ij} = C\frac{Y_i Y_j}{t_{ij}}$$

X_{ij} = exports (or trade) from i to j,

C = contstant,

Y = economic mass (≈ GDP),

t = trade costs between two countries

≈ distance, adjacency, ., "policy factors".

Export (or trade) between two countries depends on their economic masses and negatively related to trade costs between them.

Intuitive gravity model of trade:

$$X_{ij} = C \frac{Y_i \; Y_j}{t_{ij}}$$

- Larger countries trade more than smaller ones
- Trade costs between two trade partners reduce trade between them.

Empirical equation for basic gravity model:

$$\ln X_{ij} = b_0 + b_1 \ln(Y_i) + b_2 \ln(Y_j) + b_3 \ln(t_{ij}) + e_{ij}$$

$$b_1, b_2 > 0; \quad b_3 < 0$$

A 1% change in Y_i is associated with a b_1% chage in X_{ij}.

Proxies for trade costs

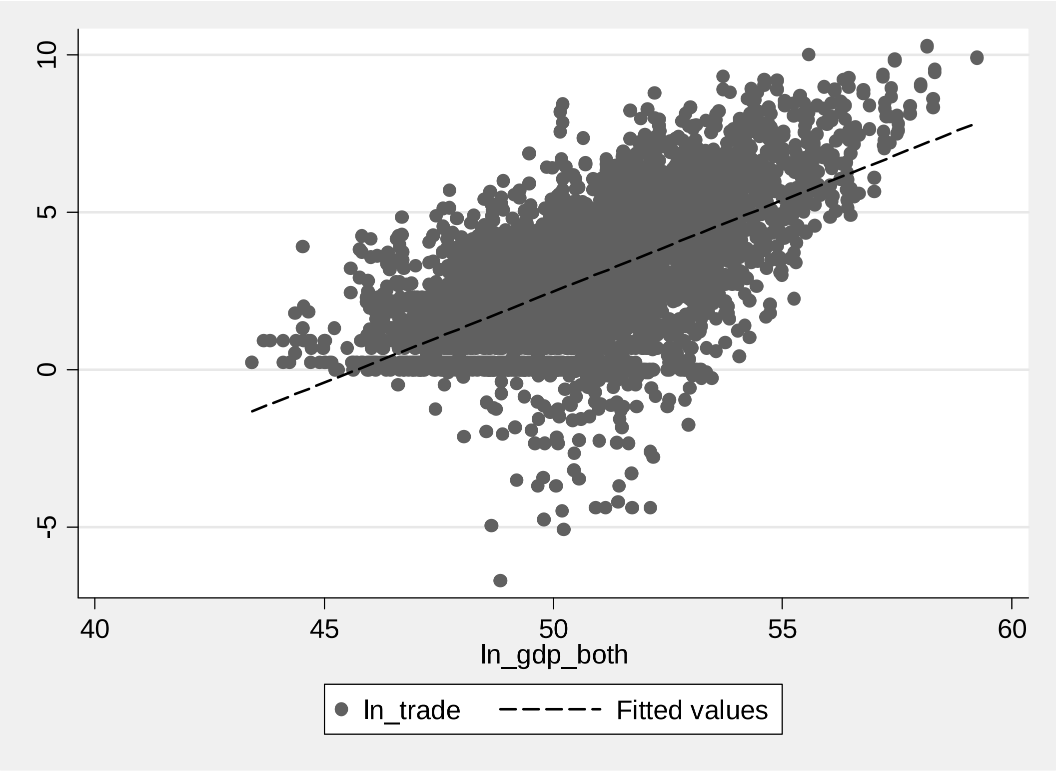
- Distance
- Adjacency
- Common language
- Colonial links
- Common currency
- Island, landlocked
- Institutions, infrastructures, migration flows,..
- Bilateral tariff barriers

Why is it so popular?

- Intuitively appealing
- Fits with some important stylized facts
- Easily to use real data to explain trade flows with respect to policy factors.
- Estimation using OLS

Trade and combined GDP

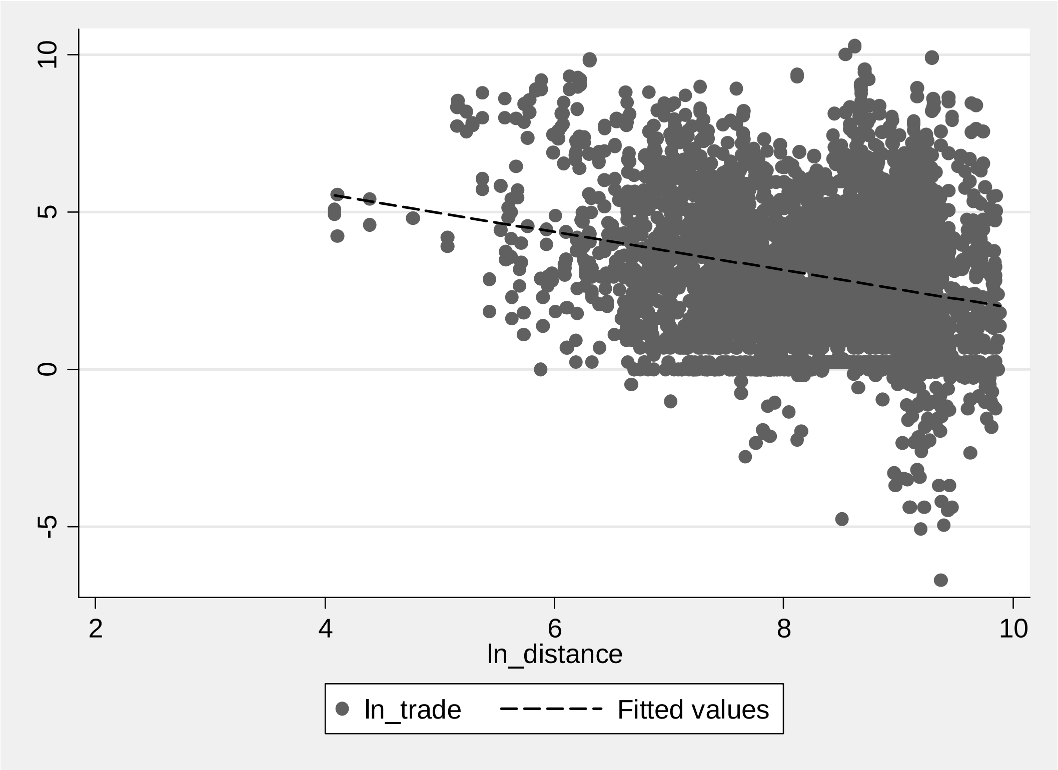
```
. gen ln_gdp_both = ln(gdp_exp*gdp_imp)
(1983 missing values generated)

. twoway (scatter ln_trade ln_gdp_both if sector == "SER") (lfit ln_trade
ln_gdp_both if sector == "SER")
```



Source: Based on services trade data provided by Shepherd (2012)

Trade and distance

```
. twoway (scatter ln_trade ln_distance if sector == "SER") (lfit ln_trade
ln_distance if sector == "SER")
```



Source: Based on services trade data provided by Shepherd (2012)

A simple OLS gravity model

```
. regress ln_trade ln_gdp_exp ln_gdp_imp ln_distance contig comlang_off colony
comcol if sector=="SER", robust cluster(dist)

Linear regression                                      Number of obs =    3884
                                                       F(  7,  2151) =  442.01
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5431
                                                       Root MSE      =  1.5281

                             (Std. Err. adjusted for 2152 clusters in dist)
------------------------------------------------------------------------------
             |               Robust
    ln_trade |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  ln_gdp_exp |   .6011672   .0132209    45.47   0.000     .5752401    .6270942
  ln_gdp_imp |   .6176062   .0142666    43.29   0.000     .5896284    .6455839
 ln_distance |  -.7385146     .03536   -20.89   0.000    -.8078579   -.6691714
      contig |   .3989524   .1829276     2.18   0.029     .0402191    .7576858
 comlang_off |   .8861328   .0993078     8.92   0.000     .6913835    1.080882
      colony |   1.202965   .1201971    10.01   0.000     .9672503     1.43868
      comcol |  -.0245067   .2018195    -0.12   0.903    -.4202883     .371275
       _cons |  -22.03706    .671738   -32.81   0.000    -23.35438   -20.71974
------------------------------------------------------------------------------
```

Interpretation

- Market-size effects

ln_trade	Coef.	Robust Std. Err.	t	P>|t|
ln_gdp_exp	.6011672	.0132209	45.47	0.000
ln_gdp_imp	.6176062	.0142666	43.29	0.000

- A 1% increase in exporter size is associated with about a 0.6% increase in bilateral trade.
- A 1% increase in importer size is associated with about a 0.6% increase in bilateral trade.
- Both effects are statistically significant at the 1% level

Interpretation

- Geography effects

ln_trade	Coef.	Robust Std. Err.	t	P>|t|
ln_distance	-.7385146	.03536	-20.89	0.000
contig	.3989524	.1829276	2.18	0.029

- A 1% increase in distance between markets is associated with about a 0.7 % decrease in bilateral trade.
- Sharing a common border is associated with a 49% increase in bilateral trade (e^0.4-1≈0.49).
- Distance is statistically significant at the 1% level, contingency is statistical significant at 5% level.

Interpretation

- History effects

ln_trade	Coef.	Robust Std. Err.	t	P>|t|
comlang_off	.8861328	.0993078	8.92	0.000
colony	1.202965	.1201971	10.01	0.000
comcol	-.0245067	.2018195	-0.12	0.903
_cons	-22.03706	.671738	-32.81	0.000

- A common official language is associated with an increase in bilateral trade of about a 145 % (e^0.9-1≈1.45).
- Having once been a colony is associated with a 232% increase in bilateral trade (e^1.2-1≈2.32).

An augmented gravity model

```
. regress ln_trade etcr_exp etcr_imp ln_gdp_exp ln_gdp_imp ln_distance contig
comlang_off colony comcol if sector == "SER", robust cluster(dist)
note: comcol omitted because of collinearity

Linear regression                                      Number of obs =     816
                                                       F(  8,   413) =  139.24
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6833
                                                       Root MSE      =  1.3835

                                  (Std. Err. adjusted for 414 clusters in dist)
------------------------------------------------------------------------------
             |               Robust
    ln_trade |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    etcr_exp |  -.3605257   .0910402    -3.96   0.000    -.5394858   -.1815657
    etcr_imp |  -.3721994   .0796389    -4.67   0.000    -.5287475   -.2156512
  ln_gdp_exp |   .7736852   .0349451    22.14   0.000     .7049927    .8423777
  ln_gdp_imp |   .8223475   .0349431    23.53   0.000      .753659     .891036
 ln_distance |  -1.114939   .0626474   -17.80   0.000    -1.238087   -.9917915
      contig |  -.5579995   .2452544    -2.28   0.023    -1.040102    -.075897
 comlang_off |   1.378263   .2090961     6.59   0.000     .9672377    1.789289
      colony |   .1771059   .2077632     0.85   0.394    -.2312993    .5855111
      comcol |          0  (omitted)
       _cons |  -27.11023   1.799003   -15.07   0.000    -30.64657   -23.57388
------------------------------------------------------------------------------
```

Interpretation

- Trade-policy effects

• The OECD’s indicators of Energy, Transport, Communication Regulations (ETCR) are used (score from 0 to 6).

•A one point increase in a country’s ETCR score is associated with about a 36-37% decrease in bilateral trade.

• Both exporter and importer ‘s ETCR scores are statistically significant at the 1% level

Examples of Applications

- Effects of regional integration on trade

Do RTAs boost trade between members?

$$\ln X_{ij} = b_0 + b_1 \ln(Y_i) + b_2 \ln(Y_j) + b_3 \ln(t_{ij}) + b_4 (dummyRTA_{ij}) + e_{ij}$$

Do RTAs reduce exports from non-members?

$$\ln X_{ij} = b_0 + b_1 \ln(Y_i) + b_2 \ln(Y_j) + b_3 \ln(t_{ij}) + \ldots$$
$$+ b_4 (dummy\ BothInRTA_{ij}) + b_5 (dummy\ OneInRTA_{ij}) + e_{ij}$$

See, World Bank (2005) for example.

Examples of Applications

- Effects of trade facilitation on trade

How much can trade facilitation boost bilateral trade?

$$\ln X_{ij} = b_0 + b_1 \ln(Y_i) + b_2 \ln(Y_j) + b_3 \ln(t_{ij}) + b_4 \ln(time_i^X) + e_{ij}$$

See, Djankov etal.(2010) for example.

Examples of Applications

- Effects of institutional weakness on trade

How does corruption affect trade?

$$\ln X_{ij} = b_0 + b_1 \ln(Y_i) + b_2 \ln(Y_j) + b_3 \ln(t_{ij}) + b_4 \ln(corruption_i) + e_{ij}$$

See, Anderson & Marcouiller(2002) for example.

Identifying “trade potential”

- The OLS estimates give the prediction of “average” trade level.
- Actually, some countries trade “more” than average, while others trade “less” than average.
- Some of the literature uses the sign and size of the error term to examine trade potential.

Identifying "trade potential"

- Estimating trade potential

$$\ln X_{ij} = b_0 + b_1 \ln(Y_i) + b_2 \ln(Y_j) + b_3 \ln(t_{ij}) + ... + e_{ij}$$

Fist step: estimate the model to get estimated coefficients

Second step: Use estimated coefficients give predicted Xij

$$\hat{X}_{ij} = \hat{b}_0 + \hat{b}_1 \ln(Y_i) + \hat{b}_2 \ln(Y_j) + \hat{b}_3 \ln(t_{ij}) + ...$$

Third: Trade potential is the gap between predicted and actual Xij

Identifying “trade potential”

- (Large) Negative errors:
  - A country could be trading more based on their economic and geographical fundamentals
  - Something is holding back trade.
- Positive errors: ??
- Keep in mind
  - The error term include also statistical noise and measurement error.
  - Do not overemphasized trade potential estimated by gravity model.
  - It may just give a first idea of what is going on with particular trade relationships. Things need to work in details of what is holding back trade.

Major weaknesses of the basic gravity:

$$\ln X_{ij} = b_0 + b_1 \ln(Y_i) + b_2 \ln(Y_j) + b_3 \ln(t_{ij}) + e_{ij}$$

The basic gravity model cannot handle the facts that:

1. Trade costs of the third party can affect trade between the two partners.
2. Relative trade costs (relative prices, to be exact) matter, not absolute trade costs

A consequence:

- The OLS basic gravity models encounter the omitted variables bias

Ex. trade creation and trade diversion are not captured by the basic gravity model.

Theoretical gravity models

The theoretical gravity model

- A number of papers try to fix those problems by laying theoretical foundations to gravity model
- For the starting point, we will focus on Anderson &Van Wincoop (AvW),2003. “The gravity with gravitas”.
- There are several theoretical gravity models developed for particular purposes.
  - Ex. Helpman et.al. (2008), Chaney (2008).
  - Showing complex empirical issues that OLS cannot handle
  - Sources such as the gravity course on Ben Shepherd’s website at http://www.developing-trade.com/ provide rich details.

AvW (2003)

- The most formal benchmark for theoretical gravity model so far
  - Bringing the gravity model a step closer to GE effects
  - Accounting for “relative price effects” on trade flow

*Things affecting **“relative price”** can influence bilateral trade flow. No matter the “things” happen between the two trading partners or happen with third parties.*

Building gravity on micro-foundation

- Consumers have love of variety preferences with CES structure
- Producers are under monopolistic competition
- Introducing trade costs, and related domestic and foreign prices
- Closing the model with macroeconomic accounting identities

AvW (2003) empirical model

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - \log Y^k + (1-\sigma_k)\left[\log \tau_{ij}^k - \log \Pi_i^k - \log P_j^k\right]$$

X_{ij}^k Sector-k exports from country i to j

Y^k World income from sector k

Y_i^k Exporting country's income from sector k

E_j^k Importing country's expenditure on sector k

τ_{ij}^k Sector-k trade costs between country i and country j

AvW (2003) empirical model

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - \log Y^k + (1-\sigma_k)\left[\log \tau_{ij}^k - \log \Pi_i^k - \log P_j^k\right]$$

Outward multi-lateral resistance (MTR): looking from export sides

$$\Pi_i^k = \sum_{j=1}^{C} \left\{\frac{\tau_{ij}^k}{P_j^k}\right\}^{1-\sigma_k} \frac{E_j^k}{Y^k}$$

Exports from country i to j depend on trade costs across ALL possible export markets.

Inward multi-lateral resistance (MTR): looking from import side

$$P_{\mathrm{j}}^k = \sum_{i=1}^{C} \left\{\frac{\tau_{ij}^k}{\Pi_i^k}\right\}^{1-\sigma_k} \frac{Y_i^k}{Y^k}$$

Imports of country j from country i depend on trade costs across ALL possible suppliers.

AvW (2003) empirical implications

- The intuitive gravity model have **omitted variables**.
  - Serious consequences (bias and inconsistency)
- "**Relative trade costs**" matter, NOT absolute trade costs,
  - Two types of trade costs have to be taken into account
    - Trade costs between i and j
    - Trade costs of i and j with third parties
- For each observation, trade flows are **"unidirectional"** trade flows, NOT total trade of a country pair.
- Variables are in **"nominal"** terms, NOT real terms
  - (Special) price indices are captured separately in the MRT terms
- Using **"aggregate"** GDP, NOT GDP per capita

AvW (2003) empirical implications

- Need the estimated "trade costs", NOT just distance in gravity models.
  - The base line can be

$$\log \tau_{ij}^{k} = b_1 \log distance_{ij} + b_2 contig + b_3 comlang\_off + b_4 colony + b_5 comcol$$

  - Policy-related variables can be augmented into the trade cost functions.
- $\hat{b}$ is NOT purely trade cost elasticity, but combined with elasticity of substitution (σ_k)

Estimating the AvW gravity model

$$\log X_{ij} = \log Y_i + \log Y_j - \log Y + (1-\sigma)\left[\log \tau_{ij} - \log \Pi_i - \log P_j\right]$$

- The **“multilateral trade resistance (MTR)”** terms are empirically unobservable
- Two possible strategies to estimate the model
  - Fixed effects (FE) estimation
  - Baier and Bergstrand (2009) approach

FE estimation

The theoretical gravity model:

$$\log X_{ij} = \log Y_i + \log Y_j - \log Y + (1-\sigma)\left[\log \tau_{ij} - \log \Pi_i - \log P_j\right]$$

FE transformation of the model:

$$\log X_{ij} = C + F_i + F_j + (1-\sigma)\left[\log \tau_{ij}\right]$$

where

$$C = -\log Y$$

$$F_i = \log Y_i - \log \Pi_i$$

Exporter fixed effects

$$F_j = \log Y_j - \log P_j$$

Importer fixed effects

$$\log \tau_{ij} = b_1 \log distance_{ij} + b_2 contig_{ij} + b_3 comlang\_off_{ij} + b_4 colony_{ij} + b_5 comcol_{ij}$$

Estimated trade cost

Estimation procedures

1. Generate (numerical) fixed effects dummies
   - Exporter & Importer dummies
   - Year dummies (if relevant)
   - Sector dummies (if relevant, but may be impractical)
2. Assuming key OLS assumptions are fine, we estimate the FE model with OLS
   - No multicolinearity
   - Homoskedasticity
   - E is uncorrelated with any independent variables

Estimation procedures

3. Variables that vary in the same dimensions as the FEs CANNOT be included in the model
    - Having multicolinearity problems with FEs
        - ex. MFN tariffs, ETCR scores
    - MUST be transformed into a new variable that vary bilaterally before including in the model,
    - or using alternative approaches.

Estimating a FE gravity model with OLS in Stata

```
. egen exporters = group(exp)

. egen importers = group(imp)

. regress ln_trade ln_distance contig comlang_off colony comcol i.exporters
i.importers if sector=="SER", robust cluster(dist)
```

Alternatively,

```
. egen exporters = group(exp)

. egen importers = group(imp)

. quietly tabulate exporters, generate(exp_dum_)

. quietly tabulate importers, generate(imp_dum_)

. regress ln_trade ln_distance contig comlang_off colony comcol exp_dum_* imp_dum_*
if sector=="SER", robust cluster(dist)
```

Taking account of MTR

- How serious is this OV bias empirically?

A model with exporter and importer FEs

R-squared = 0.7681
Root MSE = 1.1333

(Std. Err. adjusted for 2329 clusters in dist)

ln_trade	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
ln_distance	-1.014767	.0469219	-21.63	0.000	-1.10678	-.9227543
contig	.235591	.202185	1.17	0.244	-.1608905	.6320725
comlang_off	.3982351	.0936922	4.25	0.000	.2145062	.5819639
colony	1.173628	.1159908	10.12	0.000	.9461722	1.401084
comcol	-.088625	.2584496	-0.34	0.732	-.5954404	.4181904

An intuitive model

R-squared = 0.5431
Root MSE = 1.5281

(Std. Err. adjusted for 2152 clusters in dist)

ln_trade	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
ln_gdp_exp	.6011672	.0132209	45.47	0.000	.5752401	.6270942
ln_gdp_imp	.6176062	.0142666	43.29	0.000	.5896284	.6455839
ln_distance	-.7385146	.03536	-20.89	0.000	-.8078579	-.6691714
contig	.3989524	.1829276	2.18	0.029	.0402191	.7576858
comlang_off	.8861328	.0993078	8.92	0.000	.6913835	1.080882
colony	1.202965	.1201971	10.01	0.000	.9672503	1.43868
comcol	-.0245067	.2018195	-0.12	0.903	-.4202883	.371275
_cons	-22.03706	.671738	-32.81	0.000	-23.35438	-20.71974

Taking account of MTR

- Distance coefficients
  - The intuitive model = -0.738***
  - The FE model = -1.015***
  - The difference is statistically significant at the 1% level

Augmented theoretical gravity models

- Policy variables often vary only in the exporter or importer dimensions
  - Perfectly collinear with corresponding FEs
- Possible Solutions:
  - Transforming the policy variables (FE approach)
  - Using random effects (RE approach)
    - Ex. Egger (2002), Carrère (2006)
    - Not recommend due to a strong assumption.
  - Baier and Bergstrand (BB) approach
    - See Baier and Bergstrand (2009)

An augmented FE model

```
. gen etcr_both = etcr_exp*etcr_imp
(23562 missing values generated)

. regress ln_trade etcr_both ln_distance contig comlang_off colony comcol
i.exporters i.importers if sector=="SER", robust cluster(dist)
note: comcol omitted because of collinearity

Linear regression                                      Number of obs =     816
                                                       F( 63,   413) =   58.69
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8646
                                                       Root MSE      =  .93726

                               (Std. Err. adjusted for 414 clusters in dist)
------------------------------------------------------------------------------
             |               Robust
    ln_trade |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   etcr_both |    -.30427   .0917592    -3.32   0.001    -.4846434   -.1238967
 ln_distance |  -.8979641   .1073258    -8.37   0.000    -1.108937   -.6869912
      contig |   .3251201   .2648705     1.23   0.220    -.1955424    .8457826
 comlang_off |   .2087727   .1919348     1.09   0.277    -.1685182    .5860636
      colony |   .4613652   .2723341     1.69   0.091    -.0739685     .996699
      comcol |          0  (omitted)
             |
```

Disadvantages of the FE approach

- Losing insights on interested policy impacts
- Dimensionality constraints
  - Dummies for sectors, exporter-sector, and importer-sector
  - A large number of dummies if they are time-variant
  - Sectoral-variation of trade costs
  - A trick: reduce dimensions by estimating separate models for each sector

RE approach

- Advantages:
  - No dimensionality constraints
  - Can include policy variables that collinear with FEs
- Disadvantages
  - A strong assumption: MTRs must be normally distributed
    - Otherwise, RE estimates are not consistent
    - FE estimates are always consistent even if the true model is RE

Baier and Bergstrand (BB) Approach

- Taking account of MTR without using dummies
- (Policy) variables that varies by exporter or importer can be directly included
- Doing a 1st order Taylor series approximation of the MTR terms.

BB (2009) approach

Original form of Theoretical Gravity

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - \log Y^k + (1-\sigma_k)\left[\log \tau_{ij}^k - \log \Pi_i^k - \log P_j^k\right]$$

Baier and Bergstrand transformation

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - \log Y^k + (1-\sigma_k)\left[\log \tau_{ij}^{k\,*}\right]$$

by using the 1st order Taylor series approximation of MTR:

$$\log \tau_{ij}^{k\,*} = \log \tau_{ij}^k - \sum_{j=1}^{N} \theta_j^k \log \tau_{ij}^k - \sum_{i=1}^{N} \theta_i^k \log \tau_{ji}^k + \sum_{i=1}^{N}\sum_{j=1}^{N} \theta_i \theta_j \log \tau_{ij}^k$$

weighted by GDP shares

$$\theta_i^k = \frac{Y_i^k}{Y^k}$$

BB estimation procedures

1. Calculate the weight terms

2. Calculate log τ_{ij}* for EACH trade-cost variable

$$\ln dist_{ij}^{*} = \ln dist_{ij} - \sum_i \theta_i \ln dist_{ij} - \sum_j \theta_j \ln dist_{ij} + \sum_i \sum_j \theta_i \theta_j \ln dist_{ij}$$

$$contig_{ij}^{*} = contig_{ij} - \sum_i \theta_i contig_{ij} - \sum_j \theta_j contig_{ij} + \sum_i \sum_j \theta_i \theta_j contig_{ij}$$

3. Estimate the BB gravity model with OLS

$$\log X_{ij}^k = \log Y_i^k + \log E_j^k - \log Y^k + (1-\sigma_k)\left[\log \tau_{ij}^{k\,*}\right]$$

Points to keep in mind

- Need to apply the Taylor approximation to ALL variables in the trade cost function

$$\log \tau_{ij}^{k} = b_1 \log distance_{ij} + b_2 contig + b_3 comlang\_off f + b_4 colony + b_5 comcol$$

- Possible endogeneity problems when using GDP weights
  - BB (2009) recommend using simple averages rather than the GDP-weighted averages

BB with simple averages in Stata

Find the weight term: $\theta_i^k = \dfrac{1}{N_i^k}$

```
sum exporters importers if sector=="SER"

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   exporters |      7003    101.2138    56.35636          1        218
   importers |      7003    100.4535    56.00826          1        218
```

Calculate log τ_{ij}*

```
. egen temp1 = mean(ln_distance), by(exp sector)
(303 missing values generated)

. egen temp2 = mean(ln_distance), by(imp sector)
(303 missing values generated)

. egen temp3 = sum(ln_distance), by(sector)

. gen ln_distance_star = ln_distance - temp1 - temp2 + (1/218^2)*temp3
(606 missing values generated)
```

BB with simple averages in Stata

A simplified BB model

```
. reg  ln_trade ln_distance_star ln_gdp_exp ln_gdp_imp if sector=="SER",robust cluster(dist)

Linear regression                                      Number of obs =     3884
                                                       F(  3,  2151) =   971.38
                                                       Prob > F      =   0.0000
                                                       R-squared     =   0.5019
                                                       Root MSE      =   1.5947

                                (Std. Err. adjusted for 2152 clusters in dist)
------------------------------------------------------------------------------
             |               Robust
    ln_trade |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ln_distanc~r |  -1.142375    .050619   -22.57   0.000    -1.241642   -1.043108
```

A simplified FE model

```
. regress ln_trade ln_distance i.exporters i.importers if sector=="SER", robust
cluster(dist)

Linear regression                                      Number of obs =     4184
                                                       F(379,  2328) =        .
                                                       Prob > F      =        .
                                                       R-squared     =   0.7483
                                                       Root MSE      =   1.1801

                                (Std. Err. adjusted for 2329 clusters in dist)
------------------------------------------------------------------------------
             |               Robust
    ln_trade |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 ln_distance |  -1.128312   .0462662   -24.39   0.000    -1.219039   -1.037585
             |
```

Often-made mistakes

Things to do and don't

1. Do aiming for theoretical gravity models, Not the intuitive ones

2. Do using "unidirectional" trade flows on LHS, Not the total trade of a country pair.
   - X_{ij} or M_{ij} (NOT T_{ij})

3. Do using nominal terms, Not real terms
   - Price indices are already captured by the MRT terms

4. Do using aggregate GDP, Not GDP per capita

Things to do and don't

5. The "estimated trade cost" is needed, Not just distance.

   ex. $$\log \tau_{ij}^{k} = b_1 \log distance_{ij} + b_2 contig + b_3 comlang\_off + b_4 colony + b_5 comcol$$

6. Always think about dimensions of your dataset
   - How many dimensions of your variables?
   - exporter, importer, sector, time (and interactions among them)

7. Control variables that vary in the same dimensions as the FEs CANNOT be included in FE models.
   - **Do** transform them if FE approach will be used (but it comes with a cost).
   - BB approach can be an alternative.

Things to do and don't

8. If using BB approach, Do apply the Taylor approximation to ALL trade cost variables.

9. Be extremely cautious when interpreting the gravity results with possible endogeneity problems

What next?

- We have so far primarily used OLS as the estimation methodology for a variety of gravity models.

- It is important for applied researchers to know that the gravity literature has undergone a series of development to produce policy research that is increasingly credible and robust.

- The workshop will continue with real applications of the gravity techniques for policy research. It will show examples of using specific empirical techniques with reference to the recent gravity literature which is important for producing a sound basis for drawing policy conclusion.

Reference

- Empirical examples uses the dataset on bilateral trade in services compiled by Francoise et.al. (2009). The original trade data has been re-aggregated and compiled by Ben Shepherd (2012) to include GDP data are from the WDIs. Geographical data are from CEPII. Policy data are from the OECD).
- The completed lists of references are provided in the reading list

Thank you!

Training materials of this session are drawn largely from Ben Shepherd (2012).

Available at http://www.unescap.org/tid/publication/tipub2645.asp