14.581 MIT PhD International Trade
—Lecture 16: Gravity Models (Empirics)—

Dave Donaldson

Spring 2011

Plan for Today's Lecture on Gravity Model Empirics

- **We will begin with some general lessons about the 'fit' of gravity equations in settings where we have reasonable proxies for (some) trade costs.**
- But most gravity equation estimation has been for the purposes of determining the size of barriers to trade (and determinants of these barriers).
- So we will then review various ways in which researchers have attempted to measure the size of barriers to trade, and the determinants of barriers to trade:
  1. Direct measurement.
  2. Using trade flows to infer trade costs (gravity equations).
  3. Using price dispersion and price gaps to infer trade costs
  4. Other work on trade costs.

Goodness of Fit of Gravity Equations

- Lai and Trefler (2002, unpublished) discuss (among other things) the fit of the gravity equation.
- Recall from the previous lecture the notation in Anderson and van Wincoop (2004), but study imports (M) into i from j rather than exports:

$$M_{ij}^k = \frac{E_i^k Y_j^k}{Y^k} \left(\frac{\tau_{ij}^k}{P_i^k \Pi_j^k} \right)^{1-\epsilon^k}$$

  - Where P_i^k and Π_j^k are price indices.

Goodness of Fit of Gravity Equations

$$M_{ij}^k = \frac{E_i^k Y_j^k}{Y^k} \left(\frac{\tau_{ij}^k}{P_i^k \Pi_j^k} \right)^{1-\epsilon^k}$$

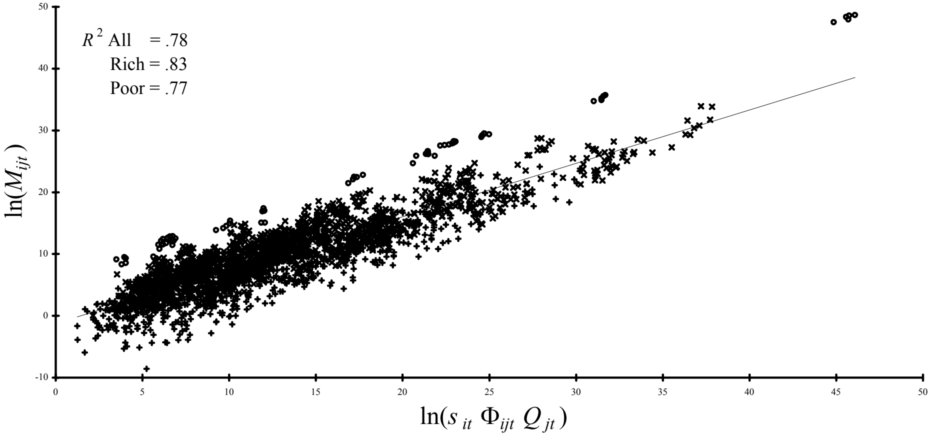
- Lai and Trefler (2002) discuss the fit of this equation, and then divide up the fit into 3 parts (using their notation):
  1. $Q_j^k \equiv Y_j^k$. Fit from this, they argue, is uninteresting due to the "data identity" that $\sum_i M_{ij}^k = Y_j^k$.
  2. $s_i^k \equiv E_i^k$. Fit from this, they argue, is somewhat interesting as it's due to homothetic preferences. But not that interesting.
  3. $\Phi_{ij}^k \equiv \left(\frac{\tau_{ij}^k}{P_i^k \Pi_j^k} \right)^{1-\epsilon^k}$. This, they argue, is the interesting bit of the gravity equation. It includes the partial-equilibrium effect of trade costs τ_{ij}^k, as well as all general equilibrium effects (in P_i^k and Π_j^k).

Lai and Trefler (2002): Other Notes

- Other notes on their estimation procedure:
  - They use 3-digit manufacturing industries (28 industries), every 5 years from 1972-1992, 14 importers (OECD) and 36 exporters. (Big constraint is data on tariffs.)
  - They estimate trade costs τ_{ij}^k as simply equal to tariffs.
  - They estimate one parameter ϵ^k per industry k.
  - They also allow for unrestricted taste-shifters by country (fixed over time).
  - Note that the term Φ_{ij}^k is highly non-linear in parameters.

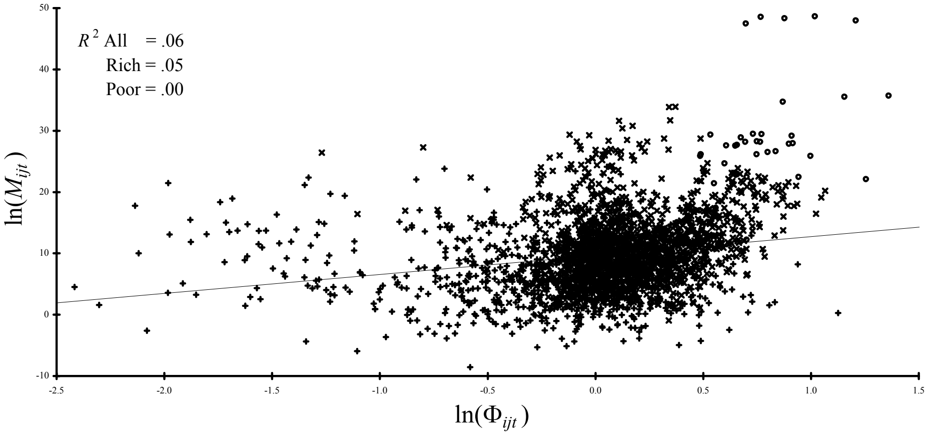
Lai and Trefler (2002): Results

Overall fit, pooled cross-sections



Lai and Trefler (2002): Results

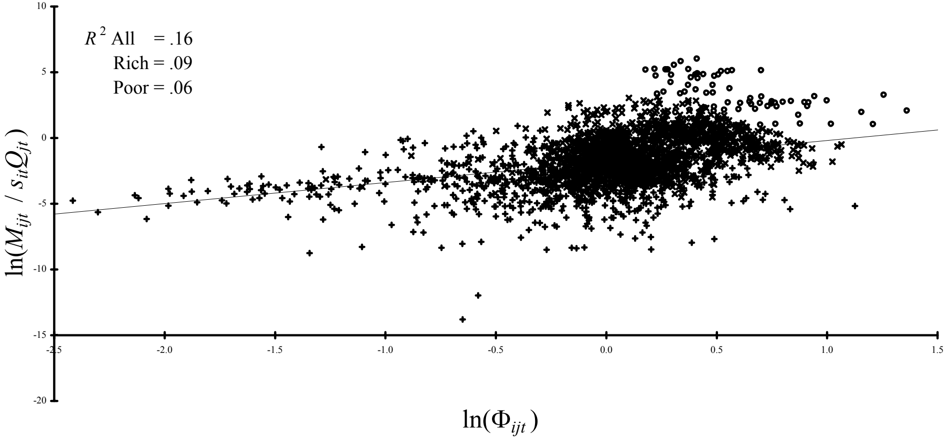
Fit from just Φ_{ijt}^{k}, pooled cross-sections



Courtesy of Daniel Trefler and Huiwen Lai. Used with permission.

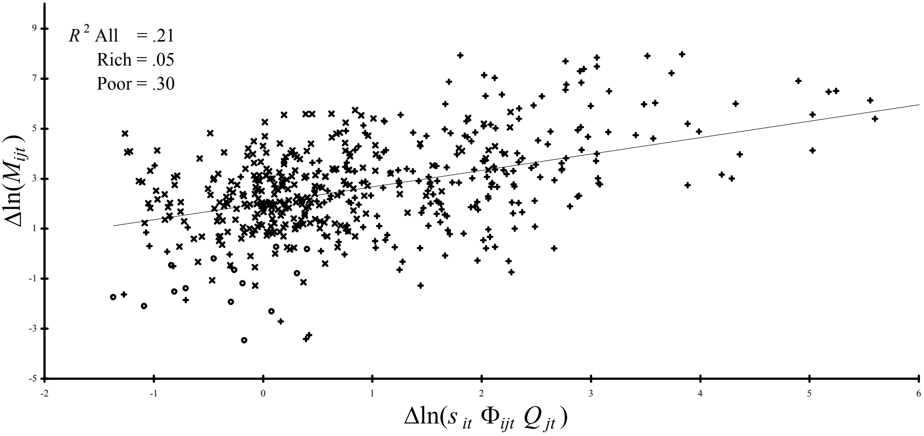
Lai and Trefler (2002): Results

Fit from just Φ_{ijt}^k, but controlling for s_{it}^k and Q_{jt}^k, pooled cross-sections



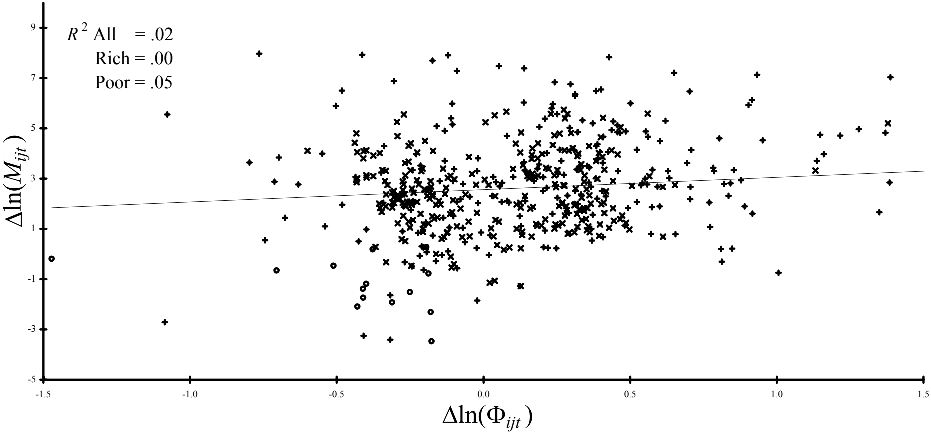
Lai and Trefler (2002): Results

Overall fit, long differences



Lai and Trefler (2002): Results

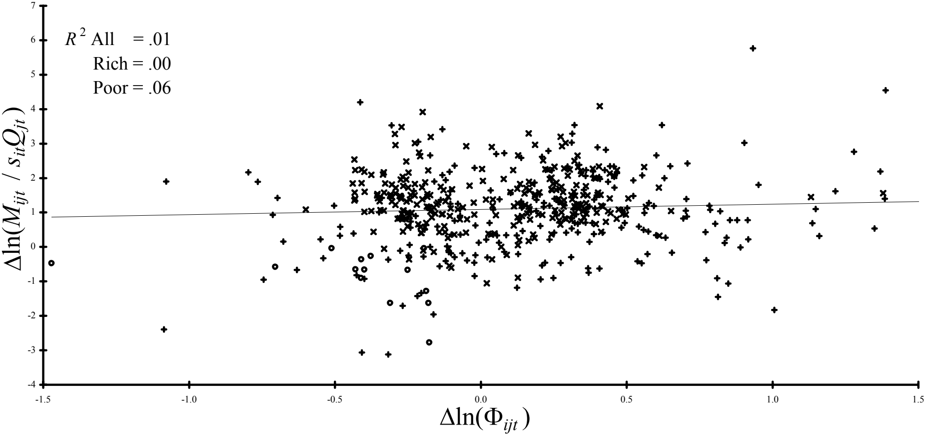
Fit from just Φ^k_{ijt}, long differences



Courtesy of Daniel Trefler and Huiwen Lai. Used with permission.

Lai and Trefler (2002): Results

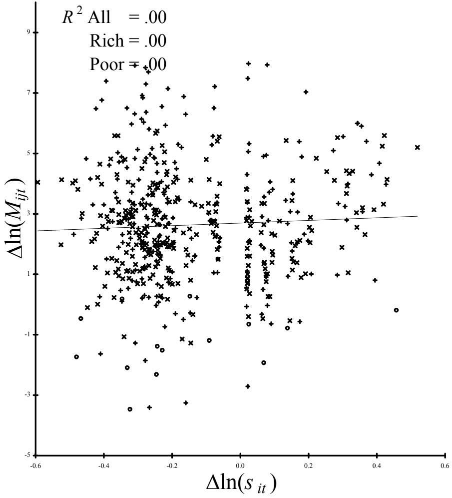
Fit from just Φ_{ijt}^k, but controlling for s_{it}^k and Q_{jt}^k, long differences

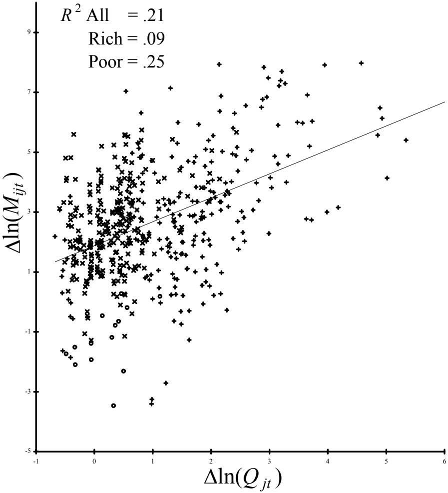


Courtesy of Daniel Trefler and Huiwen Lai. Used with permission.

Lai and Trefler (2002): Results

Exploring whether fit over long differences is driven by s_{it}^{k} (homotheticity) or Q_{jt}^{k} ("data identity")





Plan for Today's Lecture on Gravity Model Empirics

- We will begin with some general lessons about the 'fit' of gravity equations in settings where we have reasonable proxies for (some) trade costs.
- **But most gravity equation estimation has been for the purposes of determining the size of barriers to trade (and determinants of these barriers).**
- So we will then review various ways in which researchers have attempted to measure the size of barriers to trade, and the determinants of barriers to trade:
    1. Direct measurement.
    2. Using trade flows to infer trade costs (gravity equations).
    3. Using price dispersion and price gaps to infer trade costs
    4. Other work on trade costs.

Measuring Trade Costs: What do we mean by 'trade costs'?

- The sum total of all of the costs that impede trade from origin to destination.
- This includes:
  - Tariffs and non-tariff barriers (quotas etc).
  - Transportation costs.
  - Administrative hurdles.
  - Corruption.
  - Contractual frictions.
  - The need to secure trade finance (working capital while goods in transit).
- NB: There is no reason that these 'trade costs' occur only on international trade.

Introduction: Why care about trade costs?

1. They enter many modern models of trade, so empirical implementations of these models need an empirical metric for trade costs.
2. There are clear features of the international trade data that seem hard (but not impossible) to square with a frictionless world.
3. As famously argued by Obstfeld and Rogoff (Brookings, 2000), trade costs may explain 'the six big puzzles of international macro'.
4. Trade costs clearly matter for welfare calculations.
5. Trade costs could be endogenous and driven by the market structure of the trading sector; this would affect the distribution of gains from trade. (A monopolist on transportation could extract all of the gains from trade.)

Are Trade Costs ‘Large’?

- There is considerable debate (still unresolved) about this question.
- Arguments in favor:
    - Trade falls very dramatically with distance (see Figures to follow shortly).
    - Clearly haircuts are not very tradable but a song on iTunes is. Everything else is in between.
    - Contractual frictions of sale at a distance (Avner Grief’s ‘Fundamental Problem of Exchange’, No Trade theorems, etc) seem potentially severe.
    - Commonly heard claim that a fundamental problem in developing countries is their ‘sclerotic’ infrastructure (ie ports, roads, etc). (For a colorful description, see 2005 *Economist* article on traveling with a truck driver around Cameroon.)

Are Trade Costs 'Large'?

- Arguments against:
  - Inter- and intra-national shipping rates aren't that high: in March 2010 (even at relatively high gas prices) a California-Boston refrigerated truck journey cost around $\$5,000$. Fill this with grapes and they will sell at retail for around $\$100,000$.
  - Tariffs are not that big (nowadays).
  - Repeated games and reputations/brand names get around any high stakes contractual issues.
- Surprisingly little hard evidence has been brought to bear on these issues.

Trade Falls with Distance: Leamer (JEL 2007)

From Germany. Visual evidence for the gravity equation

Leamer: A Review of Thomas L Friedman's The World is Flat 111

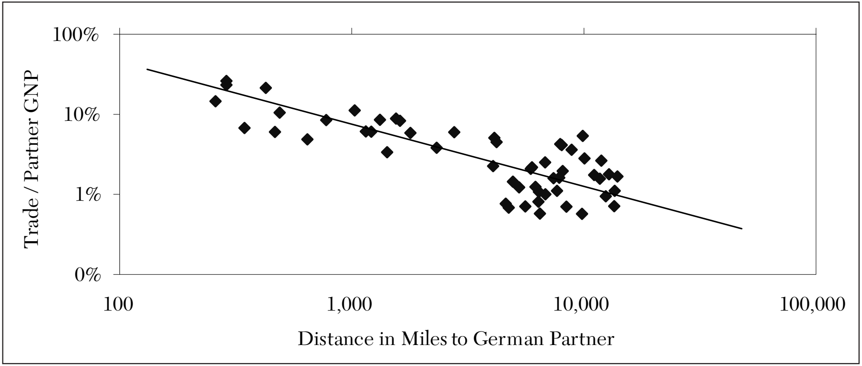


Figure 8. West German Trading Partners, 1985

Trade Falls with Distance: Eaton and Kortum (2002)

OECD manufacturing in 1995

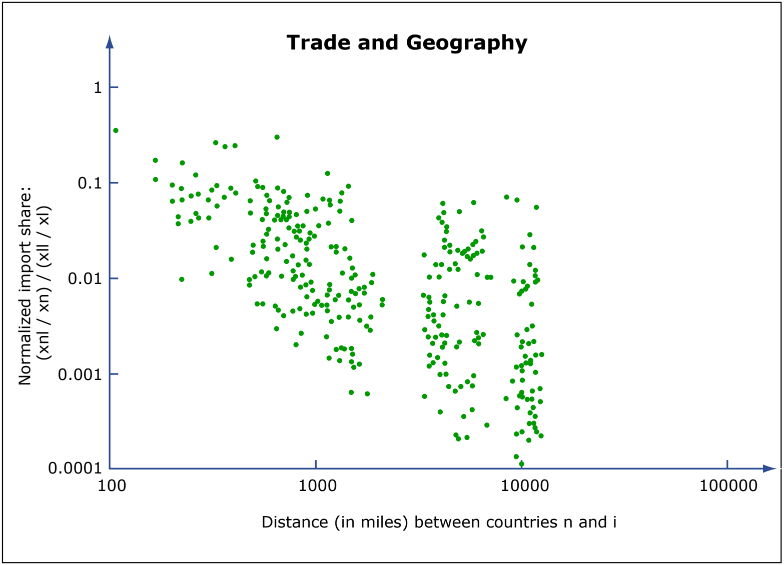
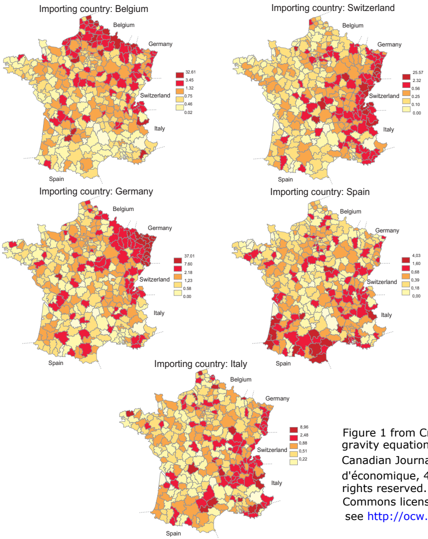


Image by MIT OpenCourseWare.

Trade Falls with Distance: Inside France

Crozet and Koenig (2009): Intensive Margin

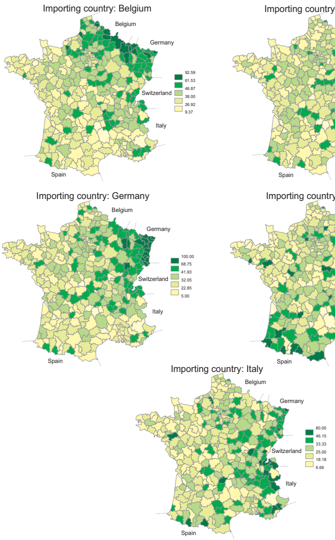
Figure 1: Mean value of individual-firm exports (single-region firms, 1992)



Trade Falls with Distance: Inside France

Crozet and Koenig (2009): Extensive Margin

Figure 2: Percentage of firms which export (single-region firms, 1992)



Trade Falls with Distance: Inside the US

Hilberry and Hummels (EER 2008) using zipcode-to-zipcode CFS data

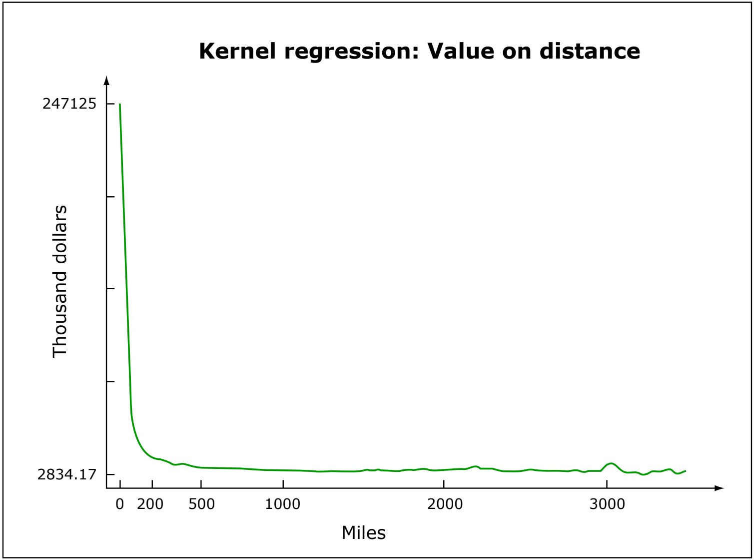


Image by MIT OpenCourseWare.

Plan for Today's Lecture on Gravity Model Empirics

- We will begin with some general lessons about the 'fit' of gravity equations in settings where we have reasonable proxies for (some) trade costs.
- But most gravity equation estimation has been for the purposes of determining the size of barriers to trade (and determinants of these barriers).
- **So we will then review various ways in which researchers have attempted to measure the size of barriers to trade, and the determinants of barriers to trade:**
  1. **Direct measurement.**
  2. Using trade flows to infer trade costs (gravity equations).
  3. Using price dispersion and price gaps to infer trade costs
  4. Other work on trade costs.

Direct Measurement of Trade Costs

- The simplest way to measure TCs is to just go out there and measure them directly.
- Many components of TCs are probably measurable. But many aren't.
- Still, this sort of descriptive evidence is extremely valuable for getting a sense of things.
- Sources of this sort of evidence (there is probably much more):
  - Hummels (JEP, 2007) survey on transportation.
  - Anderson and van Wincoop (JEL, 2004) survey on trade costs.
  - Limao and Venables (WBER 2001) on shipping.
  - Barron and Olken (JPE 2009) on bribes and trucking in Indonesia.
  - Fafchamps (2004 book) on traders and markets in Africa.

Direct Measures: Hummels (2007)

Air shipping prices falling.

Direct Measures: AvW (2004) Survey

- Anderson and van Wincoop (2004) survey trade costs in great detail.
- They begin with descriptive, 'direct' evidence on:
  - Tariffs—but this is surprisingly hard. (It is genuinely scandalous how hard it is to get good data on the state of the world's tariffs.)
  - NTBs—much harder to find data. And then there are theoretical issues such as whether quotas are binding.
  - Transportation costs (mostly now summarized in Hummels (2007)).
  - Wholesale and retail distribution costs (which clearly affect both international and intranational trade).

Direct Measures: AvW (2004)

Tariffs

Simple and Trade-Weighted Tariff Averages - 1999

Country	Simple Average	TW Average	Country	Simple Average	TW Average
Argentina	14.8	11.3	Jamaica	18.8	16.7
Australia	4.5	4.1	Japan	2.4	2.9
Bahamas	0.7	0.8	Korea	9.1	5.9
Bahrain	7.8	—	Mexico	17.5	6.6
Bangladesh	22.7	21.8	Montserrat	18.0	—
Barbados	19.2	20.3	New Zealand	2.4	3.0
Belize	19.7	14.9	Nicaragua	10.5	11.0
Bhutan	15.3	—	Paraguay	13.0	6.1
Bolivia	9.7	9.1	Peru	13.4	12.6
Brazil	15.5	12.3	Philippines	9.7	—
Canada	4.5	1.3	Romania	15.9	8.3
Chile	10.0	10.0	Saudi Arabia	12.2	—
Colombia	12.2	10.7	Singapore	0.0	0.0
Costa Rica	6.5	4.0	Slovenia	9.8	11.4
Czech Republic	5.5	—	South Africa	6.0	4.4
Dominica	18.5	15.8	St. Kitts	18.7	—
Ecuador	13.8	11.1	St. Lucia	18.7	—
European Union	3.4	2.7	St. Vincent	18.3	—
Georgia	10.6	—	Suriname	18.7	—
Grenada	18.9	15.7	Switzerland	0.0	0.0
Guyana	20.7	—	Taiwan	10.1	6.7
Honduras	7.5	7.8	Trinidad	19.1	17.0
Hong Kong	0.0	0.0	Uruguay	4.9	4.5
India	30.1	—	USA	2.9	1.9
Indonesia	11.2	—	Venezuela	12.4	13.0

Notes: The data are from UNCTAD's TRAINS database (Haveman repackaging).
A "—" indicates that trade data for 1999 are unavailable in TRAINS.

Direct Measures: AvW (2004)

NTB 'coverage ratios' (% of product lines that are subject to an NTB).

Non-Tariff Barriers 1999

Country	NTB ratio (narrow)	TW NTB ratio (narrow)	NTB ratio (broad)	TW NTB ratio (broad)
Algeria	.001	.000	.183	.388
Argentina	.260	.441	.718	.756
Australia	.014	.006	.225	.351
Bahrain	.009	–	.045	–
Bhutan	.041	–	.045	–
Bolivia	.014	.049	.179	.206
Brazil	.108	.299	.440	.603
Canada	.151	.039	.307	.198
Chile	.029	.098	.331	.375
Colombia	.049	.144	.544	.627
Czech Republic	.001	–	.177	–
Ecuador	.065	.201	.278	.476
European Union	.008	.041	.095	.106
Guatemala	.000	.000	.348	.393
Hungary	.013	.034	.231	.161
Indonesia	.001	–	.118	–
Lebanon	.000	–	.000	–
Lithuania	.000	.000	.191	.196
Mexico	.002	.000	.580	.533
Morocco	.001	–	.066	–
New Zealand	.000	.004	.391	.479
Oman	.006	.035	.134	.162
Paraguay	.018	.108	.256	.385
Peru	.021	.094	.377	.424
Poland	.001	.050	.133	.235
Romania	.001	.000	.207	.185
Saudi Arabia	.014	–	.156	–
Slovenia	.030	.019	.393	.408
South Africa	.000	.002	.113	.161
Taiwan	.057	.074	.138	.207
Tunisia	.000	.000	.317	.598
Uruguay	.052	.098	.354	.470
USA	.015	.055	.272	.389
Venezuela	.131	.196	.382	.333

Notes: The data are from UNCTAD's TRAINS database (Haveman repackaging). The "narrow" category includes, quantity, price, quality and advance payment NTBs, but does not include threat measures such as antidumping investigations and duties. The "broad" category includes quantity, price, quality, advance payment and threat measures. The ratios are calculated based on six-digit HS categories. A "-" indicates that trade data for 1999 are not available.

Direct Measures: AvW (2004)

MFA: An example of a case/industry where good quota data exists. Deardorff and Stern (1998) converted to tariff equivalents.

Tariff Equivalents of U.S. MFA Quotas, 1991 and 1993 (Percent)

Sector	1991	1993				
	Rent Tar Eq.	Rent Tar Eq.	S Tariff	TW Tariff	Rent + TW Tariff	%US Imports
Textiles:						
Broadwoven fabric mills	8.5	9.5	14.4	13.3	22.8	0.48
Narrow fabric mills	3.4	3.3	6.9	6.7	10.0	0.22
Yarn mills and textile finishing	5.1	3.1	10.0	8.5	11.6	0.06
Thread mills	4.6	2.2	9.5	11.8	14.0	0.01
Floor coverings	2.8	9.3	7.8	5.7	15.0	0.12
Felt and textile goods, n.e.c.	1.0	0.1	4.7	6.2	6.3	0.06
Lace and knit fabric goods	3.8	5.9	13.5	11.8	17.7	0.04
Coated fabrics, not rubberized	2.0	1.0	9.8	6.6	7.6	0.03
Tire cord and fabric	2.3	2.4	5.1	4.4	6.8	0.08
Cordage and twine	3.1	1.2	6.2	3.6	4.8	0.03
Nonwoven fabric	0.1	0.2	10.6	9.5	9.7	0.04
Apparel and fab. textile products:						
Women's hosiery, except socks	5.4	2.3	–	–	–	–
Hosiery, n.e.c.	3.5	2.4	14.9	15.3	17.7	0.04
App'l made from purchased mat'l	16.8	19.9	13.2	12.6	32.5	5.71
Curtains and draperies	5.9	12.1	11.9	12.1	24.2	0.01
House furnishings, n.e.c.	8.3	13.9	9.3	8.2	22.1	0.27
Textile bags	5.9	9.0	6.4	6.6	15.6	0.01
Canvas and related products	6.3	5.2	6.9	6.4	11.6	0.03
Pleating, stitching, ... embroidery	5.2	7.6	8.0	8.1	15.7	0.02
Fabricated textile products, n.e.c.	9.2	0.6	5.2	4.8	5.4	0.37
Luggage	2.6	10.4	12.1	10.8	21.2	0.28
Women's handbags and purses	1.0	3.1	10.5	6.7	9.8	0.44

Notes: "**S**" indicates "simple" and "**TW**" indicates "trade-weighted." Rent equivalents for U.S. imports from Hong Kong were estimated on the basis of average weekly Hong Kong quota prices paid by brokers, using information from International Business and Economic Research Corporation. For countries that do not allocate quota rights in public auctions, export prices were estimated from Hong Kong export prices, with adjustments for differences in labor costs and productivity. Sectors and their corresponding SIC classifications are detailed in USITC (1995) Table D1. Quota tariff equivalents are reproduced from Deardorff and Stern (1998), Table 3.6 (Source USITC 1993,1995). Tariff averages, trade-weighted tariff avarages and U.S. import percentages are calculated using data from the UNCTAD TRAINS dataset. SIC to HS concordances from the U.S. Census Bureau are used.

Direct Measures: AvW (2004)

Domestic distribution costs (measured from I-O tables).

Distribution Margins for Household Consumption and Capital Goods

Select Product Categories	Aus. 95	Bel. 90	Can. 90	Ger. 93	Ita. 92	Jap. 95	Net. 90	UK 90	US 92
Rice	1.239	1.237	1.867	1.423	1.549	1.335	1.434	1.511	1.435
Fresh, frozen beef	1.485	1.626	1.544	1.423	1.605	1.681	1.640	1.390	1.534
Beer	1.185	1.435	1.213	1.423	1.240	1.710	1.373	2.210	1.863
Cigarettes	1.191	1.133	1.505	1.423	1.240	1.398	1.230	1.129	1.582
Ladies' clothing	1.858	1.845	1.826	2.039	1.562	2.295	1.855	2.005	2.159
Refrigerators, freezers	1.236	1.586	1.744	1.826	1.783	1.638	1.661	2.080	1.682
Passenger vehicles	1.585	1.198	1.227	1.374	1.457	1.760	1.247	1.216	1.203
Books	1.882	1.452	1.294	2.039	1.778	1.665	1.680	1.625	1.751
Office, data proc. mach.	1.715	1.072	1.035	1.153	1.603	1.389	1.217*	1.040	1.228
Electronic equip., etc.	1.715	1.080	1.198	1.160	1.576	1.432	1.224*	1.080	1.139
Simple Average (125 categories)	**1.574**	**1.420**	**1.571**	**1.535**	**1.577**	**1.703**	**1.502**	**1.562**	**1.681**

Notes: The table is reproduced from Bradford and Lawrence, "Paying the Price: The Cost of Fragmented International Markets", Institute of International Economics, forthcoming (2003). Margins represent the ratio of purchaser price to producer price. Margins data on capital goods are not available for the Netherlands, so an average of the four European countries' margins is used.

Direct Measures: Djankov, Freund and Pham (ReStat, 2010)

'Doing business' style survey on freight forwarding firms around the world.

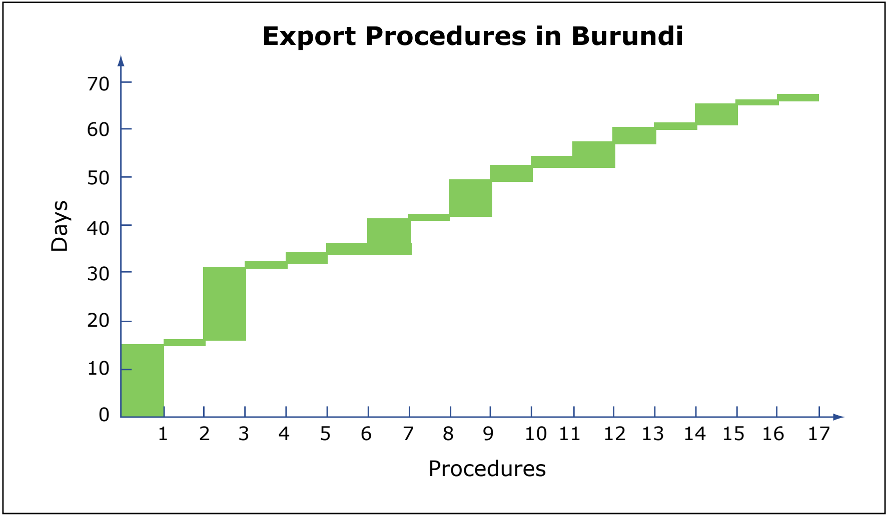


Image by MIT OpenCourseWare.

Direct Measures: Djankov, Freund and Pham (ReStat, 2010)

'Doing business' style survey on freight forwarding firms around the world.

Descriptive Statistics by Geographic Region Required Time for Exports

	Mean	Standard Deviation	Minimum	Maximum	Number of Observation
Africa and Middle East	41.83	20.41	10	116	35
COMESA	50.10	16.89	16	69	10
CEMAC	77.50	54.45	39	116	2
EAC	44.33	14.01	30	58	3
ECOWAS	41.90	16.43	21	71	10
Euro-Med	26.78	10.44	10	49	9
SADC	36.00	12.56	16	60	8
Asia	25.21	11.94	6	44	14
ASEAN 4	22.67	11.98	6	43	6
CER	10.00	2.83	8	12	2
SAFTA	32.83	7.47	24	44	6
Europe	22.29	17.95	5	93	34
CEFTA	22.14	3.24	19	27	7
CIS	46.43	24.67	29	93	7
EFTA	14.33	7.02	7	21	3
FLL FTA	14.33	9.71	6	25	3
European union	13.00	8.35	5	29	14
Western Hemisphere	26.93	10.33	9	43	15
Andean community	28.00	7.12	20	34	4
CACM	33.75	9.88	20	43	4
MERCOSUR	29.50	8.35	22	39	4
NAFTA	13.00	4.58	9	18	3
Total Sample	**30.40**	**19.13**	**5**	**116**	**98**

Note: Seven countries belong to more than one regional agreement

Source: Data on time delays were collected by the doing business team of the World Bank/IFC. They are available at www.doingbusiness.org.

Direct Measures: Barron and Olken (JPE 2009)

Survey of truckers in Aceh, Indonesia.

Summary Statistics

	Both Roads	Meulaboh Road	Banda Aceh Road
Total expenditures during trip (rupiah)	2,901,345 (725,003)	2,932,687 (561,736)	2,863,637 (883,308)
Bribes, extortion, and protection payments	361,323 (182,563)☐	415,263 (180,928)	296,427 (162,896)
Payments at checkpoints	131,876 (106,386)	201,671 (85,203)	47,905 (57,293)
Payments at weigh stations	79,195 (79,405)	61,461 (43,090)	100,531 (104,277)
Convoy fees	131,404 (176,689)	152,131 (147,927)	106,468 (203,875)
Coupons/protection fees	18,848 (57,593)	_	41,524 (79,937)
Fuel	1,553,712 (477,207)	1,434,608 (222,493)	1,697,010 (637,442)
Salary for truck driver and assistant	275,058 (124,685)	325,514 (139,233)	214,353 (65,132)
Loading and unloading of cargo	421,408 (336,904)	471,182 (298,246)	361,523 (370,621)
Food, lodging, etc.	148,872 (70,807)	124,649 (59,067)	178,016 (72,956)
Other	140,971 (194,728)	161,471 (236,202)	116,308 (124,755)
Number of checkpoints	20 (13)	27 (12)	11 (6)
Average payment at checkpoint	6,262 (3,809)	7,769 (1,780)	4,421 (4,722)
Number of trips	282	154	128

Note: Standard deviations are in parentheses. Summary statistics include only those trips for which salary information was available. All figures are in October 2006 rupiah (US$1.00 = Rp. 9,200).

Direct Measures: Barron and Olken (JPE 2009)

Survey of truckers in Aceh, Indonesia.

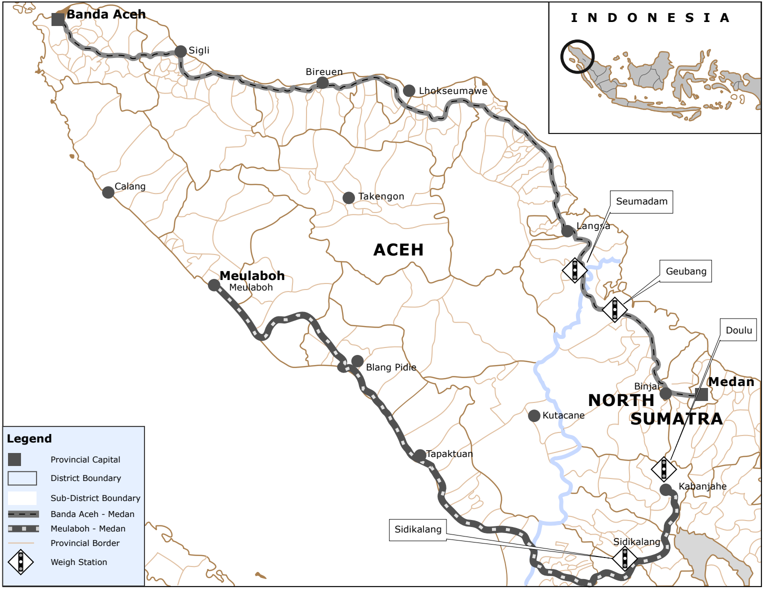


Image by MIT OpenCourseWare.

Direct Measures: Barron and Olken (JPE 2009)

Survey of truckers in Aceh, Indonesia.

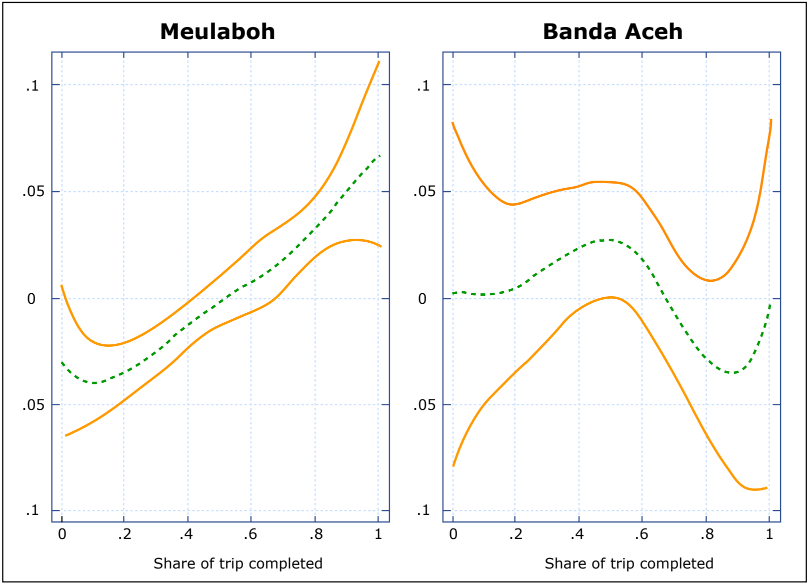


Image by MIT OpenCourseWare.

Plan for Today's Lecture on Gravity Model Empirics

- We will begin with some general lessons about the 'fit' of gravity equations in settings where we have reasonable proxies for (some) trade costs.
- But most gravity equation estimation has been for the purposes of determining the size of barriers to trade (and determinants of these barriers).
- **So we will then review various ways in which researchers have attempted to measure the size of barriers to trade, and the determinants of barriers to trade:**
  1. Direct measurement.
  2. **Using trade flows to infer trade costs (gravity equations).**
  3. Using price dispersion and price gaps to infer trade costs
  4. Other work on trade costs.

Measuring Trade Costs from Trade Flows

- Descriptive statistics can only get us so far. No one ever writes down the full extent of costs of trading and doing business afar.
  - For example, in the realm of transportation-related trade costs: the full transportation-related cost is not just the freight rate (which Hummels (2007) presents evidence on) but also the time cost of goods in transit, etc.
- The most commonly-employed method (by far) for measuring the full extent of trade costs is the gravity equation.
  - This is a particular way of inferring trade costs from trade flows.
  - Implicitly, we are comparing the amount of trade we see in the real world to the amount we'd expect to see in a frictionless world; the 'difference'—under this logic—is trade costs.
  - Gravity models put a lot of structure on the model in order to very transparently back out trade costs.

Estimating τ_{ij}^k from the Gravity Equation: 'Residual Approach'

- One natural approach would be to use the above structure to back out what trade costs τ_{ij}^k must be. Let's call this the 'residual approach'.
- Head and Ries (2001) propose a way to do this:
  - Suppose that intra-national trade is free: $\tau_{ii}^k = 1$. This can be thought of as a normalization of all trade costs (eg assume that AvW (2004)'s 'distributional retail/wholesale costs' apply equally to domestic goods and international goods (after the latter arrive at the port).
  - And suppose that inter-national trade is symmetric: $\tau_{ij}^k = \tau_{ji}^k$.
  - Then we have the 'phi-ness' of trade:

$$\phi_{ij}^k \equiv (\tau_{ij}^k)^{1-\varepsilon^k} = \sqrt{\frac{X_{ij}^k X_{ji}^k}{X_{ii}^k X_{jj}^k}} \tag{1}$$

Estimating τ_{ij}^k from the Gravity Equation: 'Residual Approach'

- There are some drawbacks of this approach:
  - We have to be able to measure internal trade, X_{ii}^k. (You can do this if you observe gross output or final expenditure in each i and k, and re-exporting doesn't get misclassified into the wrong sector.)
  - We have to know ε. (But this is actually a common drawback in most gravity approaches).

Residual Approach to Measuring Trade Costs

Jacks, Meissner and Novy (2010): plots of $\widehat{\tau}_{ijt}$ not $\widehat{\phi}_{ijt}$

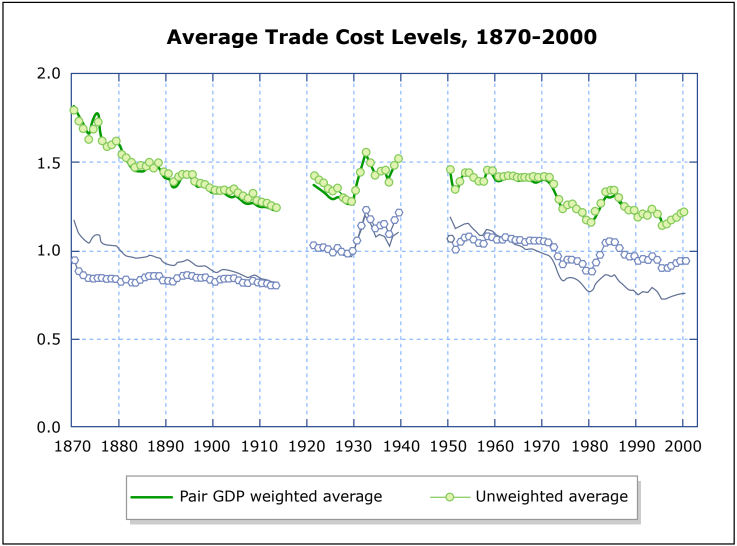


Image by MIT OpenCourseWare.

Estimating τ_{ij}^{k} from the Gravity Equation: 'Determinants Approach'

- A more common approach to measuring τ_{ij}^{k} is to give up on measuring the full τ, and instead parameterize τ as a function of observables.
- The most famous implementation of this is to model TCs as a function of distance (D_{ij}):
    - Assume $\tau_{ij}^{k} = \beta D_{ij}^{\rho}$, to make live easy when estimating in logs.
    - So we give up on measuring the full set of τ_{ij}^{k}'s, and instead estimate just the elasticity of TCs with respect to distance, ρ.
    - How do we know that trade costs fall like this in distance? Eaton and Kortum (2002) use a spline estimator.
- But equally, one could imagine including a whole host of m 'determinants' $z(m)$ of trade costs:
    - $\tau_{ij}^{k} = \prod_m (z(m)_{ij}^{k})^{\rho_m}$.
- This functional form doesn't really have any microfoundations (that I know of).
    - But this functional form certainly makes the estimation of ρ_m in a gravity equation very straightforward.

Anderson and van Wincoop (AER, 2003)

- An important message about how one actually estimates the gravity equation was made by AvW (2003).
- Suppose you are estimating the general gravity model:

$$\ln X_{ij}^k(\boldsymbol{\tau}, \mathbf{E}) = A_i^k(\boldsymbol{\tau}, \mathbf{E}) + B_j^k(\boldsymbol{\tau}, \mathbf{E}) + \varepsilon^k \ln \tau_{ij}^k + \nu_{ij}^k. \quad (2)$$

- Suppose you assume $\tau_{ij}^k = \beta D_{ij}^{\rho^k}$ and try to estimate ρ^k.
- Aside: Note that you can't actually estimate ρ^k here! All you can estimate is $\delta^k \equiv \varepsilon^k \rho^k$. But with outside information on ε^k (in some models it is the CES parameter, which maybe we can estimate from another study) you can back out ε^k.

Anderson and van Wincoop (AER, 2003)

- Suppose you are estimating the general gravity model:

$$\ln X_{ij}^k(\boldsymbol{\tau}, \mathbf{E}) = A_i^k(\boldsymbol{\tau}, \mathbf{E}) + B_j^k(\boldsymbol{\tau}, \mathbf{E}) + \varepsilon^k \ln \tau_{ij}^k + \nu_{ij}^k. \quad (3)$$

  - Note how A_i^k and B_j^k (which are equal to $Y_i^k(\Pi_i^k)^{\varepsilon^k - 1}$ and $E_j^k(P_j^k)^{\varepsilon^k - 1}$ respectively in the AvW (2004) system) depend on τ_{ij}^k too.
  - Even in an endowment economy where Y_i^k and E_j^k are exogenous this is a problem. The problem is the P_j^k and Π_i^k terms.
  - These terms are both price indices, which are very hard to get data on.
  - So a naive regression of X_{ij}^k on E_j^k, Y_i^k and τ_{ij}^k is often performed (this is AvW's 'traditional gravity') instead.
  - AvW (2003) pointed out that this is wrong. The estimate of ρ will be biased by OVB (we've omitted the P_j^k and Π_i^k terms and they are correlated with τ_{ij}^k).

Anderson and van Wincoop (AER, 2003)

- How to solve this problem?
  - AvW (2003) propose non-linear least squares:
    - The functions $(\Pi_i^k)^{1-\varepsilon^k} \equiv \sum_j \left(\frac{\tau^k}{P_j^k}\right)^{1-\varepsilon^k} \frac{E_j^k}{Y^k}$ and $(P_j^k)^{1-\varepsilon^k} \equiv \sum_i \left(\frac{\tau^k}{\Pi_i^k}\right)^{1-\varepsilon^k} \frac{Y_i^k}{Y^k}$ are known.
    - These are non-linear functions of the parameter of interest (ρ), but NLS can handle that.
  - A simpler approach (first in Leamer (1997)) is usually pursued instead though:
    - The terms $A_i^k(\boldsymbol{\tau}, \mathbf{E})$ and $B_j^k(\boldsymbol{\tau}, \mathbf{E})$ can be partialled out using α_i^k and α_j^k fixed effects.
    - Note that (ie avoid what Baldwin calls the 'gold medal mistake') if you're doing this regression on panel data, we need separate fixed effects α_{it}^k and α_{jt}^k in each year t.

Anderson and van Wincoop (AER, 2003)

- This was an important general point about estimating gravity equations
    - And it is a nice example of general equilibrium empirical thinking.
- AvW (2003) applied their method to revisit McCallum (AER, 1995)'s famous argument that there was a huge 'border' effect within North America:
    - This is an additional premium on crossing the border, controlling for distance.
    - Ontario appears to want to trade far more with Alberta (miles away) than New York (close, but over a border).
- The problem is that, as AvW (2003) showed, McCallum (1995) didn't control for the endogenous terms $A_i^k(\boldsymbol{\tau}, \mathbf{E})$ and $B_j^k(\boldsymbol{\tau}, \mathbf{E})$ and this biased his results.

Anderson and van Wincoop (AER, 2003): Results

Re-running McCallum (1995)'s specification

TABLE 1—MCCALLUM REGRESSIONS

	McCallum regressions			Unitary income elasticities		
Data	(i) CA–CA CA–US	(ii) US–US CA–US	(iii) US–US CA–CA CA–US	(iv) CA–CA CA–US	(v) US–US CA–US	(vi) US–US CA–CA CA–US
Independent variable						
ln y_i	1.22	1.13	1.13	1	1	1
	(0.04)	(0.03)	(0.03)			
ln y_j	0.98	0.98	0.97	1	1	1
	(0.03)	(0.02)	(0.02)			
ln d_{ij}	−1.35	−1.08	−1.11	−1.35	−1.09	−1.12
	(0.07)	(0.04)	(0.04)	(0.07)	(0.04)	(0.03)
Dummy–Canada	2.80		2.75	2.63		2.66
	(0.12)		(0.12)	(0.11)		(0.12)
Dummy–U.S.		0.41	0.40		0.49	0.48
		(0.05)	(0.05)		(0.06)	(0.06)
Border–Canada	16.4		15.7	13.8		14.2
	(2.0)		(1.9)	(1.6)		(1.6)
Border–U.S.		1.50	1.49		1.63	1.62
		(0.08)	(0.08)		(0.09)	(0.09)
$\bar{R}^2$	0.76	0.85	0.85	0.53	0.47	0.55
Remoteness variables added						
Border–Canada	16.3		15.6	14.7		15.0
	(2.0)		(1.9)	(1.7)		(1.8)
Border–U.S.		1.38	1.38		1.42	1.42
		(0.07)	(0.07)		(0.08)	(0.08)
$\bar{R}^2$	0.77	0.86	0.86	0.55	0.50	0.57

Notes: The table reports the results of estimating a McCallum gravity equation for the year 1993 for 30 U.S. states and 10 Canadian provinces. In all regressions the dependent variable is the log of exports from region i to region j. The independent variables are defined as follows: y_i and y_j are gross domestic production in regions i and j; d_{ij} is the distance between regions i and j; *Dummy–Canada* and *Dummy–U.S.* are dummy variables that are one when both regions are located in respectively Canada and the United States, and zero otherwise. The first three columns report results based on nonunitary income elasticities (as in the original McCallum regressions), while the last three columns assume unitary income elasticities. Results are reported for three different sets of data: (i) state–province and interprovincial trade, (ii) state–province and interstate trade, (iii) state–province, interprovincial, and interstate trade. The border coefficients *Border–U.S.* and *Border–Canada* are the exponentials of the coefficients on the respective dummy variables. The final three rows report the border coefficients and $\bar{R}^2$ when the remoteness indices (3) are added. Robust standard errors are in parentheses.

Anderson and van Wincoop (AER, 2003): Results

Using theory-consistent (NLS) specification

TABLE 2—ESTIMATION RESULTS

		Two-country model	Multicountry model
Parameters	$(1-\sigma)\rho$	−0.79	−0.82
		(0.03)	(0.03)
	$(1-\sigma)\ln b_{US,CA}$	−1.65	−1.59
		(0.08)	(0.08)
	$(1-\sigma)\ln b_{US,ROW}$		−1.68
			(0.07)
	$(1-\sigma)\ln b_{CA,ROW}$		−2.31
			(0.08)
	$(1-\sigma)\ln b_{ROW,ROW}$		−1.66
			(0.06)
Average error terms:	US–US	0.06	0.06
	CA–CA	−0.17	−0.02
	US–CA	−0.05	−0.04

Notes: The table reports parameter estimates from the two-country model and the multicountry model. Robust standard errors are in parentheses. The table also reports average error terms for interstate, interprovincial, and state–province trade.

Other elements of Trade Costs

- Many determinants of TCs have been investigated in the literature.
- AvW (2004) summarize these:
  - Tariffs, NTBs, etc.
  - Transportation costs (directly measured). Roads, ports. (Feyrer (2009) on Suez Canal had this feature).
  - Currency policies.
  - Being a member of the WTO.
  - Language barriers, colonial ties.
  - Information barriers. (Rauch and Trindade (2002).)
  - Contracting costs and insecurity (Evans (2001), Anderson and Marcoulier (2002)).
  - US CIA-sponsored coups. (Easterly, Nunn and Sayananth (2010).)
- Aggregating these trade costs together into one representative number is not trivial.
  - Anderson and Neary (2005) have outlined how to solve this problem (conditional on a given theory of trade).

AvW (2004): Summary of Gravity Results

Tariff Equivalent of Trade Costs

	Method	Data	Reported by authors	($\sigma = 5$)	($\sigma = 8$)	($\sigma = 10$)
All Trade Barriers						
Head and Ries (2001) U.S.-Canada, 1990-1995	new	disaggr.	48 ($\sigma = 7.9$)	97	47	35
Anderson and van Wincoop (2003) U.S.-Canada, 1993	new	aggr.		91	46	35
Eaton and Kortum (2002) 19 OECD countries, 1990 750-1500 miles apart	new	aggr.	48-63 ($\sigma = 9.28$)	123-174	58-78	43-57
National Border Barriers						
Wei (1996) 19 OECD countries, 1982-1994	trad.	aggr.	5 ($\sigma = 20$)	26-76	14-38	11-29
Evans (2003a) 8 OECD countries, 1990	trad.	disaggr.	45 ($\sigma = 5$)	45	30	23
Anderson and van Wincoop (2003) U.S.-Canada, 1993	new	aggr.	48 ($\sigma = 5$)	48	26	19
Eaton and Kortum (2002) 19 OECD countries, 1990	new	aggr.	32-45 ($\sigma = 9.28$)	77-116	39-55	29-41
Language Barrier						
Eaton and Kortum (2002) 19 OECD countries, 1990	new	aggr.	6 ($\sigma = 9.28$)	12	7	5
Hummels (1999) 160 countries, 1994	new	disaggr.	11 ($\sigma = 6.3$)	12	8	6
Currency Barrier						
Rose and van Wincoop (2001) 143 countries, 1980 and 1990	new	aggr.	26 ($\sigma = 5$)	26	14	11

A Potential Concern About Identification

- The above methodology identified tau (or its determinants) only by assuming trade separability. This seems potentially worrying.
- In particular, there is a set of taste or technology shocks that can rationalize any trade cost vector you want.
  - Eg if we allowed each country i to have its own taste for varieties of k that come from country j (this would be a 'demand shock' shifter in the utility function for i, a_{ij}^k) then this would mean everywhere we see τ_{ij}^k above should really be
  - In general a_{ij}^k might just be noise with respect to determining τ_{ij}^k. But if a_{ij}^k is spatially correlated, as τ_{ij}^k is, then we're in trouble.

A Potential Concern About Identification

- To take an example from the Crozet and Koenigs (2009) maps, do Alsaciens trade more with Germany (relative to how the rest of France trades with Germany) because:
  - They have low trade costs (proximity) for getting to Germany?
  - They have tastes for similar goods?
  - There is no barrier to factor mobility here. German barbers might even cut hair in France.
  - Integrated supply chains choose to locate near each other.
    - Ellison, Glaeser and Kerr (AER, 2009) look at this 'co-agglomeration' in the US.
    - Hummels and Hilberry (EER, 2008) look at this on US trade data by checking whether imports of a zipcode's goos are correlated with the upstream input demands of that zipcode's industry-mix.
    - Rossi-Hansberg (AER, 2005) models this on a spatial continuum (a line).
    - Yi (AER, 2010) argues that this explains much of the 'border effect' that remains even in AvW (2003).

Hilberry and Hummels (EER 2008) using zipcode-to-zipcode US data

Is it really plausible that trade costs fall this fast with distance?

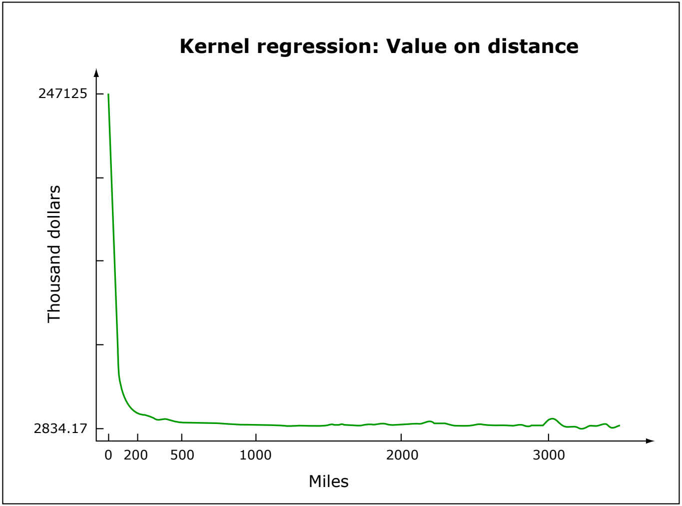
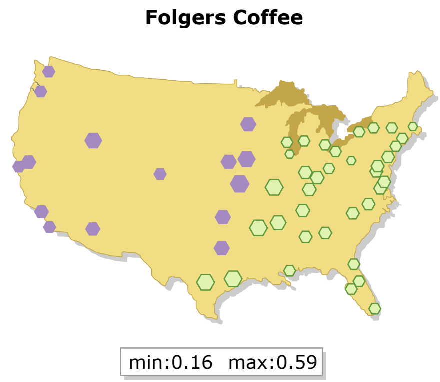


Image by MIT OpenCourseWare.

Bronnenberg, Dube (JPE 2009): Endogenous Tastes?



The joint geographic distribution of share levels and early entry across U.S. markets in ground coffee. The areas of the circles are proportional to share levels. Shaded circles indicate that a brand locally moved first.

Image by MIT OpenCourseWare.

Bronnenberg, Dube (JPE 2009): Endogenous Tastes?

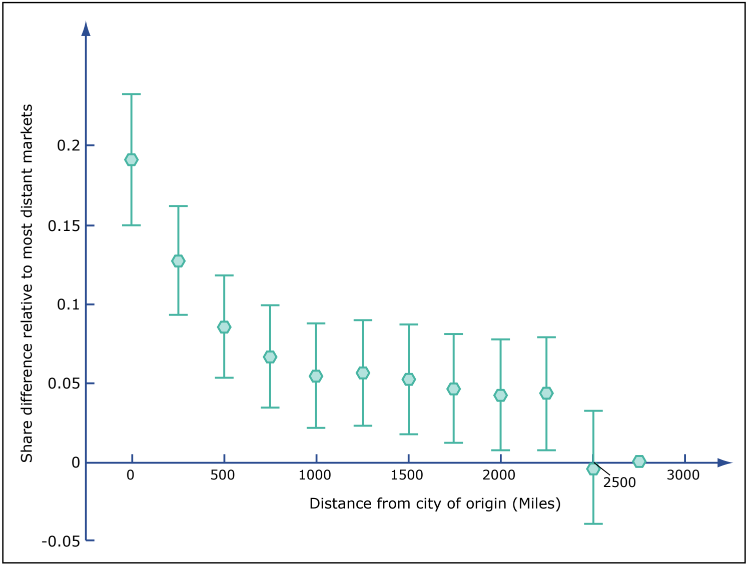


Image by MIT OpenCourseWare.

Puzzling Findings from Gravity Equations

- Trade costs seem very large.
- The decay with respect to distance seems particularly dramatic.
- The distance coefficient has not been dying.
- One sees a distance and a 'border' effect on eBay too:
  - Hortascu, Martinez-Jerez and Douglas (AEJ 2009).
  - Blum and Goldfarb (JIE, 2006) on digital products. But only for 'taste-dependent digital goods': music, games, pornography.

Disidier and Head (ReStat, 2008)

The exaggerated death of distance?

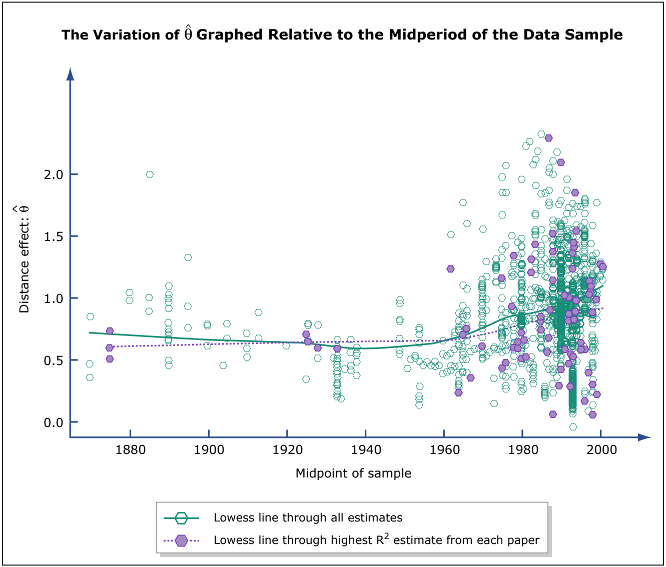


Image by MIT OpenCourseWare.

Plan for Today's Lecture on Gravity Model Empirics

- We will begin with some general lessons about the 'fit' of gravity equations in settings where we have reasonable proxies for (some) trade costs.
- But most gravity equation estimation has been for the purposes of determining the size of barriers to trade (and determinants of these barriers).
- **So we will then review various ways in which researchers have attempted to measure the size of barriers to trade, and the determinants of barriers to trade:**
  1. Direct measurement.
  2. Using trade flows to infer trade costs (gravity equations).
  3. **Using price dispersion and price gaps to infer trade costs.**
  4. Other work on trade costs.

Price Gap Approaches

- This method for estimating trade costs has received far less work among trade economists.
- The core idea is that if there is free arbitrage (assumed in most trade models anyway) then the price for any identical good k at any two points i and j in space must reflect a no-arbitrage condition:
  - $|\ln p_i^k - \ln p_j^k| \leq \tau_{ij}^k$.
  - This holds with equality if there is some good being traded from i to j: ie if $X_{ij}^k > 0$.

Price Gap Approaches

- There are 2 big challenges in using this method:
  - We clearly need to be careful that good k is the exact same good when it is for sale in i and j. (This is harder than just ensuring that it's the same barcode etc. An identical barcode for sale at Whole Foods comes with additional bundled services that might not be available at another sale location.)
  - Conditional on working with very finely-defined goods, it is hard to know whether $X_{ij}^k > 0$ holds. If we're not sure about this, then there are three options:
    - Work with a good that is differentiated by region of origin. Donaldson (2010) did this with 8 types of salt in India.
    - Build a model of supply and demand to tell you when i and j are trading k. (One could argue that if you do this you might as well just use all the information in your model's predicted trade flows, ie pursue the gravity approach.)
    - Or, work with the weak inequality $|\ln p_i^k - \ln p_j^k| \leq \tau_{ij}^k$ in all its generality. This is what the 'market integration' literature (very commonly seen in Economic History and Agricultural Economics) has grappled with.

Plan for Today's Lecture on Gravity Model Empirics

- We will begin with some general lessons about the 'fit' of gravity equations in settings where we have reasonable proxies for (some) trade costs.
- But most gravity equation estimation has been for the purposes of determining the size of barriers to trade (and determinants of these barriers).
- **So we will then review various ways in which researchers have attempted to measure the size of barriers to trade, and the determinants of barriers to trade:**
  1. Direct measurement.
  2. Using trade flows to infer trade costs (gravity equations).
  3. Using price dispersion and price gaps to infer trade costs.
  4. **Other work on trade costs.**

Other Work on Trade Costs

- Micro-founded models of iformation-based, network-based, or contractual friction-based models of trade costs.
  - Greif, Rauch, reputation models of buyers and sellers, favor exchange on networks (Jackson).
- Fixed costs of penetrating a foreign market (our focus has been on variable trade costs):
  - Tybout and Roberts (AER 1998 and Ecta 2008) have made significant progress here.
  - Implications of fixed costs for interpreting gravity equations. (Recall how HMR (2007) and Chaney (2008) point out that coefficient on distance in a gravity regression may be capturing both the variable and fixed costs of trading if both of these costs rise with distnace.)