JIE Summer School

# Lecture 3A: Some Observational and Theoretical Foundations

Jonathan Eaton
Pennsylvania State University

15 June 2023

# I. Observational Foundations

# The Gravity Equation in Trade

- The observation that exports to country $n$ from country $i$, $X_{ni}$ is well described by the equation:

$$X_{ni} = \frac{X_i \cdot X_n}{D_{ni}}$$

  where $X_j$ is some measure of the "mass" of country $j$ and $D_{ni}$ captures "bilateral resistance" between them
- In standard applications the mass of country $j$ is captured by its GDP and $D_{ni}$ by the distance between $n$ and $i$.
- Origins: Isard (1954), Tinbergen (1962),...
- Theoretical foundations: Anderson (1979), Eaton and Kortum (2002), Anderson and Van Wincoop (2003), Melitz (2003)-Chaney (2008),...

# Gravity and Merchandise Trade

From Eaton and Cecília Fieler (2022) "The Margins of Trade":

- UN COMTRADE bilateral trade data for 100 Countries in HS6 products in 2007
  - US$ values $X_{nik}$ of imports by $n$ from $i$ of product $k$
  - units (weight, usually, or counts) $Q_{nik}$ of imports by $n$ from $i$ of product $k$
  - allowing us to infer unit values (prices) $p_{nik} = X_{nik} / Q_{nik}$
- World Bank: GDP and GDP per capita
- CEPII: distance and other bilateral indicators

# The Basic Regression

| dependent variable → | value | (s.e.) |
|---|---|---|
| exporter GDP | 1.356 | (0.054) |
| importer GDP | 1.110 | (0.034) |
| distance | -1.190 | (0.083) |
| | | |
| R-squared | 0.670 | |
| number of observations | 9,479 | |

All variables are in logs.

# Dissecting Gravity

- A huge amount of work has been done on the econometrics of the gravity equation and how it relates to theory
- For today, let's accept that the gravity equation is a robust relationship connecting aggregate trade, GDP, and distance
- So if we break down aggregate trade and total GDP into various pieces, we can ask how the individual pieces contribute to gravity.

# The Margins of GDP

- Define $y_i$ as per capita income and $L_i$ as population
- so that

$$\log GDP_i = \log y_i + \log L_i$$

# The Margins of Trade

Expanding on Hummels and Klenow (2005):

- **Extensive margin**

$$E_{ni} = \frac{\text{number of products exported to } n \text{ from } i}{\text{total number of products in data}}$$

- **Price** is an impoter-exporter fixed effect $\log P_{ni}$ in the regression:

$$\log p_{nik} = \log P_{ni} + \delta_k + \epsilon_{nik}$$

  where $p_{nik}$ is the unit value of country $n$'s imports from country $i$ in product $k$ and $\delta_k$ are product fixed effects

- **Quantity**

$$\log Q_{ni} = \log X_{ni} - \log E_{ni} - \log P_{ni}$$

  $X_{ni}$ is the value of the trade flow from $i$ to $n$

# Gravity on the Margins

| dependent variable → | value | extensive margin | quantity | price |
|---|---|---|---|---|
| exporter GDP | 1.36 | 0.88 | 0.45 | 0.03 |
| importer GDP | 1.11 | 0.40 | 0.66 | 0.05 |
| distance | -1.19 | -0.72 | -0.51 | 0.03 |
| exporter GDP per capita | 1.35 | 0.92 | 0.33 | 0.10 |
| exporter population | 1.36 | 0.85 | 0.55 | -0.03 |
| importer GDP per capita | 1.09 | 0.46 | 0.51 | 0.13 |
| importer population | 1.13 | 0.35 | 0.80 | -0.02 |
| distance | -1.20 | -0.68 | -0.62 | 0.10 |
| number of observations | 9,479 | 9,479 | 9,479 | 9,479 |

All variables are in logs. Standard errors in EF Appendix.

# Takeaways

- The GDP per capita and population breakdown doesn't matter for for total trade. (Elasticities on each are similar to each other and about the same as for total GDP.)
- But higher GDP per capita is associated with a higher price margin in exporting (elasticity 0.10) and importing (elasticity 0.13)
- Could these effects be the result of product selection: Countries sell higher priced products to richer countries and countries buy higher priced products from richer countries?

# Takeaways

- The GDP per capita and population breakdown doesn't matter for for total trade. (Elasticities on each are similar to each other and about the same as for total GDP.)
- But higher GDP per capita is associated with a higher price margin in exporting (elasticity 0.10) and importing (elasticity 0.13)
- Could these effects be the result of product selection: Countries sell higher priced products to richer countries and countries buy higher priced products from richer countries?
- No.

# Takeaways

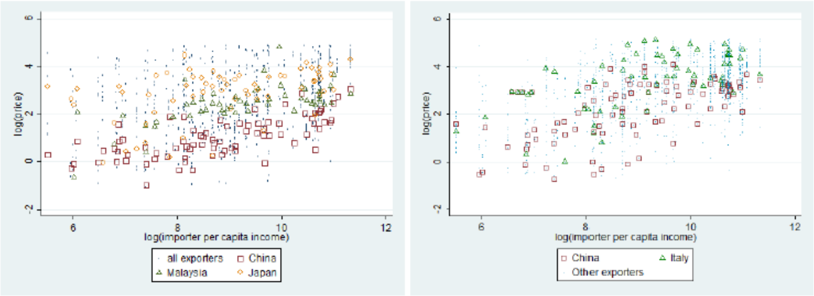
- The GDP per capita and population breakdown doesn't matter for for total trade. (Elasticities on each are similar to each other and about the same as for total GDP.)
- But higher GDP per capita is associated with a higher price margin in exporting (elasticity 0.10) and importing (elasticity 0.13)
- Could these effects be the result of product selection: Countries sell higher priced products to richer countries and countries buy higher priced products from richer countries?
- No.
- The elasticities are the same or higher at the HS6 product level:

# Price Regressions

**Dependent variable is the price for each importer, exporter, and product.**

| **independent variable** ↓ | pooled by exporter-product (1) | pooled by importer-product (2) | pooled by product (3) | income interaction (4) | Rauch (1999) differentiated products[b] (5) | manufacturing only (6) |
|---|---|---|---|---|---|---|
| exporter GDP per capita | 0.171 | | 0.174 | 0.175 | 0.180 | 0.186 |
| | (0.018) | | (0.018) | (0.018) | (0.019) | (0.019) |
| exporter population | -0.043 | | -0.042 | -0.042 | -0.053 | -0.053 |
| | (0.023) | | (0.022) | (0.022) | (0.025) | (0.025) |
| importer GDP per capita | | 0.116 | 0.121 | 0.127 | 0.136 | 0.125 |
| | | (0.013) | (0.013) | (0.012) | (0.013) | (0.013) |
| importer population | | -0.016 | -0.010 | -0.011 | -0.0028 | -0.0036 |
| | | (0.013) | (0.012) | (0.012) | (0.014) | (0.013) |
| distance | 0.104 | 0.108 | 0.085 | 0.080 | 0.080 | 0.076 |
| | (0.016) | (0.012) | (0.015) | (0.016) | (0.016) | (0.017) |
| absolute difference in GDP per capita[a] | | | | 0.020 | | |
| | | | | (0.013) | | |
| product-exporter fixed effect | no | yes | no | no | no | no |
| product-importer fixed effect | yes | no | no | no | no | no |
| product fixed effect | no | no | yes | yes | yes | yes |
| R-squared | 0.825 | 0.836 | 0.788 | 0.788 | 0.776 | 0.788 |
| number of observations | 4,552,967 | 4,552,967 | 4,552,967 | 4,552,967 | 3,165,101 | 2,554,996 |

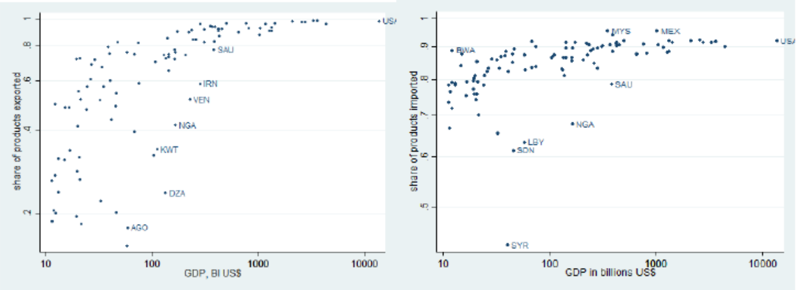
# Two Examples




(a) HS871493 (b) HS620990

bicycle hubs and spokes (HS871493) baby garments (HS620990)

# Extensive Margins: Products per Country




(a) EM of exporting (b) EM of importing

Note different scales in y-axes

# Extensive Margins: Countries (out of 100) per Product

- Exporters per HS6 product
  - mean: 65
  - 10th percentile: 35
  - 90th percentile: 91
- Importers per HS6 product
  - mean: 84
  - 10th percentile: 46
  - 90th percentile: 100

# Gravity, Market Share, and Market Size

- Return to the basic gravity equation

$$X_{ni} = \frac{X_i \cdot X_n}{D_{ni}}$$

- Posit that $X_n$ is $n$'s purchases from all countries, including $n$ itself, so that:

$$X_n = \sum_i X_{ni}$$

- Define $\pi_{ni} = X_{ni}/X_n$ as $i$'s share of sales in market $n$
- and decompose $n$'s imports from $i$ as:

$$X_{ni} = \pi_{ni} \cdot X_n$$

the product of market share and market size

# Sellers, Buyers, and Relationships

From Eaton, Sam Kortum, and Francis Kramarz (2022) "Firm-to-Firm Trade"

- French customs data on the sales of French firms to individual buyers in 24 other EU destinations in 2005, giving us, for each destination $n$:

| | |
|---|---|
| number of French sellers | $N_{nF}$ |
| number of local buyers | $F_{nF}$ |
| buyers/seller | $\bar{b}_{nF}$ |
| sellers/buyer | $\bar{s}_{nF}$ |
| number of relationships | $R_{nF}$ |
| sales/relationship | $\bar{x}_{nF}$ |

- Some identities:

$$R_{nF} = N_{nF}\bar{b}_{nF} = F_{nF}\bar{s}_{nF}$$

$$X_{nF} = R_{nF}\bar{x}_{nF}$$

# Some Regressions

Table: **French Firm Entry into EU Destinations**

| | $\ln R_{nF}$ | $\ln \bar{x}_{nF}$ | $\ln N_{nF}$ | $\ln \bar{b}_{nF}$ | $\ln F_{nF}$ | $\ln \bar{s}_{nF}$ |
|---|---|---|---|---|---|---|
| constant | -2.80 | 2.80 | -1.39 | -1.41 | -4.38 | 1.58 |
| | (0.99) | (0.99) | (0.59) | (0.55) | (0.87) | (1.24) |
| market size | 0.81 | 0.19 | 0.47 | 0.34 | 0.83 | -0.02 |
| | (0.06) | (0.06) | (0.04) | (0.03) | (0.05) | (0.01) |
| French market share | 1.02 | -0.02 | 0.64 | 0.38 | 0.85 | 0.17 |
| | (0.19) | (0.19) | (0.11) | (0.11) | (0.17) | (0.05) |
| Number of Observations | 24 | 24 | 24 | 24 | 24 | 24 |
| $R^2$ | 0.92 | 0.33 | 0.91 | 0.86 | 0.93 | 0.40 |

# Takeaways

- Relationships fully account for French market share (elasticity is 1.02)
- Relationships account for a 0.81 share of market size, with sales per relationship accounting for the rest: Larger markets have larger relationships!
- In either case, a little more than half the increase in relationships is accounted for by more French sellers, the rest by more buyers per seller
- In markets where the French market share is larger, a buyer has more French sellers.

# II. Theoretical Foundations

# Some Primitives

- Concepts behind a vast number of papers in international trade, spacial economics, growth,...
- Kortum, EK (various), Melitz, Buera and Oberfield,.....
- including two papers providing a theoretical explanation for the observations above, which we'll turn to next.
- The goal here is to show the deep connections among the distributions that show up repeatedly in this literature: the Pareto, Poisson, and Fréchet (with the binomial in between)

# An Idea

- An idea for producing a good using inputs
- Efficiency: Output $q$ per unit of inputs
- Pareto distribution of $q$:

$$\Pr\left[Q \leq q\right] = \begin{array}{cc} 1 - \left(\frac{q}{\underline{q}}\right)^{-\theta} & q \geq \underline{q} \\ 0 & q \leq \underline{q} \end{array}$$

with shape parameter $\theta > 0$ and location parameter $\underline{q} > 0$

# Some Properties of the Pareto Distribution I

- Often convenient: the complementary or tail distribution:

$$\Pr\left[Q \geq q\right] = \begin{cases} \left(\frac{q}{\underline{q}}\right)^{-\theta} & q \geq \underline{q} \\ 1 & q \leq \underline{q} \end{cases}$$

- The upper tail is Pareto with shape parameter $\theta$ wherever it's truncated from below (fractality)
- The Pareto distribution is easy to integrate into economic models, and describes some types of data very well.
- For low $\theta$, it has a "fat tail".

# Some Properties of the Pareto Distribution II

- Mean:

$$E[q] = \underline{q}^{\theta}\theta \int_{\underline{q}}^{\infty} q^{-\theta} dq = \frac{\theta}{\theta - 1}\underline{q}$$

  defined only for $\theta > 1$

- Median:

$$.5 = \left(\frac{q_{med}}{\underline{q}}\right)^{-\theta}$$

  so that:

$$q_{med} = 2^{1/\theta}\underline{q}$$

- Both the mean and the median, when it exists, approach the lower bound $\underline{q}$ as $\theta \to \infty$.

# From Efficiency to Unit Costs

- If a bundle of inputs costs $w$ then the unit cost associated with an idea with quality $q$ is

$$c = w/q$$

and the distribution of the associated unit cost is:

$$G(c) = \Pr[C \le c] = \Pr\left[Q \ge \frac{w}{c}\right] = \begin{cases} \left(\frac{c}{\bar{c}}\right)^{-\theta} & c \le \bar{c} \\ 1 & c \ge \bar{c} \end{cases}$$

where $\bar{c} = w/\underline{q}$.

# Putting in Space

- Say there are $N$ locations labelled $i, n = 1, ..., N$ each with a wage $w_i$ separated by iceberg trade costs $d_{ni}$
- An idea with efficiency $z$ in location $i$ can deliver to $n$ at unit cost:

$$c = \frac{w_i d_{ni}}{z}$$

# The Accumulation of Ideas

- Say that $N_i$ ideas have arrived at location $i$, each with quality drawn independently from the Pareto distribution above
- Define

$$p_q = \left( \frac{q}{\underline{q}} \right)^{-\theta},$$

  the probability that an idea is better than $q$
- The number of ideas with quality at least $q \geq \underline{q}$ is $N_{i,q}$, which is distributed binomially:

$$\Pr\left[N_{i,q} = n\right] = \binom{N_i}{n} p_q^n (1 - p_q)^{N_i - n}$$

# The Expected Number of Good Ideas I

- Define:
$$T_i = N_i \underline{q}^{\theta}$$
which can remain finite as $N_i \to \infty$ by sending $\underline{q} \to 0$
- Define:
$$\lambda_{i,q} = N_i p_q = T_i q^{-\theta}$$
the expected number of ideas with quality better than $q$, where $q \geq \underline{q}$
- so that:
$$p_q = \frac{\lambda_{i,q}}{N_i}$$

# The Expected Number of Good Ideas II

- Substitute into the probability above to get:

$$\begin{aligned}
\Pr\left[N_{i,q}=n\right] &= \frac{N_i!}{(N_i-n)!n!}p_q^n(1-p_q)^{N_i-n} \\
&= \frac{N_i!}{(N_i-n)!n!}\left(\frac{\lambda_{i,q}}{N_i}\right)^n\left(1-\frac{\lambda_{i,q}}{N_i}\right)^{N_i-n} \\
&= \frac{\lambda_{i,q}^n}{n!}\left(1-\frac{\lambda_{i,q}}{N_i}\right)^{N_i}\left(1-\frac{\lambda_{i,q}}{N_i}\right)^{-n}\frac{N_i}{N_i}\cdot\frac{N_i-1}{N_i} \\
&\quad \cdot \ldots \cdot \frac{N_i-n+1}{N_i}
\end{aligned}$$

# From the Pareto and Binomial to the Poisson

- Fixing $\lambda_q$ and $n$ the limit as $N_i \to \infty$ is:

$$\Pr\left[N_{i,q} = n\right] = \frac{\lambda_{i,q}^n}{n!} e^{-\lambda_{i,q}}$$

  the Poisson distribution with parameter $\lambda_{i,q} = T_i q^{-\theta}$

- Note that, by fixing $\lambda_q$ and taking $N_i \to \infty$, we're taking $\underline{q}, p_q \to 0$

# Back to Space

- The number of ideas from $i$ that deliver to $n$ at unit cost $C \leq c$ is the number with $Q \geq w_i d_{ni} / c$ which is distributed Poison with parameter:

$$\Phi_{ni} c^{\theta}$$

where:

$$\Phi_{ni} = T_i \left( w_i d_{ni} \right)^{-\theta}$$

- The number of ideas that can deliver to $n$ from anywhere at a unit cost $C \leq c$ is distributed Poisson with parameter

$$\Phi_n = \sum_i \Phi_{ni}$$

# The Distribution of Order Statistics I

- Consider ideas in terms of their order according to efficiency:

$$Q^{(1)} > Q^{(2)} > Q^{(3)} > \ldots$$

and their corresponding unit cost

$$C^{(k)} = \frac{w}{Q^{(k)}}$$

so that:

$$C^{(1)} < C^{(2)} < C^{(3)} < \ldots$$

# The Distribution of Order Statistics II

- From the Poisson, the distribution of the $k$th best idea $Q^{(k)}$ is:

$$\Pr[Q^{(k)} \leq q] = e^{-Tq^{-\theta}} \sum_{i=0}^{k-1} \frac{\left(Tq^{-\theta}\right)^i}{i!}$$

that is, the probability that **at most** $k-1$ ideas exceed $q$.

# From the Poisson to the Fréchet

- Of particular interest is the distribution of the best idea

$$\Pr[Q^{(1)} \leq q] = e^{-Tq^{-\theta}}$$

i.e., the probability that no idea has quality better than $q$, giving us the type II extreme value or the Fréchet distribution.

# The Distribution of Unit Costs

- The corresponding distribution of the $k$th lowest cost $C^{(k)}$ is:

$$G^{(k)}(c) = \Pr[C^{(k)} \leq c] = 1 - e^{-\Phi c^{\theta}} \sum_{i=0}^{k-1} \frac{\left(\Phi c^{\theta}\right)^{i}}{i!},$$

  that is, one minus the probability that **any** of the lowest $k-1$ costs exceed $c$.

- Of particular interest is the distribution of the lowest cost $C^{(1)}$

$$\Pr[C^{(1)} \leq c] = \Pr[Q^{(1)} \geq c/w] = 1 - e^{-\Phi c^{\theta}}$$

- and the second lowest cost $C^{(2)}$

$$\begin{aligned} \Pr[C^{(2)} &\leq c] = \Pr[Q^{(2)} \geq c/w] \\ &= 1 - e^{-\Phi c^{\theta}} - \Phi c^{\theta} e^{-\Phi c^{\theta}} \end{aligned}$$