Trading Partners and Trading Volumes

Elhanan Helpman
Harvard University

Marc J. Melitz
Harvard University

Yona Rubinstein
Tel-Aviv University

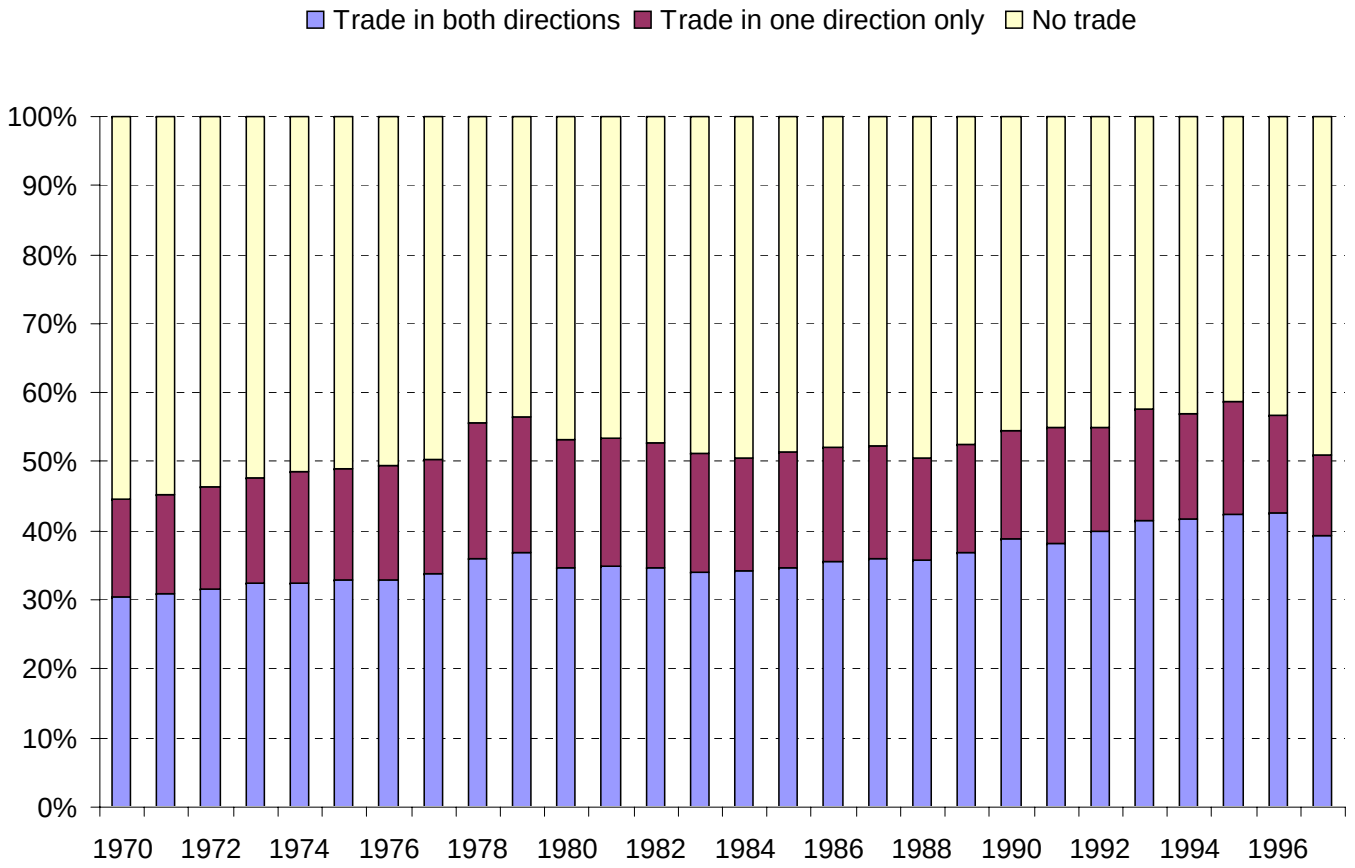
Motivation

- Estimation of gravity models of trade are an important tool in economic analysis
  - Within trade literature: measurement of trade barriers/enhancers:
    - National borders
    - Common currencies
    - Common Language
    - WTO/GATT membership
    - Free Trade Agreements
    - Cultural ties and immigration networks
  - In broader economic literature: indirect use of gravity model
    - Use effect of geography on trade as an instrument to measure contributions to economic growth

Motivation (Cont.)

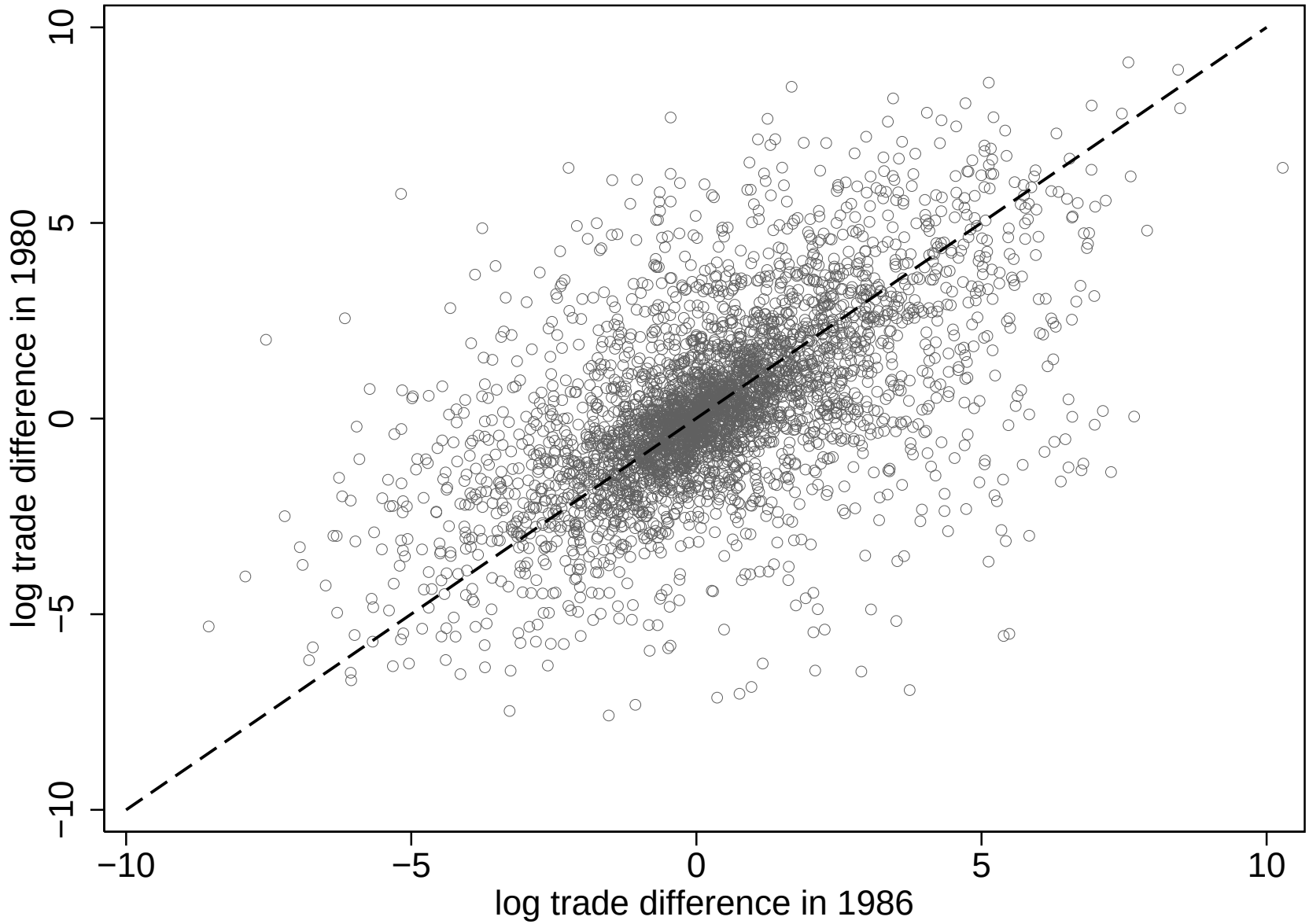
- Recent work has highlighted the importance of theoretical foundations for proper inference of empirical estimations of gravity models
- Yet (with some exceptions), theoretical foundations for gravity model assume:
  - Positive trade between all potential trade partners
  - Symmetric bi-lateral trade flows
  - Same number of traded goods (exporting firms) across destinations (i.e. no change in the extensive margin of trade across trading partners)
- Empirically, these characteristics of trade patterns are pervasive:

Evidence on Positive Trade Between Potential Trade Partners

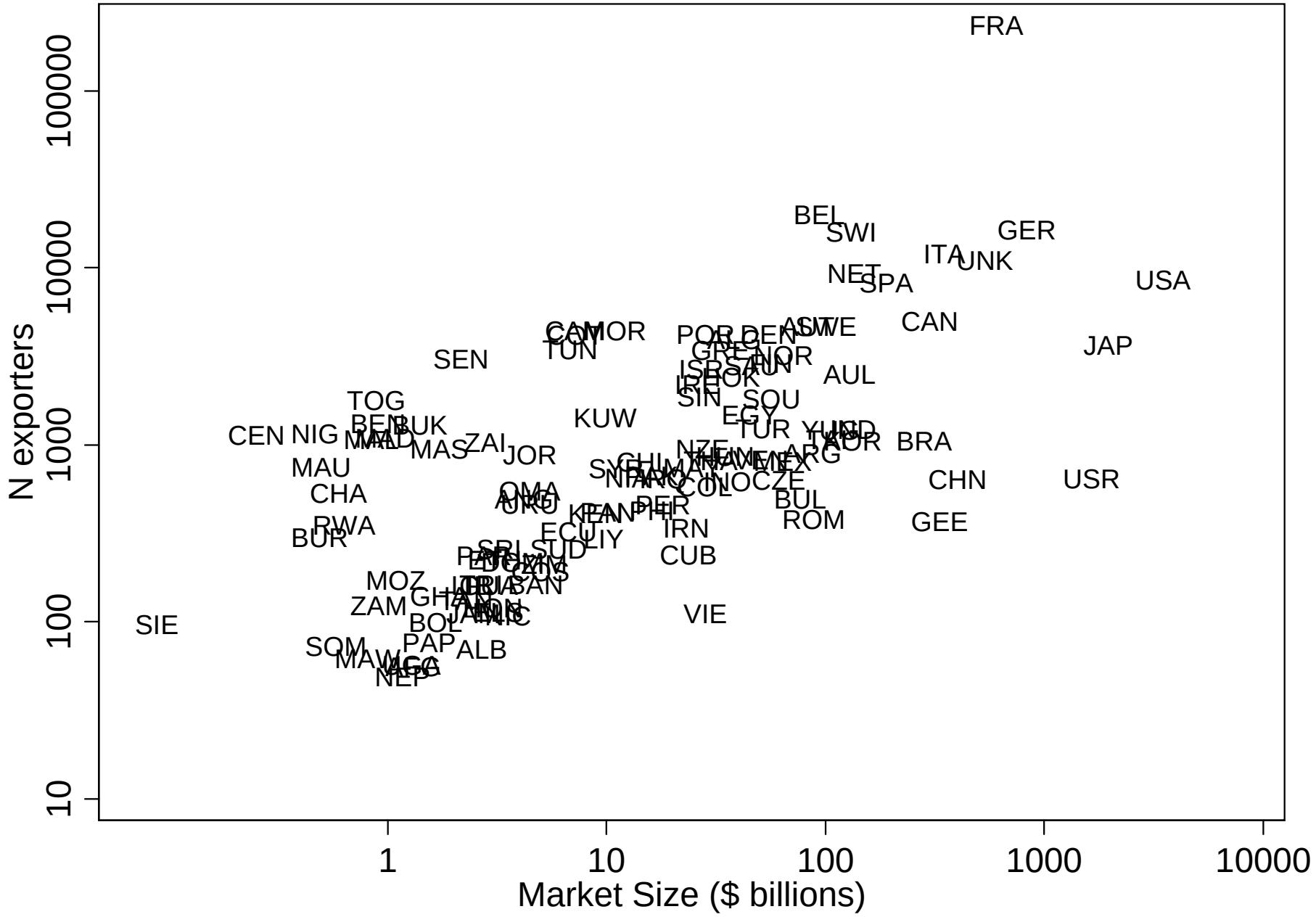


⟶ Average 7.5% growth rate of trade over these years reflects growth of trade between existing trading partners

Evidence on Bi-Lateral Trade Imbalances

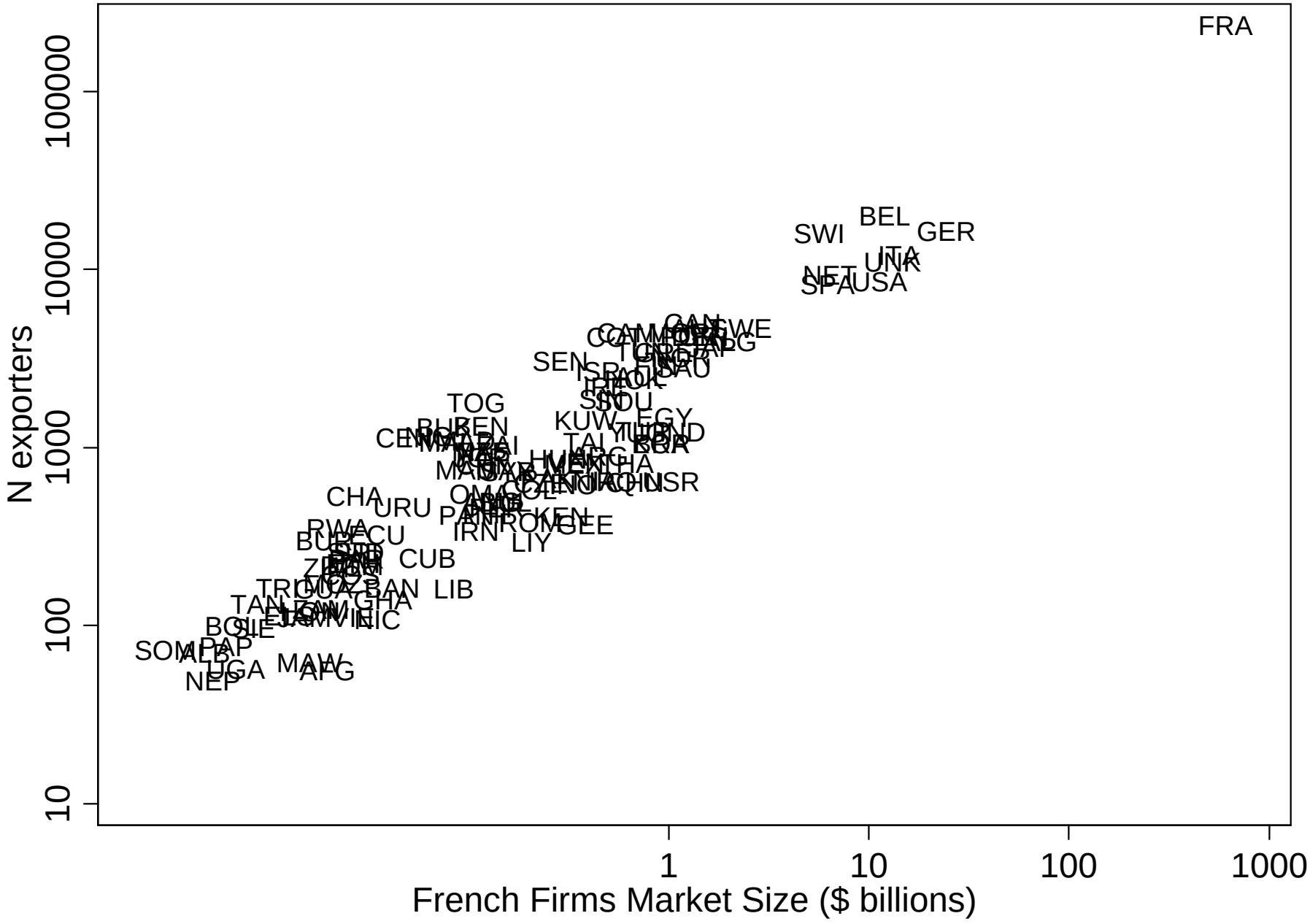


Evidence on the Extensive Margin of Trade



Source: Eaton, Kortum, and Kramarz (AER 2004, P&P)

Evidence on the Extensive Margin of Trade



Source: Eaton, Kortum, and Kramarz (AER 2004, P&P)

Relevance...

- Do these characteristics of trade patterns affect inferences based on the gravity model?
- Are these differences quantitatively important?
- Can estimation of the gravity model be amended to address these characteristics?

Contributions

- We build a theoretical “gravity” model that incorporates:
  - Potential for zero trade between countries
  - Asymmetric trade between country pairs
  - Endogenous determination of number of trade goods (exporting firms) across destinations
- We develop estimation methods tied to this theoretical framework
  - Our empirical model:
    - Predicts the formation of trading relationships (who trades with one-another)
    - Predicts the direction of trade and bi-lateral net trade imbalances *along with* the volumes of bi-lateral trade
    - Separately identifies the effects of trade barriers on the intensive and extensive margins of trade

Model Highlights

- Countries produce and consume a continuum of differentiated varieties
- Each firms produces its own individual variety
- Monopolistic competition between all firms active in any given country/market
- Firms are heterogeneous as they produce with different productivity levels
- Firms face both fixed and variable (iceberg) trade costs
- Due to fixed trade costs, only a portion of more productive firms (possibly none) export to any destination
- Product variety endogenously varies across countries

Model Structure

- Country j's preferences are given by

$$u_j = \left[\int_l x_j(l)^\alpha dl\right]^{1/\alpha}, \quad 0 < \alpha < 1$$

  - CES with elasticity $\varepsilon = 1/(1-\alpha)$
  - P_j is CES price index (defined over varieties l available in country j)

- Mass N_j of firms in country j who produce with different unit input requirements indexed by a and characterized by cdf $G(a)$ with support on $[a_L, a_H]$
  - Cost of input bundle is c_j (also indexes country's technology)

- Firms from country j selling to country i face:
  - Iceberg trade cost $\tau_{ij} > 1$
  - Fixed trade cost $f_{ij} > 0$ (measured in input units)

- If firm from j exports to i, then its export sales and profits (earned in j) are given by:

$$r_{ij}(a) = \left(\frac{\tau_{ij} c_j a}{\alpha P_i}\right)^{1-\varepsilon} Y_i$$

$$\pi_{ij}(a) = (1-\alpha)\left(\frac{\tau_{ij} c_j a}{\alpha P_i}\right)^{1-\varepsilon} Y_i - c_j f_{ij}$$

  where Y_i is aggregate income of importer i

Model Implications for Bi-Lateral Trade Flows

- All firms with $a \leq a_{ij}$ export, where:

$$\pi_{ij}(a_{ij}) = 0 \iff (1-\alpha)\left(\frac{\tau_{ij}c_j a_{ij}}{\alpha P_i}\right)^{1-\varepsilon} Y_i = c_j f_{ij} \quad \text{(zero export cutoff)}$$

  - $a_{ij} < a_L$ if $\pi_{ij}(a) < 0, \forall a \quad \Longrightarrow \quad$ No exports from j to i

- Exports from j to i are given by:

$$\begin{aligned} M_{ij} &= N_j \int_{a_L}^{a_{ij}} r_{ij}(a) dG(a) \\ &= \left(\frac{\tau_{ij}c_j}{\alpha P_i}\right)^{1-\varepsilon} Y_i N_j V_{ij} \end{aligned}$$

  where $V_{ij} = \int_{a_L}^{a_{ij}} a^{1-\varepsilon} dG(a)$ if $a_{ij} > a_L$ (otherwise, $V_{ij} = 0$ and $M_{ij} = 0$)

- Parametrization assumption: $G(a) = a^k / \left(a_H^k - a_L^k\right)$
  - Productivity $1/a$ is distributed Pareto(k) – truncated on $[1/a_H, 1/a_L]$
  - Then:

$$V_{ij} \propto W_{ij} \equiv \max\left\{\left(\frac{a_{ij}}{a_L}\right)^{k-\varepsilon+1} - 1, 0\right\}$$

    (constant of proportionality does not depend on country characteristics)
  - V_{ij} and W_{ij} increase monotonically with the share of exporting firms $G(a_{ij})$

Model Implications for Bi-Lateral Trade Flows (Cont.)

- Determination of aggregate bi-lateral trade flows:

$$M_{ij} = \left(\frac{\tau_{ij} c_j}{\alpha P_i} \right)^{1-\varepsilon} Y_i N_j V_{ij},$$

  together with export cutoff condition

$$(1-\alpha) \left(\frac{\tau_{ij} c_j a_{ij}}{\alpha P_i} \right)^{1-\varepsilon} Y_i = c_j f_{ij}$$

  completely determines bi-lateral trade flows as a function of country characteristics and trade barriers τ_{ij} and f_{ij}.

- Whenever $M_{ij} > 0$, trade volumes can be written in log-linear form:

$$m_{ij} = (\varepsilon - 1) \ln \alpha - (\varepsilon - 1) \ln c_j + n_j + (\varepsilon - 1) p_i + y_i + (1 - \varepsilon) \ln \tau_{ij} + v_{ij}$$

- Assume iceberg trade cost τ_{ij} is stochastic due to unobserved iid trade costs $u_{ij} \sim N(0, \sigma_u^2)$:

$$\tau_{ij}^{\varepsilon - 1} \equiv D_{ij}^{\gamma} e^{-u_{ij}}, \quad \text{where } D_{ij} \text{ is (symmetric) distance between } i \text{ and } j$$

- Then

$$m_{ij} = \beta_0 + \lambda_j + \chi_i - \gamma d_{ij} + w_{ij} + u_{ij},$$

  where $\chi_i = (\varepsilon - 1) p_i + y_i$ is an importer fixed effect and $\lambda_j = -(\varepsilon - 1) \ln c_j + n_j$ is an exporter fixed effect

Empirical Issues

Recall the derived gravity equation:

$$m_{ij} = \beta_0 + \lambda_j + \chi_i - \gamma d_{ij} + w_{ij} + u_{ij}, \qquad u_{ij} \sim N(0, \sigma_u^2)$$

- w_{ij} is a monotonic function of a_{ij} cutoff, which is determined by country characteristics and trade barriers
- In general, $\lambda_i \neq \chi_i$ and $w_{ij} \neq w_{ji}$ so bilateral trade need not be balanced – *even* when all trade barriers are symmetric
- $E[u_{ij} \mid M_{ij} > 0] \neq 0$ and may be correlated with the regressors (Heckman selection bias)
- Most importantly, w_{ij} is unobserved

Identification Strategy – Main Idea

- w_{ij} is a monotonic function of a_{ij} – determined by the export cutoff condition
- We also do not observe a_{ij}, but we observe whenever a_{ij} is above the threshold level a_L
- Our model explains how a_{ij} is determined by a combination of observable country characteristics and some unknown parameters – which we estimate

Identification Strategy – Details

- We define a latent variable Z_{ij} as:

$$Z_{ij} = \frac{\pi_{ij}(a_L) + c_j f_{ij}}{c_j f_{ij}} = \frac{(1-\alpha)\left(P_i \frac{\alpha}{c_j \tau_{ij}}\right)^{\varepsilon-1} Y_i a_L^{1-\varepsilon}}{c_j f_{ij}}$$

  This is the ratio of variable export profits for the most productive firm (with productivity $1/a_L$) to the fixed export costs (common to all exporters) for exports from j to i.

- Furthermore,

$$M_{ij} > 0 \iff a_{ij} > a_L \iff Z_{ij} > 1$$

- Whereupon $Z_{ij} = (a_{ij}/a_L)^{\varepsilon-1}$ is a monotonic function of $W_{ij} = (a_{ij}/a_L)^{k-\varepsilon+1} - 1$

Recovering the Z_{ij}s and W_{ij}s

- We assume that the fixed export costs f_{ij} are stochastic due to unmeasured iid trade frictions $\nu_{ij} \sim N(0, \sigma_\nu^2)$
- Let $f_{ij} \equiv \exp\left(\phi_{EX,j} + \phi_{IM,i} + \kappa\phi_{ij} - \nu_{ij}\right)$, where:
  - $\phi_{IM,i}$ is a fixed trade barrier imposed by the importing country on all exporters
  - $\phi_{EX,j}$ is a measure of fixed export costs common across all export destinations
  - ϕ_{ij} is an observed measure of any additional country-pair specific fixed trade costs
- Z_{ij} can then be written in log-linear form:

$$z_{ij} = \gamma_0 + \xi_j + \zeta_i - \gamma d_{ij} - \kappa\phi_{ij} + \eta_{ij}, \qquad \eta_{ij} \equiv u_{ij} + \nu_{ij} \sim N(0, \sigma_\eta)$$

where $\xi_j \equiv -\varepsilon \log c_j + \phi_{EX,j}$ and $\zeta_i \equiv (\varepsilon - 1)\, p_i + y_i$ are exporter and importer fixed effects

Recovering the Z_{ij}s and W_{ij}s

- Recall the log-linear form:

$$z_{ij} = \gamma_0 + \xi_j + \zeta_i - \gamma d_{ij} - \kappa \phi_{ij} + \eta_{ij}, \qquad \eta_{ij} \sim N(0, \sigma_\eta)$$

  - d_{ij} and ϕ_{ij} are observed variable and fixed bilateral trade barriers
  - ξ_j and ζ_i are exporter and importer fixed effects

- Divide both sides by σ_η:

$$z^*_{ij} = \gamma^*_0 + \xi^*_j + \zeta^*_i - \gamma^* \ln d_{ij} - \kappa^* \ln \phi_{ij} + \eta^*_{ij}, \qquad \text{so} \quad \eta^*_{ij} \sim N(0, 1)$$

- We can then estimate the following Probit equation:

$$\rho_{ij} = \Pr(T_{ij} = 1 \mid \cdot) = \Phi(\gamma^*_0 + \xi^*_j + \zeta^*_i - \gamma^* \ln d_{ij} - \kappa^* \ln \phi_{ij}),$$

  - T_{ij} is a 0-1 indicator variable equal to 1 whenever $z^*_{ij} > 0$ $(M_{ij} > 0)$
  - ρ_{ij} is the conditional probability for positive exports from j to i
  - Φ is the cdf of the unit normal distribution

Recovering the Z_{ij}s and W_{ij}s and Estimating the Gravity Equation

- The probit equation yields consistent estimates for $\hat{\rho}_{ij}$ and:

$$\begin{array}{ll}
\hat{z}^*_{ij} = \Phi^{-1}\left(\hat{\rho}_{ij}\right) & \text{for } E\left[z^*_{ij} \mid .\right] \\
\hat{w}_{ij} = \ln\left\{\exp\left[\delta \hat{z}^*_{ij}\right] - 1\right\} & \text{for } E\left[w^*_{ij} \mid .\right] \quad \text{where } \delta \equiv \sigma_\eta \left(k - \varepsilon + 1\right) / \left(\varepsilon - 1\right) \\
\hat{\bar{\eta}}^*_{ij} = \phi\left(\hat{z}^*_{ij}\right) / \Phi\left(\hat{z}^*_{ij}\right) & \text{for } E\left[\eta^*_{ij} \mid ., T_{ij} = 1\right] \quad \text{(Mills Ratio)} \\
\hat{\bar{w}}_{ij} = \ln\left\{\exp\left[\delta\left(\hat{z}^*_{ij} + \hat{\bar{\eta}}^*_{ij}\right)\right] - 1\right\} & \text{for } E\left[w^*_{ij} \mid ., T_{ij} = 1\right]
\end{array}$$

- We can now use $\hat{\bar{w}}_{ij}$ and $\hat{\bar{\eta}}^*_{ij} \propto E\left[u_{ij} \mid ., T_{ij} = 1\right]$ to obtain consistent estimates of the gravity equation:

$$\begin{aligned}
m_{ij} &= \beta_0 + \lambda_j + \chi_i - \gamma \ln d_{ij} + w_{ij} + u_{ij} \\
&= \beta_0 + \lambda_j + \chi_i - \gamma \ln d_{ij} + \ln\left(\exp\left(\delta\left(\hat{z}^*_{ij} + \bar{\eta}^*_{ij}\right)\right) - 1\right) + \beta_\lambda \bar{\eta}^*_{ij} + e_{ij},
\end{aligned}$$

  where $E\left[e_{ij} \mid ., T_{ij} = 1\right] = 0$

- We thus introduce 2 new controls:
  - $\hat{z}^*_{ij}$ for the unobserved proportion of exporting firms
  - $\hat{\bar{\eta}}^*_{ij}$ for the sample selection bias
  - ... along with separate exporter and importer fixed effects

- Estimating equation is non-linear in δ so we use maximum likelihood

Data: Main Independent Bi-Lateral Variables

Distance: distance between both countries

Land border: (indicator) countries share a land border

Island: (indicator) at least one country is an island

Landlock: (indicator) at least one country is landlocked

Legal: (indicator) countries have the same legal origins

Language: (indicator) countries i and j share the same language

Colonial Ties: (indicator) one country colonized the other

Currency union: (indicator) countries use a common currency

WTO (none): (indicator) neither in the WTO

WTO (both): (indicator) both countries in the WTO

FTA: (indicator) both countries belong to the same regional trade agreement

Exclusion Restriction

- Need some observable trade barrier that does not affect per-unit cost of trade
  - i.e. a variable that is in ϕ_{ij} that does not enter in d_{ij}
  - other variables can have both a fixed and a per-unit component
- We use two different sources for this excluded variable:
  - One "ad-hoc" source (main advantage: available for all countries):
    - The percent of the population in each country that shares the same religion
  - One source directly tied to firm-level fixed costs across countries:
    - The cost of starting a firm (# of days, # of procedures, cost as % of GDP per capita)
    - We interact these variables for each country pair to create a bilateral fixed cost index
    - Indicator variable set to 1 if both countries have above median levels of fixed costs

Data Sources

Trade data

- The bilateral trade flows are from Feenstra's "World Trade Flows, 1970-1992" and "World Trade Flows, 1980-1997".
- 161 countries
- 161 x 160 = 25,760 importer-exporter potential pairs per year.

Country-level data

- Population and real GDP per capita have been obtained from four standard sources:
  - Penn World Tables, World Bank, IMF and the CIA's World Factbook
- Data on legal system were provided by Florencio Lopez-de-Silanes
- Currency union were taken from Glick and Rose (2001). FTA and WTO data are from Rose (2003)
- Fixed cost data from World Bank survey

Benchmark Gravity and Probit: Table 1

	1986	
	1	2
Variables	m_ij (OLS)	T_ij (Probit)
Distance	-1.176	-0.263
	(0.031)**	(0.012)**
Land border	0.458	-0.148
	(0.147)**	(0.047)*
Legal	0.486	0.038
	(0.050)**	(0.014)*
Language	0.176	0.113
	(0.061)**	(0.016)**
Religion	0.102	0.104
	(0.096)	(0.025)**
Colonial Ties	1.299	0.128
	(0.120)**	(0.117)
Currency Union	1.364	0.190
	(0.255)**	(0.052)**
FTA	0.759	0.494
	(0.222)**	(0.020)**
WTO (none)		
WTO (both)		
Observations	11,146	24,649
R-Squared	0.709	0.587

Exporter, Importer, and year fixed effects

Robust standard errors (clustering by country pair) * significant at 5%; ** significant at 1%

Benchmark Gravity and Probit: Table 1

	1986		1980s	
	1	2	5	6
Variables	m_ij (OLS)	T_ij (Probit)	m_ij (OLS)	T_ij (Probit)
Distance	-1.176	-0.263	-1.200	-0.246
	(0.031)**	(0.012)**	(0.024)**	(0.008)**
Land border	0.458	-0.148	0.364	-0.146
	(0.147)**	(0.047)*	(0.131)**	(0.032)**
Legal	0.486	0.038	0.407	0.028
	(0.050)**	(0.014)*	(0.040)**	(0.009)**
Language	0.176	0.113	0.203	0.108
	(0.061)**	(0.016)**	(0.047)**	(0.011)**
Religion	0.102	0.104	-0.038	0.098
	(0.096)	(0.025)**	(0.077)	(0.016)**
Colonial Ties	1.299	0.128	1.326	0.116
	(0.120)**	(0.117)	(0.110)**	(0.082)
Currency Union	1.364	0.190	1.409	0.206
	(0.255)**	(0.052)**	(0.187)**	(0.026)**
FTA	0.759	0.494	0.976	0.495
	(0.222)**	(0.020)**	(0.214)**	(0.018)**
WTO (none)			-0.068	-0.056
			(0.058)	(0.013)**
WTO (both)			0.303	0.093
			(0.042)**	(0.013)**
Observations	11,146	24,649	110,697	248,060
R-Squared	0.709	0.587	0.682	0.551

Exporter, Importer, and year fixed effects
Robust standard errors (clustering by country pair) * significant at 5%; ** significant at 1%

Two Stage Estimation: Table 2

	1986		
	T_ij	m_ij	
Variables	(Probit)	Benchmark	ML
Distance	-0.660 (0.029)**	-1.181 (0.031)**	-0.801 (0.030)**
Land border	-0.382 (0.129)*	0.468 (0.146)**	0.831 (0.139)**
Legal	0.096 (0.034)*	0.490 (0.050)**	0.388 (0.049)**
Language	0.284 (0.042)**	0.187 (0.061)*	0.024 (0.06)
Religion	0.261 (0.063)**	--	--
Colonial Ties	0.325 (0.305)	1.299 (0.121)**	1.003 (0.114)**
Currency Union	0.492 (0.143)**	1.356 (0.256)**	1.026 (0.258)**
FTA	1.985 (0.315)**	0.756 (0.222)**	0.386 (0.171)*
WTO (none)	--	--	--
WTO (both)	--	--	--
delta (from w_hat)	--	--	0.716 (0.060)**
eta_hat	--	--	0.399 (0.063)**
Observations	24,649	11,146	11,146

Exporter, Importer, and year fixed effects
Robust standard errors (clustering by country pair) * significant at 5%; ** significant at 1%

Two Stage Estimation: Table 2

	1986			1980s		
	T_ij	m_ij		T_ij	m_ij	
Variables	(Probit)	Benchmark	ML	(Probit)	Benchmark	ML
Distance	-0.660	-1.181	-0.801	-0.618	-1.198	-0.822
	(0.029)**	(0.031)**	(0.030)**	(0.021)**	(0.024)**	(0.024)**
Land border	-0.382	0.468	0.831	-0.380	0.360	0.702
	(0.129)*	(0.146)**	(0.139)**	(0.089)**	(0.131)**	(0.123)**
Legal	0.096	0.490	0.388	0.071	0.406	0.327
	(0.034)*	(0.050)**	(0.049)**	(0.022)**	(0.040)**	(0.039)**
Language	0.284	0.187	0.024	0.273	0.198	0.033
	(0.042)**	(0.061)*	(0.06)	(0.027)**	(0.047)**	(0.046)
Religion	0.261	--	--	0.245	--	--
	(0.063)**			(0.040)**		
Colonial Ties	0.325	1.299	1.003	0.293	1.326	1.061
	(0.305)	(0.121)**	(0.114)**	(0.211)	(0.110)**	(0.106)**
Currency Union	0.492	1.356	1.026	0.531	1.412	1.034
	(0.143)**	(0.256)**	(0.258)**	(0.071)**	(0.187)**	(0.191)**
FTA	1.985	0.756	0.386	1.842	0.978	0.519
	(0.315)**	(0.222)**	(0.171)*	(0.207)**	(0.214)**	(0.148)**
WTO (none)	--	--	--	-0.143	-0.070	0.001
				(0.033)**	(0.058)	(0.058)
WTO (both)	--	--	--	0.234	0.302	0.143
				(0.032)**	(0.042)**	(0.042)**
delta (from w_hat)	--	--	0.716	--	--	0.794
			(0.060)**			(0.067)**
eta_hat	--	--	0.399	--	--	0.270
			(0.063)**			(0.049)**
Observations	24,649	11,146	11,146	248,060	110,697	110,697

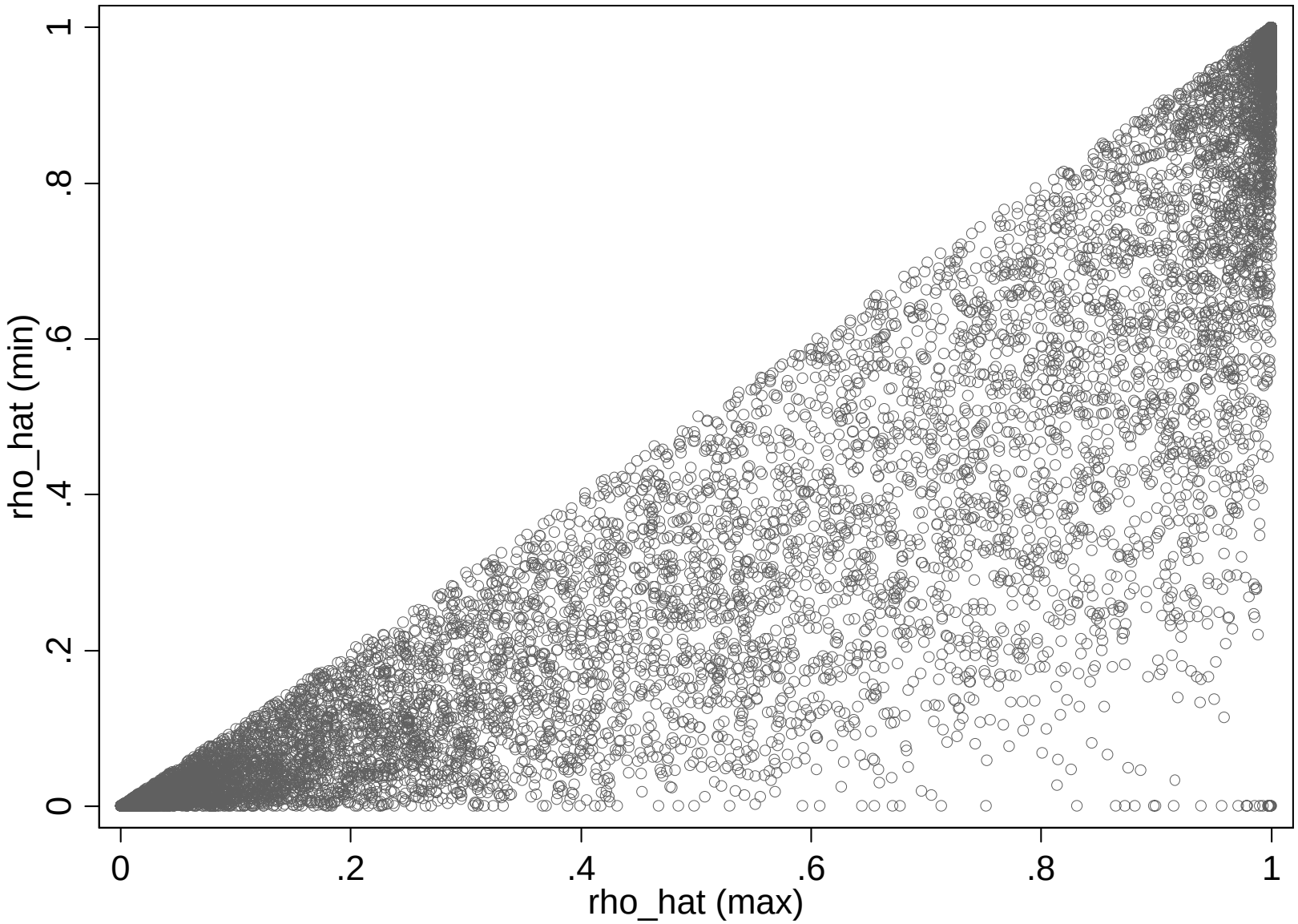
Exporter, Importer, and year fixed effects
Robust standard errors (clustering by country pair) * significant at 5%; ** significant at 1%

Decomposing the Biases: Table 3

	Dependent variable: m_ij			
Variables	Benchmark	ML	Firm Heterogeneity	Heckman Selection
Distance	-1.181 (0.031)**	-0.801 (0.030)**	-0.824 (0.036)**	-1.214 (0.031)**
Land border	0.468 (0.146)**	0.831 (0.139)**	0.807 (0.139)**	0.436 (0.149)**
Legal	0.490 (0.050)**	0.388 (0.049)**	0.420 (0.050)**	0.488 (0.050)**
Language	0.187 (0.061)**	0.024 (0.06)	-0.008 (0.061)	0.223 (0.061)**
Colonial Ties	1.299 (0.121)**	1.003 (0.114)**	1.051 (0.114)**	1.311 (0.123)**
Currency Union	1.356 (0.256)**	1.026 (0.258)**	1.028 (0.256)**	1.391 (0.257)**
FTA	0.756 (0.222)**	0.386 (0.171)*	0.502 (0.160)**	0.737 (0.235)**
delta (from w_hat)	--	0.716 (0.060)**	--	--
eta_hat	--	0.399 (0.063)**	--	0.265 (0.070)**
z_hat	--	--	0.611 (0.043)**	--
Observations	11,146	11,146	11,146	11,146
R-squared	0.709	--	0.713	0.710

Exporter, Importer, and year fixed effects
Robust standard errors (clustering by country pair) * significant at 5%; ** significant at 1%

Predictions for Asymmetric Trade: Figure 1



Predicted Asymmetries: $\hat{\rho}_{ij}$ versus $\hat{\rho}_{ji}$

Explanatory Power for Asymmetric Trade: Table 4

Variable	T_ij - T_ji
rho_hat_ij - rho_hat_ji	0.994 (0.023)**
Country Fixed Effects	No
Observations	12403
R-Squared	0.228

Variable	m_ij - m_ji	
w_hat_ij - w_hat_ji	2.073 (0.079)**	1.820 (0.320)**
Country Fixed Effects	No	Yes
Observations	4652	4652
R-Squared	0.156	0.299

* significant at 5%; ** significant at 1%

Conclusions

- Traditional Estimation of Gravity Equation confounds effects of trade barriers on the intensive and extensive margin
  - This induces serious biases in the measured effects of trade barriers
- We show how consistent estimates for these effects can be obtained using information only for aggregate trade flows
- In addition, our model:
  - Can separately measure the effects of fixed versus variable trade costs
  - Predict the formation of trading relationships
  - Predicts bi-lateral trade imbalances