14.582: International Trade II
— Lecture 5: Firm-level Heterogeneity (Empirics III) —

Plan for Today's Lecture

1. Stylized facts on extensive margin in trading behavior
2. Implications of extensive margin for gravity estimation:
    1. Helpman, Melitz and Rubinstein (2008)
    2. Crozet and Konig (2010)
3. Path-dependent sunk costs of exporting: Morales, Sheu and Zahler (2017)
4. Uncertainty and export behavior: Dickstein and Morales (2016)

Plan for Today's Lecture

1. **Stylized facts on extensive margin in trading behavior**
2. Implications of extensive margin for gravity estimation:
    1. Helpman, Melitz and Rubinstein (2008)
    2. Crozet and Konig (2010)
3. Path-dependent sunk costs of exporting: Morales, Sheu and Zahler (2017)
4. Uncertainty and export behavior: Dickstein and Morales (2016)

Intensive and Extensive Margins in Trade Flows

- With access to micro data on trade flows at the firm-level, a question that has been explored is whether trade flows expand over time (or look bigger in the cross-section) along the:
  - Intensive margin: the same firms (or product-firms) from country i export more volume (and/or charge higher prices—we could also decompose the intensive margin into these two margins) to country j.
  - Extensive margin: new firms (or product-firms) from country i are penetrating the market in country j.
- This is really just a decomposition—we should expect trade to expand along both margins.
- Recently some papers have been able to look at this.
  - A rough lesson from these exercises is that the extensive margin seems more important (in a purely 'accounting' sense, not necessarily a causal sense).

Bernard, Jensen, Redding and Schott (2007): Exporters

Data from US manufacturing firms. The coefficients in columns 2-4 sum (across columns) to those in column 1.

Table 6

Gravity and Aggregate U.S. Exports, 2000

	Log of total exports value	*Log of number of exporting firms*	*Log of number of exported products*	*Log of export value per product per firm*
Log of GDP	0.98	0.71	0.52	−0.25
	(0.04)	(0.04)	(0.03)	(0.04)
Log of distance	−1.36	−1.14	−1.06	0.84
	(0.17)	(0.16)	(0.15)	(0.19)
Observations	175	175	175	175
R^2	0.82	0.74	0.64	0.25

Sources: Data are from the 2000 Linked-Longitudinal Firm Trade Transaction Database (LFTTD).
Notes: Each column reports the results of a country-level ordinary least squares regression of the dependent variable noted at the top of each column on the covariates noted in the first column. Results for the constant are suppressed. Standard errors are noted below each coefficient. Products are defined as ten-digit Harmonized System categories. All results are statistically significant at the 1 percent level.

Bernard, Jensen, Redding and Schott (2007): Importers

Data from US manufacturing firms. The coefficients in columns 2-4 sum (across columns) to those in column 1.

Table 9

Gravity and Aggregate U.S. Imports, 2000

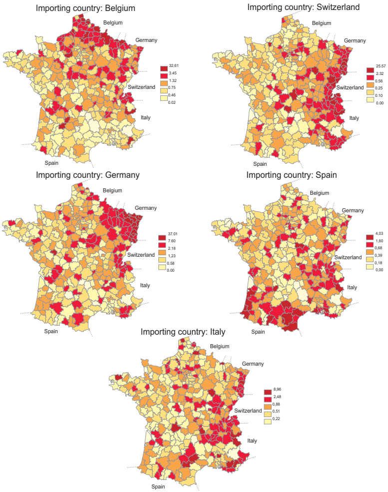
	Log of total import value	*Log of number of importing firms*	*Log of number of imported products*	*Log of import value per product per firm*
Log of GDP	1.14***	0.82***	0.71***	−0.39***
	(0.06)	(0.03)	(0.03)	(0.05)
Log of Distance	−0.73***	−0.43***	−0.61***	0.31
	(0.27)	(0.15)	(0.15)	(0.24)
Observations	175	175	175	175
R^2	0.69	0.78	0.74	0.25

Sources: Data are from the 2000 Linked-Longitudinal Firm Trade Transaction Database (LFTTD).
Notes: Each column reports the results of a country-level ordinary least squares regression of the dependent variable noted at the top of each column on the covariates listed on the left. Results for constants are suppressed. Standard errors are noted below each coefficient. Products are defined as ten-digit Harmonized System categories.
*, **, and *** represent statistical significance at the 10, 5, and 1 percent levels, respectively.

CK (2010): Intensive margin

Data from French manufacturing firms trading internationally, by domestic region j.

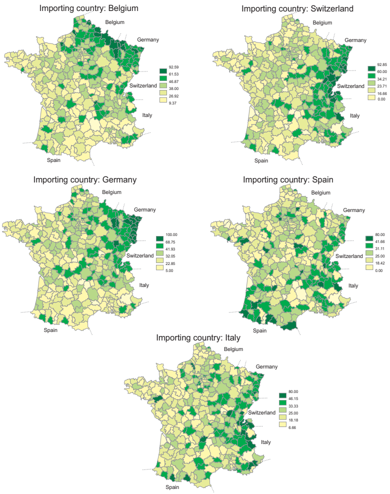
Figure 1: Mean value of individual-firm exports (single-region firms, 1992)



CK (2010): Extensive margin

Data from French manufacturing firms trading internationally, by domestic region j. (NB: Extensive margin here based only on included firms, which had over 20 workers.)

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



Crozet and Koenig (CJE, 2010)

Data from French manufacturing firms trading internationally, by domestic region j.

Table 2: Decomposition of French aggregate industrial exports (34 industries - 159 countries - 1986 to 1992)

	All firms $>$ 20 employees		Single-region firms $>$ 20 employees	
	(1)	(2)	(3)	(4)
	Average Shipment ln (M_{kjt}/N_{kjt})	Number of Shipments ln (N_{kjt})	Average Shipment ln (M_{kjt}/N_{kjt})	Number of Shipments ln (N_{kjt})
ln (GDP_{kj})	0.461^a	0.417^a	0.421^a	0.417^a
	(0.007)	(0.007)	(0.007)	(0.008)
ln ($Dist_j$)	-0.325^a	-0.446^a	-0.363^a	-0.475^a
	(0.013)	(0.009)	(0.012)	(0.009)
$Contig_j$	-0.064^c	-0.007	0.002	0.190^a
	(0.035)	(0.032)	(0.038)	(0.036)
$Colony_j$	0.100^a	0.466^a	0.141^a	0.442^a
	(0.032)	(0.025)	(0.035)	(0.027)
$French_j$	0.213^a	0.991^a	0.188^a	1.015^a
	(0.029)	(0.028)	(0.032)	(0.028)
N	23553	23553	23553	23553
R^2	0.480	0.591	0.396	0.569

Note: These are OLS estimates with year and industry dummies. Robust standard errors in parentheses with ^{a}, ^{b} and ^{c} denoting significance at the 1%, 5% and 10% level respectively.

Hilberry and Hummels (EER, 2008)

Data on intra-national US commodity shipping (zipcode-to-zipcode, with firm identifiers).

Table 2. Decomposing Spatial Frictions
(5-digit zip code data)

	dist	dist^2	ownzip	ownstate	constant	Adj. R^2	N	ε_D
value (T_{ij})	**-0.137** **(0.009)**	**-0.004** **(0.001)**	**1.102** **(0.030)**	**-0.024** **(0.007)**	**-13.393** **(0.026)**	0.01	1290788	-0.187
# of shipments (N_{ij})	**-0.294** **(0.002)**	**0.017** **(0.000)**	**0.883** **(0.008)**	**0.043** **(0.002)**	**-1.413** **(0.007)**	0.10	1290840	-0.081
# of trading pairs (N_{ij}^F)	-0.159 (0.002)	0.008 (0.000)	0.540 (0.007)	0.029 (0.002)	-0.888 (0.006)	0.05	1290840	-0.059
# of commodities (N_{ij}^k)	-0.135 (0.001)	0.009 (0.000)	0.342 (0.003)	0.014 (0.001)	-0.525 (0.003)	0.10	1290840	-0.022
avg. value ($\overline{PQ}_{ij}$)	**0.157** **(0.008)**	**-0.021** **(0.001)**	**0.219** **(0.028)**	**-0.067** **(0.006)**	**-11.980** **(0.024)**	0.00	1290788	-0.106
avg. price ($\bar{P}_{ij}$)	-0.032 (0.007)	0.036 (0.001)	-0.115 (0.024)	-0.154 (0.006)	0.021 (0.020)	0.08	1290788	0.419
avg. weight ($\bar{Q}_{ij}$)	0.189 (0.011)	-0.058 (0.001)	0.334 (0.037)	0.087 (0.009)	-12.001 (0.031)	0.05	1290788	-0.537

Notes:
1. Regression of (log) shipment value and its components from equations (7) and (8) on geographic variables. Dependent variables in left hand column. Coefficients in right-justified rows sum to coefficients in left justified rows.
2. Standard errors in parentheses.
3. ε_D is the elasticity of trade with respect to distance, evaluated at the sample mean distance of 523 miles.

Plan for Today's Lecture

1. Stylized facts on extensive margin in trading behavior
2. **Implications of extensive margin for gravity estimation:**
   1. **Helpman, Melitz and Rubinstein (2008)**
   2. Crozet and Konig (2010)
3. Path-dependent sunk costs of exporting: Morales, Sheu and Zahler (2017)
4. Uncertainty and export behavior: Dickstein and Morales (2016)

Helpman, Melitz and Rubenstein (QJE, 2008)

- What does the difference between intensive and extensive margins imply for the estimation of gravity equations?
  - Gravity equations are often used as a tool for measuring trade costs and the determinants of trade costs—we will see more of this in Lecture 15.
- HMR (2008) started wave of thinking about gravity equation estimation in the presence of extensive/intensive margins.
  - They use aggregate international trade (so, technically, this paper doesn't belong in a lecture on "empirical work on firm-level heterogeneity"!) to explore implications of a heterogeneous firm model for gravity equation estimation.
  - The Melitz (2003) model is simplified and used as a tool to understand, estimate, and correct for biases in gravity equation estimation.

HMR (2008): Zeros in Trade Data

- HMR start with the observation that there are lots of 'zeros' in international trade data, even when aggregated up to total bilateral exports.
  - Baldwin and Harrigan (2008) and Johnson (2008) look at this in a more disaggregated (sectoral) manner and find (unsurprisingly) far more zeros.
- Zeros are also problematic.
  - A typical analysis of trade flows is based on the gravity equation (in logs), which can't incorporate $X_{ij} = 0$.
  - Indeed, other models of the gravity equation (Armington, Krugman, Eaton-Kortum, Melitz-with-Pareto) don't have any zeros in them (due to symmetric CES demand, unbounded productivities, and finite trade costs).

HMR (2008)

The extent of zeros, even at the aggregate export level

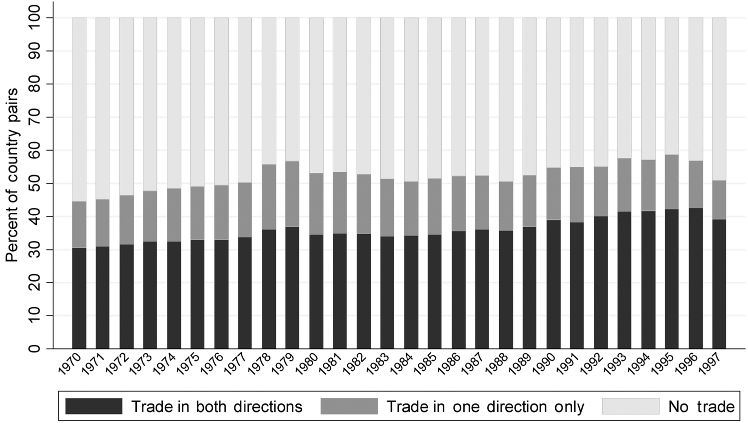


FIGURE I
Distribution of Country Pairs Based on Direction of Trade
Note. Constructed from 158 countries.

HMR (2008)

The growth of trade in recent decades is not due to the death of zeros

FIGURE II
Aggregate Volume of Exports of All Country Pairs and of Country Pairs That Traded in Both Directions in 1970

A Gravity Model with Zeroes

- HMR work with a multi-country version of Melitz (2003)—similar to Chaney (2008).
- Set-up:
  - Monopolistic competition, CES preferences (ε), one factor of production (unit cost c_j), one sector.
  - Both variable (iceberg τ_{ij}) and fixed (f_{ij}) costs of exporting.
  - Heterogeneous firm-level productivities $1/a$ drawn from truncated Pareto, $G(a)$. This is the feature that allows for zeroes to exist (at finite trade costs).
- Some firms in j sell in country i iff $a \leq a_{ij}$, where the cutoff productivity (a_{ij}) is defined by:

$$\kappa_1 \left(\frac{\tau_{ij} c_j a_{ij}}{P_i} \right)^{1-\varepsilon} Y_i = c_j f_{ij} \tag{1}$$

An Augmented Gravity Equation

- HMR (2008) derive an "augmented" gravity equation, for those observations that are non-zero, of the form:

$$\ln(M_{ij}) = \beta_0 + \alpha_i + \alpha_j - \gamma \ln d_{ij} + w_{ij} + u_{ij} \tag{2}$$

- Where:
  - M_{ij} is imports (into i from j)
  - d_{ij} is distance (or potentially other observable shifters of trade costs).
  - w_{ij} is the "augmented" part, which is a term accounting for selection.
  - u_{ij} represents unobserved components of trade costs

Two Sources of Bias

- The HMR (2008) theory suggests (and solves) two sources of bias in the typical estimation of gravity equations (which neglects w_{ij}).
- First: Omitted variable bias due to the presence of w_{ij}:
  - In a model with heterogeneous firm productivities and fixed costs of exporting (i.e. a Melitz (2003) model), only highly productive firms will penetrate distant markets.
  - So distance (d_{ij}) does two things: it raises the price at which any firm can sell (thus reducing demand along the intensive margin) in a distant market; and it changes the productivity (and hence the price and hence the amount sold) of the firms entering a distant market.
  - This means that d_{ij} is correlated with w_{ij}.
  - Therefore, if one aims to estimate γ but neglects to control for w_{ij} the estimate of γ will be biased (due to OVB).

Two Sources of Bias

- The HMR (2008) theory suggests (and solves) two sources of bias in the typical estimation of gravity equations (which neglects w_{ij}).
- Second: A selection effect induced by only working with non-zero trade flows:
    - HMR's gravity equation, like those before it, can't be estimated on the observations for which $M_{ij} = 0$.
    - The HMR theory tells us that the existence of these 'zeros' is not as good as random with respect to d_{ij}, so econometrically this 'selection effect' needs to be corrected/controlled for.
    - Intuitively, the problem is that far away destinations are less likely to be profitable, so the sample of zeros is selected on the basis of d_{ij}.
    - This calls for a standard Heckman (1979) selection correction.

HMR (2008): Two-step Estimation

Two-step estimation to solve bias

1. Estimate probit for zero trade flow or not:
   - Include exporter and importer fixed effects, and d_{ij}.
   - Can proceed with just this, but then identification (in Step 2) is achieved purely off of the normality assumption.
   - To 'strengthen' identification, need additional variable that enters Probit in step 1, but does not enter Step 2.
   - Theory says this should be a variable that affects the fixed cost of exporting, but not the variable cost.
   - HMR use Djankov et al (QJE, 2002)'s 'entry regulation' index. Also try 'common religion dummy.'

2. Estimate gravity equation on positive trade flows:
   - Include inverse Mills ratio (standard Heckman procedure trick) to control for selection problem (Second source of bias)
   - Also include empirical proxy for w_{ij} based on estimate of entry equation in Step 1 (to fix First source of bias).

HMR (2008): Results (traditional gravity estimation)

TABLE I
BENCHMARK GRAVITY AND SELECTION INTO TRADING RELATIONSHIPS

	1986		1980s			
		(Probit)		(Probit)		(Probit)
Variables	m_{ij}	T_{ij}	m_{ij}	T_{ij}	m_{ij}	T_{ij}
Distance	-1.176^{**}	-0.263^{**}	-1.201^{**}	-0.246^{**}	-1.200^{**}	-0.246^{**}
	(0.031)	(0.012)	(0.024)	(0.008)	(0.024)	(0.008)
Land border	0.458^{**}	-0.148^{**}	0.366^{**}	-0.146^{**}	0.364^{**}	-0.146^{**}
	(0.147)	(0.047)	(0.131)	(0.032)	(0.131)	(0.032)
Island	-0.391^{**}	-0.136^{**}	-0.381^{**}	-0.140^{**}	-0.378^{**}	-0.140^{**}
	(0.121)	(0.032)	(0.096)	(0.022)	(0.096)	(0.022)
Landlock	-0.561^{**}	-0.072	-0.582^{**}	-0.087^{**}	-0.581^{**}	-0.087^{**}
	(0.188)	(0.045)	(0.148)	(0.028)	(0.147)	(0.028)
Legal	0.486^{**}	0.038^{**}	0.406^{**}	0.029^{**}	0.407^{**}	0.028^{**}
	(0.050)	(0.014)	(0.040)	(0.009)	(0.040)	(0.009)
Language	0.176^{**}	0.113^{**}	0.207^{**}	0.109^{**}	0.203^{**}	0.108^{**}
	(0.061)	(0.016)	(0.047)	(0.011)	(0.047)	(0.011)
Colonial ties	1.299^{**}	0.128	1.321^{**}	0.114	1.326^{**}	0.116
	(0.120)	(0.117)	(0.110)	(0.082)	(0.110)	(0.082)
Currency union	1.364^{**}	0.190^{**}	1.395^{**}	0.206^{**}	1.409^{**}	0.206^{**}
	(0.255)	(0.052)	(0.187)	(0.026)	(0.187)	(0.026)
FTA	0.759^{**}	0.494^{**}	0.996^{**}	0.497^{**}	0.976^{**}	0.495^{**}
	(0.222)	(0.020)	(0.213)	(0.018)	(0.214)	(0.018)
Religion	0.102	0.104^{**}	-0.018	0.099^{**}	-0.038	0.098^{**}
	(0.096)	(0.025)	(0.076)	(0.016)	(0.077)	(0.016)
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	110,697	248,060
R^2	0.709	0.587	0.682	0.551	0.682	0.551

Notes. Exporter, importer, and year fixed effects. Marginal effects at sample means and pseudo R^2 reported for Probit. Robust standard errors (clustering by country pair).
+ Significant at 10%.
*Significant at 5%.
**Significant at 1%.

HMR (2008): Results (gravity estimation with correction)

TABLE II
BASELINE RESULTS

	1986 reduced sample					
		m_{ij}				
					Indicator variables	
Variables	(Probit) T_{ij}	Benchmark	NLS	Polynomial	50 bins	100 bins
Distance	-0.213^{**}	-1.167^{**}	-0.813^{**}	-0.847^{**}	-0.755^{**}	-0.789^{**}
	(0.016)	(0.040)	(0.049)	(0.052)	(0.070)	(0.088)
Land border	-0.087	0.627^{**}	0.871^{**}	0.845^{**}	0.892^{**}	0.863^{**}
	(0.072)	(0.165)	(0.170)	(0.166)	(0.170)	(0.170)
Island	-0.173^{*}	-0.553^{*}	-0.203	-0.218	-0.161	-0.197
	(0.078)	(0.269)	(0.290)	(0.258)	(0.259)	(0.258)
Landlock	-0.053	-0.432^{*}	-0.347^{*}	$-0.362+$	$-0.352+$	$-0.353+$
	(0.050)	(0.189)	(0.175)	(0.187)	(0.187)	(0.187)
Legal	0.049^{**}	0.535^{**}	0.431^{**}	0.434^{**}	0.407^{**}	0.418^{**}
	(0.019)	(0.064)	(0.065)	(0.064)	(0.065)	(0.065)
Language	0.101^{**}	$0.147+$	-0.030	-0.017	-0.061	-0.036
	(0.021)	(0.075)	(0.087)	(0.077)	(0.079)	(0.083)
Colonial ties	-0.009	0.909^{**}	0.847^{**}	0.848^{**}	0.853^{**}	0.838^{**}
	(0.130)	(0.158)	(0.257)	(0.148)	(0.152)	(0.153)
Currency union	0.216^{**}	1.534^{**}	1.077^{**}	1.150^{**}	1.045^{**}	1.107^{**}
	(0.038)	(0.334)	(0.360)	(0.333)	(0.337)	(0.346)
FTA	0.343^{**}	0.976^{**}	0.124	0.241	-0.141	0.065
	(0.009)	(0.247)	(0.227)	(0.197)	(0.250)	(0.348)
Religion	0.141^{**}	0.281^{*}	0.120	0.139	0.073	0.100
	(0.034)	(0.120)	(0.136)	(0.120)	(0.124)	(0.128)
Regulation costs	-0.108^{**}	-0.146				
	(0.036)	(0.100)				
R. costs (days & proc.)	-0.061^{*}	$-0.216+$				
	(0.031)	(0.124)				
δ (from $\hat{\bar{w}}^{*}_{ij}$)			0.840^{**}			
			(0.043)			
$\hat{\bar{\eta}}^{*}_{ij}$			0.240^{*}	0.882^{**}		
			(0.099)	(0.209)		
$\hat{\bar{z}}^{*}_{ij}$				3.261^{**}		
				(0.540)		
$\hat{\bar{z}}^{*2}_{ij}$				-0.712^{**}		
				(0.170)		
$\hat{\bar{z}}^{*3}_{ij}$				0.060^{**}		
				(0.017)		
Observations	12,198	6,602	6,602	6,602	6,602	6,602
R^2	0.573	0.693		0.701	0.704	0.706

Notes. Exporter and importer fixed effects. Marginal effects at sample means and pseudo R^2 reported for Probit. Regulation costs are excluded variables in all second stage specifications. Bootstrapped standard errors for NLS; robust standard errors (clustering by country pair) elsewhere.
+Significant at 10%.
*Significant at 5%.
**Significant at 1%.

Plan for Today's Lecture

1. Stylized facts on extensive margin in trading behavior
2. **Implications of extensive margin for gravity estimation:**
   1. Helpman, Melitz and Rubinstein (2008)
   2. **Crozet and Konig (2010)**
3. Path-dependent sunk costs of exporting: Morales, Sheu and Zahler (2017)
4. Uncertainty and export behavior: Dickstein and Morales (2016)

Crozet and Koenig (CJE, 2010)

- CK (2010) conduct a similar exercise to HMR (2008), but with French firm-level data.
  - This is attractive—after all, the main point that HMR (2008) is making is that firm-level realities matter for aggregate flows.
- CK's firm data has exports to foreign countries in it (CK focus only on adjacent countries: Belgium, Switzerland, Germany, Spain and Italy).

CK (2010): Identification

- But interestingly, CK also know where the firm is in France.
- So they try to separately identify the effects of variable and fixed trade costs by assuming:
  - Variable trade costs are proportional to distance. Since each firm is a different distance from, say, Belgium, there is cross-firm variation here.
  - Fixed trade costs are homogeneous across France for a given export destination. (That is, it costs just as much to figure out how to sell to the Swiss whether your French firm is based in Geneva or Normandy).

CK (2010): The model and estimation I

- The model is deliberately close to Chaney (2008), which is a particular version of the Melitz (2003) model but with (unbounded) Pareto-distributed firm productivities (with shape parameter γ).
- In Chaney (2008) the elasticity of trade flows with respect to variable trade costs (proxies for by distance here, if we assume $\tau_{ij} = \theta D_{ij}^{\delta}$ where $D =$ distance) can be subdivided into the:
  - Extensive elasticity: $\varepsilon_{D_{ij}}^{EXT_j} = -\delta\left[\gamma - (\sigma - 1)\right]$. CK estimate this by regressing firm-level entry (ie a Probit) on firm-level distance D_{ij} and a firm fixed effect. This is analogous to HMR's first stage.
  - Intensive elasticity: $\varepsilon_{D_{ij}}^{INT_j} = -\delta(\sigma - 1)$. CK estimate this by regressing firm-level exports on firm-level distance D_{ij} and a firm fixed effect. This is analogous to HMR's second stage.

CK (2010): The model and estimation I

- Recall that γ is the Pareto parameter governing firm heterogeneity.
- The above two equations (HMR's first and second stage) don't separately identify δ, σ and γ.
  - So to identify the model, CK bring in another equation which is the slope of the firm size (sales) distribution.
  - In the Chaney (2008) model this will behave as: $X_i = \lambda c_i^{-[\gamma-(\sigma-1)]}$, where c_i is a firm's marginal cost and X_i is a firm's total sales.
  - With an Olley and Pakes (1996) TFP estimate of $1/c_i$, CK estimate $[\gamma - (\sigma - 1)]$ and hence identify the entire system of 3 unknowns.

CK (2010): Results (each industry separately)

Table 3: The structural parameters of the gravity equation (Firm-level estimations)

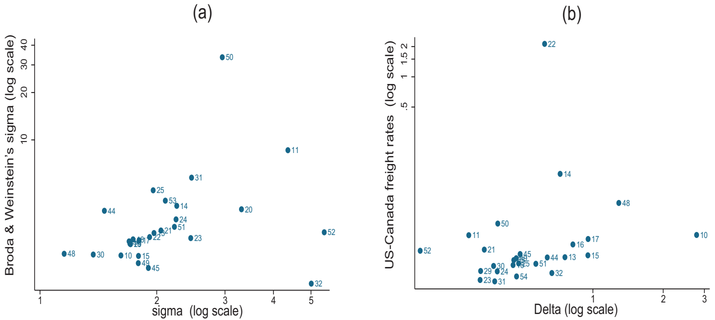
Industry	Code	(1) $P[Export > 0]$ $-\delta\gamma$	(2) Export value $-\delta(\sigma - 1)$	(3) Pareto$^{\#}$ $-[\gamma - (\sigma - 1)]$	(4) γ	(5) σ	(6) δ
Iron and steel	10	-5.51^a	-1.71^a	-1.36	1.98	1.62	2.78
Steel processing	11	-1.5^a	-0.99^a	-1.74	5.1	4.36	0.29
Metallurgy	13	-2.14^a	-0.73^a	-1.85	2.82	1.97	0.76
Minerals	14	-2.98^a	-0.91^a	-2.86	4.11	2.25	0.72
Ceramic and building mat.	15	-2.63^a	-0.76^a	-1.97	2.76	1.79	0.95
Glass	16	-2.33^a	-0.58^a	-2.13	2.84	1.7	0.82
Chemicals	17	-1.81^a	-0.76^a	-1.09	1.89	1.8	0.95
Speciality chemicals	18	-0.97^a	-0.34^a	-1.39	2.13	1.74	0.46
Pharmaceuticals	19	-1.19^a	-0.14	-1.4			
Foundry	20	-1.72^a	-0.85^a	-2.37	4.68	3.31	0.37
Metal work	21	-1.19^a	-0.36^a	-2.43	3.48	2.05	0.34
Agricultural machines	22	-2.06^a	-0.57^a	-2.39	3.31	1.92	0.62
Machine tools	23	-1.29^a	-0.48^a	-2.47	3.92	2.45	0.33
Industrial equipment	24	-1.25^a	-0.48^a	-1.97	3.21	2.24	0.39
Mining/civil egnring eqpmt	25	-1.37^a	-0.46^a	-1.9	2.86	1.96	0.48
Office equipment	27	-0.52^a	-1.02	-1.57			
Electrical equipment	28	-0.8^a	-0.14	-2.34			
Electronical equipment	29	-0.77^a	-0.24^a	-1.63	2.34	1.71	0.33
Domestic equipment	30	-0.94^a	-0.14^a	-2.13	2.51	1.37	0.38
Transport equipment	31	-1.4^a	-0.55^a	-2.23	3.69	2.46	0.38
Ship building	32	-3.69^a	-2.67^a	-1.52	5.53	5.01	0.67
Aeronautical building	33	-0.78^a	-0.13	-3.27			
Precision instruments	34	-1.07^a	0.08	-1.63			
Textile	44	-1.17^a	-0.3^a	-1.37	1.84	1.47	0.64
Leather products	45	-1.24^a	-0.44^a	-1.63	2.53	1.9	0.49
Shoe industry	46	-0.42^a	-0.29^a	-2.3	7.31	6.01	0.06
Garment industry	47	-0.33^a	0.13	-1.04			
Mechanical woodwork	48	-2.14^a	-0.2^a	-1.5	1.65	1.15	1.29
Furniture	49	-1.43^a	-0.37^a	-2.25	3.04	1.79	0.47
Paper & Cardboard	50	-1.45^a	-0.76^a	-1.76	3.71	2.95	0.39
Printing and editing	51	-1.4^a	-0.7^a	-1.24	2.46	2.22	0.57
Rubber	52	-1.26^a	-0.8^a	-2.52	6.93	5.41	0.18
Plastic processing	53	-1.24^a	-0.51^a	-1.6	2.7	2.11	0.46
Miscellaneous	54	-0.91^a	-0.33^a	-1.22	1.92	1.7	0.47
Trade-weighted mean		**-1.41**	**-0.53**	**-1.86**	**3.09**	**2.25**	**0.58**

^{a}, ^{b} and ^{c} denote significance at the 1%, 5% and 10% level respectively. $^{\#}$: All coefficients in this

CK (2010): Results (do the parameters make sense?)

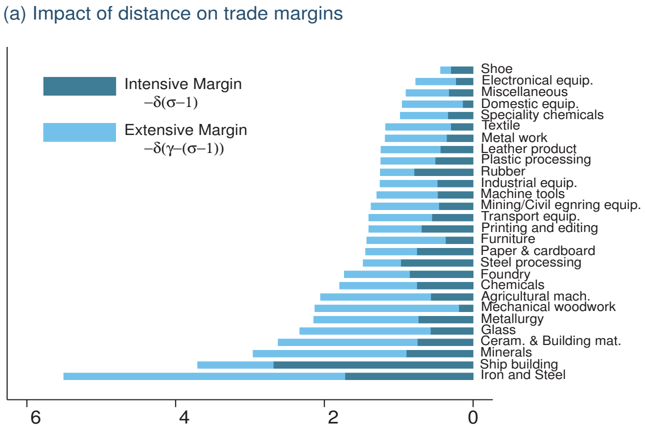
NB: the "Broda and Weinstein (2003)" here is now Broda and Weinstein (QJE, 2006)

Figure 3: Comparison of our results for σ and δ with those of Broda and Weinstein (2003)



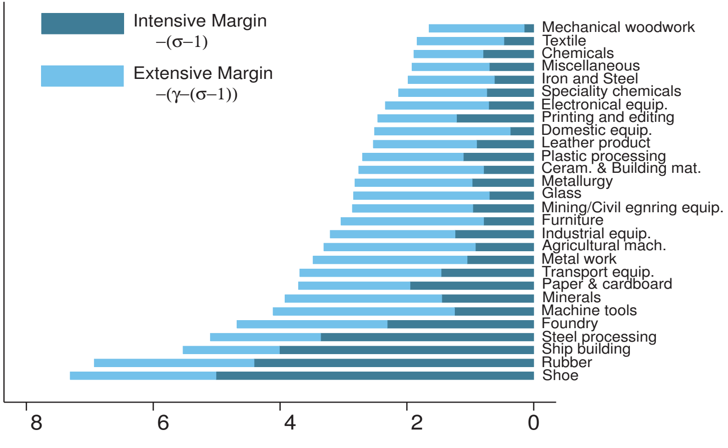
CK (2010): Results (what do the parameters imply about the two margins?)

Figure 4: The estimated impact of trade barriers and distance on trade margins, by industry



CK (2010): Results (what do the parameters imply about the two margins?)

(b) Impact of a tariff on trade margins



Plan for Today's Lecture

1. Stylized facts on extensive margin in trading behavior
2. Implications of extensive margin for gravity estimation:
    1. Helpman, Melitz and Rubinstein (2008)
    2. Crozet and Konig (2010)
3. **Path-dependent sunk costs of exporting: Morales, Sheu and Zahler (2017)**
4. Uncertainty and export behavior: Dickstein and Morales (2016)

Morales, Sheu and Zahler (2017): Introduction

- Exporting firms continuously enter and exit foreign markets.
- Descriptive evidence shows that firms are more likely to enter countries similar to their prior export destinations.
  - Lawless (2009, 2013); Albornoz et al. (2012); Chaney (2014); Defever et al. (2015); Meinen (2015).
- MSZ ask whether this cross-country correlation in firms' entry patterns is due to sunk entry costs in a market being smaller for firms that have previously exported to similar markets.
- They denote this *path dependence* in entry costs as "*extended gravity*".

MSZ (2017): Moment Inequalities

- The standard approach to the estimation of entry models relies on deriving choice probabilities from the theoretical framework and finding the parameter values that maximize the likelihood of the entry choices observed in the data (e.g. Das, Roberts, Tybout, ECMA, 2007).
- This approach is not feasible in MSZ's case. Evaluating these probabilities involves examining the dynamic implications of every possible choice a firm may make; i.e. of every possible bundle of export destinations.
- Given the cardinality of the potential choice set (for a given number of countries N, this set includes 2^N elements), computing the value function for each of its elements is infeasible unless very strong simplifying assumptions are imposed on the firm's actual choice set and state vector.
- As an example, even if the firm's actual choice set were to include only 20 destinations and the expected profits of exporting to each them were to depend only on one state variable that can take only 5 values, the state vector would still include $5^{20} \approx 10^{13}$ distinct elements.

MSZ (2017): Data

- Covers the period 1995-2005 and comes from two separate sources.
  1. Chilean customs database: exports for every firm, country, and year.
  2. Chilean Annual Industrial Survey: domestic sales, employment, value added, for all firms.
- Sector: chemicals, among the top two Chilean manufacturing sectors by volume of exports in every sample year.
- Per-year average number of firm-country pairs with positive exports is approximately 650; out of which 150 correspond to entering firms, and 125 correspond to exiting firms.
- Export events are generated by a per-year average of 110 firms exporting to around 70 countries in total.

MSZ (2017): Data

- Gravity variables relate Chile to each destination.
- Dummy variables that equal one if these destinations do *not* share Chile's border, continent, language, or similar income per capita: "Grav. Border", "Grav. Cont.", "Grav. Lang.", and "Grav. GDPpc".
- *Extended gravity* variables relate each potential destination to a firm's *prior* export bundle.
- Dummies for sharing border, continent, language, or similar income per capita with at least one country the firm exported to in the previous year, and not with Chile itself: "Ext. Grav. Border", "Ext. Grav. Cont.", "Ext. Grav. Lang.", and "Ext. Grav. GDPpc". For example:

$$
\begin{aligned}
(\text{Ext. Grav. Lang})_{ijt} = {} & (1 - d_{ijt-1}) \times \\
& \sum_{j'=1}^{J} \{ d_{ij't-1} \times \text{language}(j, j') \times (1 - \text{language}(h, j')) \},
\end{aligned}
$$

Demand, Variable Trade Costs, and Market Structure

- Firm i faces an isoelastic demand function in every market j: $q_{ijt} = p_{ijt}^{-\eta} P_{jt}^{\eta-1} Y_{jt}$.
- Marginal production cost is constant: $a_{it} w_t$.
- Marginal cost of selling in a foreign destination is constant: $\tau_{ijt} a_{it} w_t$.
- Exporters behave as monopolistically competitive firms in every market: they set their price p_{ijt} optimally taking P_{jt} as given.
- Given these assumptions on demand, variable trade and production costs, and market structure, the revenue a firm i obtains if it exports to a market j is

$$r_{ijt} \equiv p_{ijt} q_{ijt} = \left[\frac{\eta}{\eta - 1} \frac{\tau_{ijt} a_{it} w_t}{P_{jt}} \right]^{1-\eta} Y_{jt},$$

and the potential export profits gross of fixed and sunk costs are

$$r_{ijt} - \tau_{ijt} a_{it} w_t q_{ijt} = \eta^{-1} r_{ijt}.$$

Demand, Variable Trade Costs, and Market Structure

- Model the impact of variable trade costs τ_{ijt} on export revenues as

$$(\tau_{ijt})^{1-\eta} = \tau_{ijt}^{o} + \varepsilon_{ijt}^{\tau},$$

with τ_{ijt}^{o} a function of observed covariates and parameters and ε_{ijt}^{τ} unobserved:

$$\tau_{ijt}^{o} = \tau_{jt}\tau_{i}(a_{it})^{\xi_a}(d_{ijt-1})^{\xi_d}(X_{ijt}^{e})^{\xi_e},$$

with

$$X_{ijt}^{e} \equiv \big((\text{Ext. Grav. Border})_{ijt}, (\text{Ext. Grav. Cont.})_{ijt}, (\text{Ext. Grav. Lang.})_{ijt}, (\text{Ext. Grav. GDPpc})_{ijt}\big),$$

and

$$\mathbb{E}_{jt}[\varepsilon_{ijt}^{\tau} | a_{it}, d_{ijt-1}, X_{ijt}^{e}, \mathcal{B}_{it}, \mathcal{J}_{it}] = 0$$

Demand, Variable Trade Costs, and Market Structure

- Previous equations imply that we can rewrite potential export revenues as

$$r_{ijt} = r^o_{ijt} + \varepsilon^R_{ijt},$$

where r^o_{ijt} is a function of a vector of observed covariates and parameters,

$$r^o_{ijt} = \exp(\alpha_{jt} + \alpha_i + X^r_{ijt}\alpha^r), \qquad X^r_{ijt} = (d_{ijt-1}, X^e_{ijt}, \ln(r_{iht})),$$

and the unobserved component ε^R_{ijt} verifies

$$\mathbb{E}_{jt}[\varepsilon^R_{ijt} | X^r_{ijt}, \mathcal{B}_{it}, \mathcal{J}_{it}] = 0,$$

and, thus, ε^R_{ijt} is unobserved to the firm when deciding on its set of export destinations.

Fixed Export Costs

- Export fixed costs: they are independent of both the firm's previous export history and how much it sells to each destination.
- Model fixed export costs as

$$f_{ijt} = f_j^o + u_{ic_j t} + \varepsilon_{ijt}^F.$$

- The observable part of fixed costs, f_j^o, is modeled as

$$f_j^o = \gamma_0^F + \gamma_c^F(\text{Grav. Cont.})_j + \gamma_l^F(\text{Grav. Lang.})_j + \gamma_g^F(\text{Grav. GDPpc})_j.$$

- NB: impose no assumption on the distribution of $u_{ic_j t}$.
- Assume:

$$\mathbb{E}[\varepsilon_{ijt}^F | \mathcal{B}_{it}, \mathcal{J}_{it}] = 0;$$

thus, ε_{ijt}^F is unobserved to the firm when deciding on its export destinations.

Sunk Export Costs

- Sunk export costs: independent of the quantity exported to a destination (as with the fixed costs) and a firm only has to pay them if it was not exporting to this destination in the previous year (contrary to the fixed costs).
- Model sunk costs as

$$s_{ijt} = s_j^o - e_{ijt}^o + \varepsilon_{ijt}^S \qquad \text{with} \qquad \mathbb{E}[\varepsilon_{ijt}^S | \mathcal{B}_{it}, \mathcal{J}_{it}] = 0.$$

- Model the gravity term analogously to that in fixed export costs

$$s_j^o = \gamma_0^S + \gamma_c^S (\text{Grav. Cont.})_j + \gamma_l^S (\text{Grav. Lang.})_j + \gamma_g^S (\text{Grav. GDPpc})_j,$$

and the extended gravity term as

$$\begin{aligned} e_{ijt}^o = {} & \gamma_b^E (\text{Ext. Grav. Border})_{ijt} + \gamma_c^E (\text{Ext. Grav. Cont.})_{ijt} \\ & + \gamma_l^E (\text{Ext. Grav. Lang.})_{ijt} + \gamma_g^E (\text{Ext. Grav. GDPpc})_{ijt}. \end{aligned}$$

Net Export Profits

- The net static export profits of a destination j therefore are:

$$\pi_{ijt} = \eta^{-1} r_{ijt} - f_{ijt} - (1 - d_{ijt-1}) s_{ijt}.$$

- Aggregating across countries, the static profits of an export bundle b are:

$$\pi_{ibt} = \sum_{j \in b} \pi_{ijt}.$$

- Conditional on exporting to a bundle b in year t, the L_{it} periods ahead discounted sum of profits along the optimal path is

$$\Pi_{ibt,L_{it}} = \pi_{ibt} + \sum_{l=1}^{L_{it}} \delta^{l} \pi_{io_{it+l}(b)t+l},$$

where $o_{it+l}(b)$ is the optimal export bundle that firm i will choose at $t+l$ conditional on having exported to the bundle b at t.

Net Export Profits

- Let b denote a generic bundle of countries that a firm may export to.
- And let o_{it} denote the export bundle actually chosen by firm i in year t:

$$o_{it} = (d_{i1t}, \ldots, d_{ijt}, \ldots, d_{iJt}).$$

- Assume that firms choose export bundles optimally:

$$o_{it} = \underset{b \in \mathcal{B}_{it}}{\operatorname{argmax}} \, \mathbb{E}\left[\Pi_{ibt, L_{it}} | \mathcal{J}_{it}\right],$$

  where $\mathbb{E}[\cdot]$ is the expectation consistent with the data generating process.
- Therefore, o_{it} depends on three elements:
  - the discounted sum of profits, $\Pi_{ibt, L_{it}}$;
  - the consideration set or actual choice set, $\mathcal{B}_{it}$;
  - the information set about export profits in each $b \in \mathcal{B}_{it}$, $\mathcal{J}_{it}$.

Planning Horizon

- MSV (2017) only assume that

$$L_{it} \geq 1.$$

- Thus, they impose only weak restrictions on how forward-looking firms are when deciding its set of export destinations
- Compatible with firms taking into account the effect of their current choices on future profits in any of the three following ways:
  - only one period ahead, $L_{it} = 1$;
  - any finite number $p \geq 2$ of periods ahead, $L_{it} = p$;
  - an infinite number of periods ahead, $L_{it} = \infty$.
- NB: the planning horizon may be heterogeneous over time and across firms:

$$L_{it} \text{ may be different from } L_{i't'} \text{ for } i \neq i' \text{ or } t \neq t'.$$

Information Set

- MSV assume firms' expectations are rational, but leave (relatively) unrestricted the content of their information sets. Specifically, only assume that:

$$Z_{it} \subseteq \mathcal{J}_{it}, \qquad \text{where } Z_{it} \text{ is a vector of observed covariates.}$$

- Thus only impose a minimal content requirement on potential exporters' information sets. Specifically, in estimation, assume

$$Z_{it} = (Z_{ijt}, j = 1, \ldots, J),$$
$$Z_{ijt} = (f_j^o, s_j^o, e_{ijt}^o, d_{ijt-1}).$$

- Given the assumptions on the determinants of $(f_j^o, s_j^o, e_{ijt}^o)$, MSV only require firms to know whether each foreign country shares continent, language or similar income per capita with Chile, or with at least one country to which they exported in the previous year.

Consideration Set

- The potential choice set of the firm includes all possible combinations of countries. Given J possible countries, the cardinality of this set is 2^J: it is probably unrealistic to assume that firms evaluate the trade-offs, as captured in $\Pi_{ibt,L_{it}}$, of exporting to each of these potential bundles of countries.
- So MSV impose only a minimum content requirement on $\mathcal{B}_{it}$:

$$\mathcal{A}_{it} \subset \mathcal{B}_{it}, \text{ where } \mathcal{A}_{it} \text{ is known to the researcher.}$$

- Specifically, in estimation, assume

$$\mathcal{A}_{it} = \{o_{it}\} \cup \{o_{it}^{j \to j'}, \forall j' = 1, \ldots, J \text{ such that } j' \in \mathcal{A}_{ijt}\},$$
$$\mathcal{A}_{ijt} = \{j' = 1, \ldots, J \text{ such that } f_j^o = f_{j'}^o \text{ and } u_{ic_j t} = u_{ic_{j'} t}\},$$

where $o_{it}^{j \to j'}$ is the bundle generated by swapping destination j by j'.

Parameters to Estimate

- The unknown parameters of the model are:
  - the demand elasticity, η;
  - the discount factor, δ;
  - the export revenue parameters,

$$\alpha \equiv (\{\alpha_{jt}\}_{j,t}, \{\alpha_i\}_i, \alpha^r);$$

  - the fixed and sunk costs parameters,

$$\gamma \equiv (\gamma_0^F, \gamma_c^F, \gamma_l^F, \gamma_g^F, \gamma_0^S, \gamma_c^S, \gamma_l^S, \gamma_g^S, \gamma_b^E, \gamma_c^E, \gamma_l^E, \gamma_g^E);$$

  - the planning horizon, L_{it};
  - the information set, $\mathcal{J}_{it}$;
  - the consideration set, $\mathcal{B}_{it}$;
  - the distributions of the unobserved terms, u and $(\varepsilon^R, \varepsilon^F, \varepsilon^S)$;

Parameters of Interest

- MSV's parameter vector of interest, however, is:

$$\kappa \equiv (\kappa_b, \kappa_c, \kappa_l, \kappa_g) \equiv \left(\frac{\gamma_b^E}{\gamma_{all}^S}, \frac{\gamma_c^E}{\gamma_{all}^S}, \frac{\gamma_l^E}{\gamma_{all}^S}, \frac{\gamma_g^E}{\gamma_{all}^S} \right), \quad \gamma_{all}^S \equiv \gamma_0^S + \gamma_c^S + \gamma_l^S + \gamma_g^S.$$

- For a firm entering Germany, the relative reduction in sunk export costs is:
  - κ_g for a firm previously exporting to the United States,
  - κ_c for a firm previously exporting to Romania;
  - $\kappa_c + \kappa_g$ for a previous exporter to the United Kingdom;
  - $\kappa_b + \kappa_c + \kappa_g$ for a previous exporter to France;
  - $\kappa_b + \kappa_c + \kappa_l + \kappa_g$ for a prior exporter to Austria.
- Advantages of focusing only on estimating κ (instead of γ):
  - computational feasibility: not possible to identify confidence set for γ;
  - weaker assumptions are needed to derive bounds on κ than on γ; e.g. do not need to impose any normalization by scale.

One-period Deviations

- MSV apply an analogue of Euler's perturbation method.
- They compare the stream of profits along an observed sequence of bundles

$$o_{i1}^{T} = \{o_{i1} \ldots, o_{it-1}, o_{it}, o_{it+1}, \ldots, o_{iT}\},$$

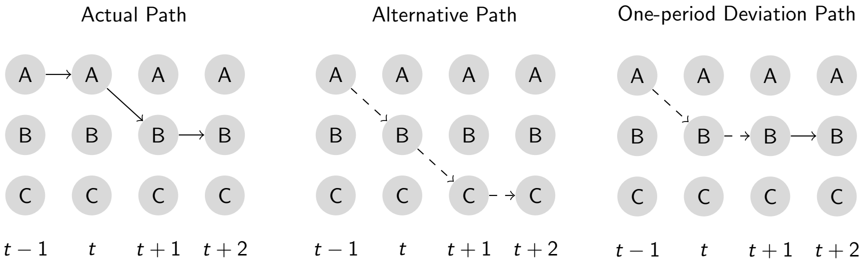
with the analogous stream along alternative sequences that differ from the observed one in only one period

$$\{o_{i1}, \ldots, o_{it-1}, o_{it}^{j \to j'}, o_{it+1}, \ldots, o_{iT}\}.$$

- Given the above model, the difference in the *realized* export profits, $\Pi_{ibt,L_{it}}$ is

$$\underbrace{\pi_{ijt} - \pi_{ij't}}_{\pi_{ijj't}} + \underbrace{\delta \sum_{j''=1}^{J} d_{ij''t+1}(\pi_{ij''t+1} - \pi_{ij''t+1}^{j \to j'})}_{\pi_{ijj't+1}}.$$

Example of One-period Deviations



- Given the model, for **any** pair (j, j') such that $o_{it}^{j \to j'} \in \mathcal{B}_{it}$ and any Z_{it}

$$\mathbb{E}\left[\pi_{ijj't} + \delta\pi_{ijj't+1} \middle| d_{ijt}(1 - d_{ij't}) = 1, Z_{it}\right] \geq 0.$$

- Conditioning set does not depend on any future action of the firm. This is the key difference with Holmes (2011) and Illanes (2016): important so that MSV can allow for the vector of expectational errors $(\varepsilon_{ijt}^R, \varepsilon_{ijt}^F, \varepsilon_{ijt}^S)$.

From Conditional to Unconditional Moment Inequalities

- These inequalities have two properties that complicate their applicability.
- First, they condition on a particular pair of actual and counterfactual choices: the number of moment inequalities is larger than MSV's sample size.
  - See Menzel (2014), Chernozhukov et al. (2014), and Bugni et al. (2016) for many moment inequality estimators
- Second, they condition on particular values of the instrument vector Z_{it}.
  - See Andrews and Shi (2013), Chernozhukov et al. (2013), and Armstrong (2014, 2015) for conditional moment inequality estimators.
- These two characteristics imply that the sample analogue of the moments described in the prior slide will sum over either none or very few observations.
- To facilitate the computation of confidence sets for κ, MSV exploit the *many conditional* moment inequalities to derive a *small number* of *unconditional* moment inequalities.
- Trade off between tightness and computational ease.

From Conditional to Unconditional Moment Inequalities

- Given

$$\mathbb{E}\left[\pi_{ijj't} + \delta\pi_{ijj't+1} \middle| d_{ijt}(1 - d_{ij't}) = 1, Z_{it}\right] \geq 0,$$

then, for any

$$\Psi(Z_{ijt}, Z_{ij't}) \geq 0, \qquad \text{for any value of } Z_{it},$$

it holds that

$$\mathbb{E}\left[\sum_{j=1}^{J} \sum_{j' \in \mathcal{A}_{ijt}} \Psi(Z_{ijt}, Z_{ij't}) d_{ijt}(1 - d_{ij't})(\pi_{ijj't} + \delta\pi_{ijj't+1})\right] \geq 0.$$

- Without imposing additional restrictions on the function $\Psi(\cdot)$, this moment will depend on all elements of the vector γ and on the structural errors u_{ic_jt}.

From Conditional to Unconditional Moment Inequalities

- If we choose instrument functions $\Psi(\cdot) \geq 0$ such that

$$\Psi(Z_{ijt}, Z_{ij't}) = 0 \quad \text{if} \quad s_j^o \neq \gamma_{all}^S \text{ or } s_{j'}^o \neq \gamma_{all}^S,$$

and the set of counterfactual choices is such that

$$\mathcal{A}_{ijt} = \{j' = 1, \ldots, J \text{ such that } f_j^o = f_{j'}^o \text{ and } u_{ic_jt} = u_{ic_{j'}t}\},$$

then the moment

$$\mathbb{E}\left[\sum_{j=1}^{J} \sum_{j' \in \mathcal{A}_{ijt}} \Psi(Z_{ijt}, Z_{ij't}) d_{ijt}(1 - d_{ij't})(\pi_{ijj't} + \delta\pi_{ijj't+1})\right]$$

depends only on observed covariates and the parameters $(\alpha, \kappa, \eta\gamma_{all}^S, \gamma_{all}^S)$, and is homogeneous of degree one in γ_{all}^S.

Estimation

- Parameters of interest are the elements of the vector κ.
- MSV base estimation of κ on the moment inequalities

$$\{\mathbb{m}_k(\alpha, \kappa, \tilde{\eta}) \geq 0, k = 1, \ldots, K\}$$

where

$$\mathbb{m}_k(\alpha, \kappa, \tilde{\eta}) \equiv$$

$$\mathbb{E}\left[\sum_{j=1}^{J} \sum_{j' \in \mathcal{A}_{ijt}} \Psi_k(Z_{ijt}, Z_{ij't}) d_{ijt}(1 - d_{ij't})(1/\gamma_{all}^{S})(\pi_{ijj't} + \delta \pi_{ijj't+1})\right] \geq 0.$$

- (Note how each of the K inequalities we use for estimation is defined by a different instrument function $\Psi_k(\cdot)$.)

Instrument Functions

- The instrument functions MSV use are:

$$\Psi_1(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = d_{ij't-1} = 0, (\text{Ext. Grav. Border})_j = 1, (\text{Ext. Grav. Border})_{j'} = 0\},$$
$$\Psi_2(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = d_{ij't-1} = 0, (\text{Ext. Grav. Border})_j = 0, (\text{Ext. Grav. Border})_{j'} = 1\},$$
$$\Psi_3(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = d_{ij't-1} = 0, (\text{Ext. Grav. Cont.})_j = 1, (\text{Ext. Grav. Cont.})_{j'} = 0\},$$
$$\Psi_4(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = d_{ij't-1} = 0, (\text{Ext. Grav. Cont.})_j = 0, (\text{Ext. Grav. Cont.})_{j'} = 1\},$$
$$\Psi_5(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = d_{ij't-1} = 0, (\text{Ext. Grav. Lang.})_j = 1, (\text{Ext. Grav. Lang.})_{j'} = 0\},$$
$$\Psi_6(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = d_{ij't-1} = 0, (\text{Ext. Grav. Lang.})_j = 0, (\text{Ext. Grav. Lang.})_{j'} = 1\},$$
$$\Psi_7(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = d_{ij't-1} = 0, (\text{Ext. Grav. GDPpc})_j = 1, (\text{Ext. Grav. GDPpc})_{j'} = 0\},$$
$$\Psi_8(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = d_{ij't-1} = 0, (\text{Ext. Grav. GDPpc})_j = 0, (\text{Ext. Grav. GDPpc})_{j'} = 1\},$$
$$\Psi_9(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = 1, d_{ij't-1} = 0\},$$
$$\Psi_{10}(\cdot) = \mathbb{1}\{s_j^o = s_{j'}^o = \gamma_{all}^S, d_{ijt-1} = 0, d_{ij't-1} = 1\}.$$

- Each of these moments is chosen with the aim of defining one bound on one of the elements of the parameter vector $(\tilde{\eta}, \kappa)$.

Estimation

- MSV estimate the vector $(\alpha, \kappa, \tilde{\eta})$ in two steps.
- In the first step, use data on export revenues and moment equalities to obtain point estimates of α.

$$\mathbb{E}_{jt}[r^o_{ijt} - \exp(\alpha_{jt} + \alpha_i + X^r_{ijt}\alpha^r)|X^r_{ijt}, d_{ijt} = 1] = 0.$$

- In the second step, use these estimates of α and the ten moment inequalities defined in the previous slide to obtain confidence sets for $(\kappa, \tilde{\eta})$.
- This two-step estimator is preferred over an alternative approach that uses only moment inequalities to estimate $(\alpha, \kappa, \tilde{\eta})$. Some reasons:
  1. Use different sources of variation to identify α and κ: information on observed export revenues conditional on foreign market participation to identify α, and information on foreign market entry and exit to identify κ.
  2. α is point identified (instead of set identified).
  3. No need to restrict the dimensionality of α for computational reasons.

Results

- Baseline results in MSV are based on the assumption that $u_{ic_jt} = u_{it}$, for all $j \in J$.
- Therefore, the only restriction they impose on the set of actual and counterfactual destinations used to form inequalities is that $s_j^o = s_{j'}^o = \gamma_{all}^S$.
- Treat u_{it} in our analysis as a firm- and year-specific fixed effect.
- Confidence sets on κ depend on a preliminary estimate α, which is used to generate a proxy for the potential export revenue that each firm may obtain in each country and year, $\hat{r}_{ijt}$.
- Compute estimates for many different specifications of the revenue estimating equation. However, the predicted export revenues $\hat{r}_{ijt}$ are reassuringly similar across the different specifications. Moment inequality estimates of κ are thus robust to the specific details of the regression used to generate our proxy for the potential export revenues.

Results

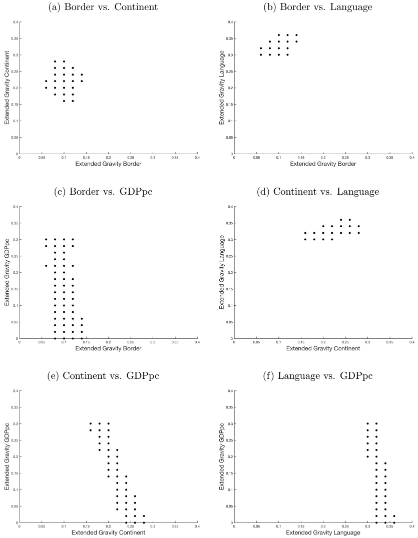
Table 4: Bounds on Individual Extended Gravity Parameters

Border	Continent	Language	GDPpc
[5.71%, 13.33%]	[19.05%, 28.57%]	[28.57%, 36.19%]	[0%, 28.57%]

Notes: This table reports bounds on the vector κ defined in equation (21). It uses the regression results described in column I of Table B.1. The confidence intervals are projections of a confidence set for $(\kappa, \tilde{\eta})$ computed following the procedure in Section 10.2 of Andrews and Soares (2010).

Projected Confidence Sets

Figure 4: Projected Confidence Set



Notes: These confidence sets are two-dimensional projections of a 5-dimensional confidence set for $(\kappa, \tilde{\eta})$ computed following the procedure in Section 10.2 of Andrews and Soares (2010).

Plan for Today's Lecture

1. Stylized facts on extensive margin in trading behavior
2. Implications of extensive margin for gravity estimation:
    1. Helpman, Melitz and Rubinstein (2008)
    2. Crozet and Konig (2010)
3. Path-dependent sunk costs of exporting: Morales, Sheu and Zahler (2017)
4. **Uncertainty and export behavior: Dickstein and Morales (2016)**

Exporting Under Uncertainty

- When taking decision to export, firms presumably face considerable uncertainty.
- Firms enter a market if the expected value of exporting is positive. So firms' choices will depend on expectations about the future evolution of their own productivity, trade policy, political stability in foreign countries, etc.
- Literature on exporting decisions (e.g. Roberts and Tybout (1997), Das et al. (2007), Arkolakis (2010), Cherkashin et al. (2010), Moxnes (2010), Eaton et al. (2011), Arkolakis (2013), Ruhl and Willis (2014)...) typically assumes that the researcher has **perfect knowledge** of the content of firms' information sets at the time of deciding whether to enter a foreign market.
- Dickstein and Morales (2016) develop partial identification tools for relaxing that assumption. Additionally, can test content of exporters' information sets. ("What do exporters know?")

Theoretical Model: Summary

- Single-agent partial equilibrium model.
- Two-periods
- First period:
  - firms decide whether to export to each possible destination market;
  - if they export, they must pay export fixed costs;
  - entry decision depends on expectations about potential export revenue upon entry and expectations are rational
  - **firms' information set used to predict export revenues is left unspecified**.
- Second period:
  - firms observe the realized demand in each destination and their realized marginal costs, determine their optimal price in each market in which they have decided to enter, and obtain export profits.
- No discounting between periods.

Demand

- Firm i faces an isoelastic demand in every country j,

$$x_{ijt} = \frac{p_{ijt}^{-\eta} Y_{jt}}{P_{jt}^{1-\eta}},$$

where p_{ijt} is the price set by i in j at t, Y_{jt} is the total expenditure, and P_{jt} is the ideal price index

$$P_{jt} = \left[\int_{i' \in A_{jt}} p_{i'jt}^{1-\eta} di \right]^{\frac{1}{1-\eta}},$$

where A_{jt} denotes the set of all firms selling in j at t.

- DM also show how one can allow for demand shifters that vary at the jt level and are common for all firms in the same country of origin.

Supply

- Every firm is the single world producer of its own variety.
- Market structure: monopolistic competition in every destination country.
- Constant marginal production cost: c_{it}.
- If a firm exports a positive amount to j, it must pay two additional costs:
  - iceberg trade costs τ_{jt};
  - fixed costs $f_{ijt} = \beta_0 + \beta_1 \boldsymbol{dist}_j + \nu_{ijt}$.
  - (in an extension, DM assume $f_{ijt} = \beta_j + \nu_{ijt}$ instead)
- In dynamic extension, allow also for sunk entry costs s_{ijt}.

Revenue Conditional on Exporting

- Conditional on entering destination market j, firm i obtains revenue

$$r_{ijt} = \left[\frac{\eta}{\eta - 1} \frac{\tau_{jt} c_{it}}{P_{jt}} \right]^{1-\eta} Y_{jt},$$

which may be rewritten as (think of "h" as denoting "home"):

$$r_{ijt} = \alpha_{jt} r_{iht},$$

with

$$\alpha_{jt} = \left(\frac{\tau_{jt}}{\tau_{ht}} \frac{P_{ht}}{P_{jt}} \right)^{1-\eta} \frac{Y_{jt}}{Y_{ht}}.$$

- The export profits that i would obtain in j if it were to export to t are

$$\pi_{ijt} = \eta^{-1} r_{ijt} - f_{ijt} = \eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j - \nu_{ijt}.$$

Information Set

- When deciding whether to export to market j in period t, firm i knows:
  - fixed costs of exporting f_{ijt};
  - an information set $\mathcal{J}_{ijt}$ that helps predict export revenues r_{ijt}.
- The set $\mathcal{J}_{ijt}$ includes any variable that helps predict demand in j,

$$(Y_{jt}, P_{jt}),$$

  or the marginal cost of selling to j, $\tau_{jt} c_{it}$.
- E.g. $\mathcal{J}_{ijt}$ may include lagged values of these variables, information on the number of competitors operating in market j (A_{jt}) or on their productivity levels ($p_{i'jt}$, for $i' \neq i$), tariffs, etc.
- Concerning the relationship between f_{ijt} and $\mathcal{J}_{ijt}$, DM assume that

$$\nu_{ijt} | (\mathcal{J}_{ijt}, dist_j) \sim \mathbb{N}(0, \sigma^2).$$

Decision to Export

- Assuming firms have rational expectations, firm i exports to j at t if

$$\mathbb{E}[\pi_{ijt}|\mathcal{J}_{ijt}, f_{ijt}] \geq 0.$$

- Therefore,

$$d_{ijt} = \mathbb{1}\{\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j - \nu_{ijt} \geq 0\},$$

  and the probability that i exports to j conditional on $(\mathcal{J}_{ijt}, dist_j)$ is

$$\mathcal{P}(d_{ijt} = 1|\mathcal{J}_{ijt}, dist_j) = \Phi\big(\sigma^{-1}\big(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j\big)\big).$$

- Not a simple probit because $\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}]$ is unobserved.
- Generally both r_{ijt} and $\mathcal{J}_{ijt}$ are unobserved.
- Parameters to estimate: $\theta^* \equiv (\beta_0, \beta_1, \sigma)$.
  - DM use $\eta = 5$ as a scale normalization.

Data Sources

- Data sources and variables used in the analysis:
  - Chilean customs database:
    - dummy for positive exports: d_{ijt},
    - firm-country specific exports: r_{ijt},
    - aggregate exports: $R_{jt} = \sum_i r_{ijt}$,
  - Chilean industrial survey:
    - domestic sales: $r_{iht} = r_{it} - \sum_j r_{ijt}$,
  - CEPII:
    - physical distance to Chile: $dist_j$.
- Unbalanced panel of firms for the sample period: 1995-2005.
- Focus on 2 (big) sectors:
  - manufacture of chemicals and chemical products;
  - food products.

Measurement of Export Revenues

- Only difficulty to estimate θ^* is that firms' expectations of revenue,

$$\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}],$$

  are unobserved.

- No matter what assumptions one imposes on the content of the information set $\mathcal{J}_{ijt}$, dealing with the unobserved firms' expectations requires the construction of a measure of potential export revenues r_{ijt}.
- Export revenue r_{ijt} is unobserved for firms that do not export, so r_{ijt} is not directly observed in the data for every firm, country and year.
- Given that $r_{ijt} = \alpha_{jt} r_{iht}$ and r_{iht} is observed for every firm and time period, obtaining a measure of r_{ijt} for every firm, country and year is equivalent to obtaining a measure of α_{jt} for every country and year.

Measurement of Export Revenues

- We denote observed export revenues as r_{ijt}^{obs} and, allowing for measurement error e_{ijt} in these observed export revenues, we can write

$$r_{ijt}^{obs} = d_{ijt}(r_{ijt} + e_{ijt}), \qquad \mathbb{E}_i[e_{ijt} | r_{iht}, d_{iht} = 1] = 0.$$

- The above model therefore predicts that

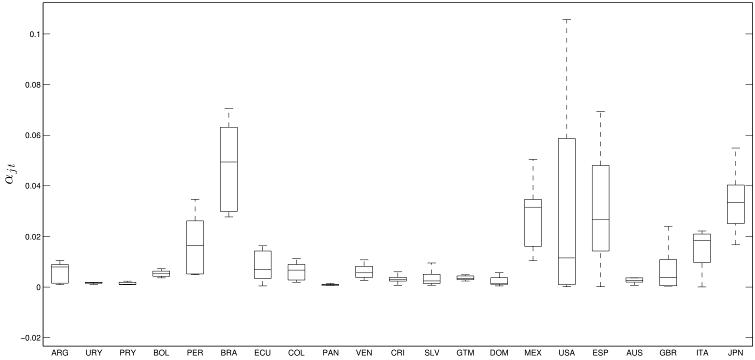
$$r_{ijt}^{obs} = d_{ijt}(\alpha_{jt} r_{iht} + e_{ijt})$$

and the following moment equality point identifies α_{jt} for every j and t

$$\mathbb{E}_i[r_{ijt}^{obs} - \alpha_{jt} r_{iht} | r_{iht}, d_{ijt} = 1] = 0.$$

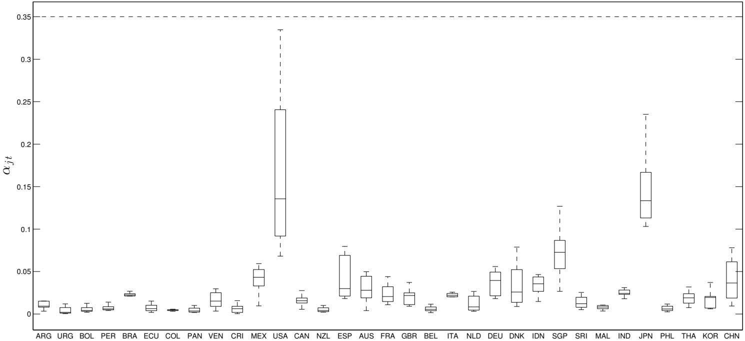
- Therefore, from now on, just treat $r_{ijt} \equiv \alpha_{jt} r_{iht}$ as known.

Market-year Revenue Shifters: Chemicals



- $\hat{\alpha}_{jt} < 0.1$ for all destinations and years.
- $\hat{\alpha}_{jt}$ relatively larger for Brasil, Mexico, Japan, Spain and US.

Country-year Revenue Shifters: Food



- $\hat{\alpha}_{jt} < 0.1$ for most destinations and years.
- $\hat{\alpha}_{jt}$ relatively larger for US, Japan, Singapore, Mexico, and Spain.

Assuming an Information Set

- The model above does not specify the content of the information set $\mathcal{J}_{ijt}$.
- These information sets are generally unobservable in standard datasets.
- Under the assumption that

$$\mathcal{J}^a_{ijt} = \mathcal{J}_{ijt},$$

the export probability conditional on $(\mathcal{J}^a_{ijt}, dist_j)$ is

$$\mathcal{P}(d_{ijt} = 1 | \mathcal{J}^a_{ijt}, dist_j) = \Phi\big(\sigma^{-1}\big(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}^a_{ijt}] - \beta_0 - \beta_1 dist_j\big)\big),$$

and one can estimate $\mathbb{E}[r_{ijt}|\mathcal{J}^a_{ijt}]$ non-parametrically and θ^* using ML.
- The ML estimator is asymptotically unbiased if and only if the researcher has a perfect proxy for firms' unobserved expectations; i.e.

$$\mathbb{E}[r_{ijt}|\mathcal{J}^a_{ijt}] = \mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}].$$

Bias when $\mathcal{J}_{ijt} \subset \mathcal{J}_{ijt}^a$

- Bias when the assumed information set $\mathcal{J}_{ijt}^a$ is too large; i.e. when the distribution of $\mathcal{J}_{ijt}$ conditional on $\mathcal{J}_{ijt}^a$ is degenerate.
- Perfect foresight is a special case: $\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}^a] = r_{ijt}$. In this case, the difference between the true agents' expectation and the researchers' proxy for it,

$$\xi_{ijt} \equiv \mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}^a] - \mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}],$$

  is equal to the expectational error that firms make when forecasting their export revenue upon entry,

$$\varepsilon_{ijt} \equiv r_{ijt} - \mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}].$$

- Assuming rational expectations implies $cov(\varepsilon_{ijt}, r_{ijt}) > 0$.
- Therefore, wrongly assuming perfect foresight is equivalent to introducing classical measurement error in firms' expectations: $cov(\varepsilon_{ijt}, \mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}^a]) > 0$.

Bias when $\mathcal{J}_{ijt} \subset \mathcal{J}^a_{ijt}$

- Understanding bias due to wrongly assuming perfect foresight is equivalent to understanding bias due to classical measurement error in probit models.
- Using proof in Yatchew and Griliches (1985), can show that: if agent's true expectations are normally distributed

  $$\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] \sim \mathrm{N}(0, \sigma_e^2),$$

  and the expectational error is normal conditional on agents' information sets

  $$\varepsilon_{ijt}|\mathcal{J}_{ijt} \sim \mathrm{N}(0, \sigma_\varepsilon^2),$$

  then there is an upward bias in the estimates of β_0, β_1 and σ.
- Bias increases in $\sigma_\varepsilon^2/\sigma_e^2$.
- Deviations from normality in the distribution of expectations or errors will alter the exact formula for the bias.

Bias when $\mathcal{J}^a_{ijt} \subset \mathcal{J}_{ijt}$

- Bias when the assumed information set $\mathcal{J}^a_{ijt}$ is too small; i.e. when the distribution of $\mathcal{J}^a_{ijt}$ conditional on $\mathcal{J}_{ijt}$ is degenerate.
- In this case, the difference between the true agents' expectation and the researcher's proxy for it is mean independent of the researcher's proxy.
- Therefore,

$$cov(\xi_{ijt}, \mathbb{E}[r_{ijt}|\mathcal{J}^a_{ijt}]) = 0.$$

- If the distribution of ν_{ijt} and that of $\nu_{ijt} + \eta^{-1}\xi_{ijt}$ differ only in their variance, then the ML estimates of β_0 and β_1 are consistent and only the estimates of σ are biased upwards. Example of this special case: when both ν_{ijt} and ξ_{ijt} are normally distributed.
- Parameters β_0 and β_1 will also be biased if the distributions of ν_{ijt} and $\nu_{ijt} + \eta^{-1}\xi_{ijt}$ differ in features other than their variance.

Empirical Application

- DM estimate θ^* under two alternative extreme assumptions on the content of firms' information sets.
- First, perfect foresight. Assume $\mathcal{J}^a_{ijt}$ such that

$$\mathbb{E}[r_{ijt}|\mathcal{J}^a_{ijt}] = r_{ijt}.$$

- Second, minimal information set. Assume $\mathcal{J}^a_{ijt}$ such that

$$\mathcal{J}^a_{ijt} \equiv (r_{iht-1}, R_{jt-1}, dist_j).$$

Average (Across Destinations) Fixed Export Costs

Table 3: Average fixed export costs

Estimator	Chemicals Argentina	Chemicals Japan	Chemicals United States	Food Argentina	Food Japan	Food United States
Perfect Foresight (MLE)	868.0 (51.7)	2,621.4 (159.4)	1,645.0 (97.6)	2,049.3 (87.2)	2,395.1 (103.9)	2,202.5 (93.5)
Minimal Information (MLE)	348.7 (12.9)	1,069.4 (40.9)	668.1 (24.2)	1,273.9 (43.1)	1,482.4 (50.3)	1,366.3 (45.5)
Moment Inequality	[79.1, 104.1]	[309.2, 420.5]	[181.3, 243.6]	[175.6, 270.1]	[269.1, 361.0]	[227.3, 308.9]

Notes: All parameters are reported in thousands of year 2000 USD and are conditional on the assumption that $\eta = 5$. For the two ML estimators, standard errors are reported in parentheses. For the moment inequality estimates, extreme points of the 95% confidence set are reported in square brackets. These confidence sets are projections of a confidence set for $(\beta_0, \beta_1, \sigma)$ computed according to the procedure described in Appendix A.5.

Exporters' Information Sets are Partially Observed

- It is generally hard to observe everything included in agents' information sets.
- But it seems feasible to observe a subset of their content.
- No matter which information sets potential exporters truly have, these might plausibly include:
  - lagged own domestic sales: r_{iht-1},
  - lagged aggregate exports from home country to each destination: R_{jt-1},
  - distance from home country: $dist_j$.
- Under the assumption that firms' information sets are only partially observed, the model parameters are only partially identified.
- DM combine two types of moment inequalities,
  - odds-based moment inequalities,
  - revealed-preference moment inequalities,

  that allow one to estimate the model parameters under the assumption that the researcher only observes part of firms' information sets.

Odds-Based Moment Inequalities

- Assume firm observes at least Z_{ijt}. That is, $Z_{ijt} \subseteq \mathcal{J}_{ijt}$.
- So then:

$$\mathcal{M}^{ob}(Z_{ijt};\theta^*) = \mathbb{E}\left[\begin{array}{c} m_l^{ob}(d_{ijt}, r_{ijt}, dist_j; \theta^*) \\ m_u^{ob}(d_{ijt}, r_{ijt}, dist_j; \theta^*) \end{array} \middle| Z_{ijt} \right] \geq 0,$$

with

$$m_l^{ob}(\cdot) = d_{ijt} \frac{1 - \Phi\big(\sigma^{-1}\big(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j\big)\big)}{\Phi\big(\sigma^{-1}\big(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j\big)\big)} - (1 - d_{ijt}),$$

$$m_u^{ob}(\cdot) = (1 - d_{ijt}) \frac{\Phi\big(\sigma^{-1}\big(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j\big)\big)}{1 - \Phi\big(\sigma^{-1}\big(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j\big)\big)} - d_{ijt}.$$

- Let $\mathcal{M}^{ob}(\cdot)$ denote the conditional **odds-based moment inequalities**.

Odds-Based Moment Inequalities - Proof for $m_l^{ob}(\cdot)$

- From model,

$$\mathbb{E}[\mathbb{1}\{\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j - \nu_{ijt} \geq 0\} - d_{ijt}|\mathcal{J}_{ijt}, dist_j] \geq 0.$$

- Given distributional assumption on ν_{ijt},

$$\mathbb{E}[\Phi(\sigma^{-1}(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j)) - d_{ijt}|\mathcal{J}_{ijt}, dist_j] \geq 0.$$

- Doing simple algebra,

$$\mathbb{E}[d_{ijt}\frac{1 - \Phi(\sigma^{-1}(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j))}{\Phi(\sigma^{-1}(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j))} - (1 - d_{ijt})|\mathcal{J}_{ijt}, dist_j] \geq 0.$$

Odds-Based Moment Inequalities - Proof for $m_l^{ob}(\cdot)$

- Given rational expectations assumption and convexity of $\Phi(\cdot)/(1-\Phi(\cdot))$,

$$\mathbb{E}[d_{ijt}\frac{1-\Phi(\sigma^{-1}(\eta^{-1}r_{ijt}-\beta_0-\beta_1 dist_j))}{\Phi(\sigma^{-1}(\eta^{-1}r_{ijt}-\beta_0-\beta_1 dist_j))}-(1-d_{ijt})|\mathcal{J}_{ijt}, dist_j] \geq 0.$$

- Given assumption that $Z_{ijt} \subseteq \mathcal{J}_{ijt}$ and Law of Iterated Expectations

$$\mathbb{E}[d_{ijt}\frac{1-\Phi(\sigma^{-1}(\eta^{-1}r_{ijt}-\beta_0-\beta_1 dist_j))}{\Phi(\sigma^{-1}(\eta^{-1}r_{ijt}-\beta_0-\beta_1 dist_j))}-(1-d_{ijt})|Z_{ijt}] \geq 0.$$

- Identical process for $m_u^{ob}(\cdot)$ but starting from

$$\mathbb{E}[d_{ijt}-\mathbb{1}\{\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}]-\beta_0-\beta_1 dist_j-\nu_{ijt} \geq 0\}|\mathcal{J}_{ijt}, dist_j] \geq 0.$$

Revealed-Preference Based Moment Inequalities

- If $Z_{ijt} \subseteq \mathcal{J}_{ijt}$, then

$$\mathcal{M}^r(Z_{ijt};\theta^*) = \mathbb{E}\left[\begin{array}{c} m_l^r(d_{ijt}, r_{ijt}, dist_j; \theta^*) \\ m_u^r(d_{ijt}, r_{ijt}, dist_j; \theta^*) \end{array} \middle| Z_{ijt}\right] \geq 0,$$

with

$$m_l^r(\cdot) = -(1-d_{ijt})\left(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j\right) + d_{ijt}\sigma \frac{\phi\left(\sigma^{-1}(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j)\right)}{\Phi\left(\sigma^{-1}(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j)\right)}$$

$$m_u^r(\cdot) = d_{ijt}\left(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j\right) + (1-d_{ijt})\sigma \frac{\phi\left(\sigma^{-1}(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j)\right)}{1-\Phi\left(\sigma^{-1}(\eta^{-1} r_{ijt} - \beta_0 - \beta_1 dist_j)\right)}.$$

- Let $\mathcal{M}^r(\cdot)$ denote the conditional **revealed-preference moment inequalities**.

Revealed-Preference Moment Inequalities - Proof for $m_u^r(\cdot)$

- From model,

$$\mathbb{E}[d_{ijt}(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j - \nu_{ijt})|\mathcal{J}_{ijt}, dist_j] \geq 0.$$

- Given distributional assumption on ν_{ijt},

$$\begin{aligned}&\mathbb{E}[d_{ijt}\big(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j\big)\\ &\quad + (1-d_{ijt})\sigma\frac{\phi\big(\sigma^{-1}(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j)\big)}{1-\Phi\big(\sigma^{-1}(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}] - \beta_0 - \beta_1 dist_j)\big)}|\mathcal{J}_{ijt}, dist_j] \geq 0.\end{aligned}$$

Revealed-Preference Moment Inequalities - Proof for $m_l^r(\cdot)$

- Given rational expectations assumption and convexity of $\phi(\cdot)/(1-\Phi(\cdot))$,

$$\mathbb{E}[d_{ijt}\big(\eta^{-1}\alpha_{jt}r_{iht}-\beta_0-\beta_1 dist_j\big) + (1-d_{ijt})\sigma\frac{\phi\big(\sigma^{-1}(\eta^{-1}\alpha_{jt}r_{iht}-\beta_0-\beta_1 dist_j)\big)}{1-\Phi\big(\sigma^{-1}(\eta^{-1}\alpha_{jt}r_{iht}-\beta_0-\beta_1 dist_j)\big)}|\mathcal{J}_{ijt}, dist_j]\geq 0.$$

- Given assumption that $Z_{ijt}\subseteq\mathcal{J}_{ijt}$ and Law of Iterated Expectations

$$\mathbb{E}[d_{ijt}\big(\eta^{-1}\alpha_{jt}r_{iht}-\beta_0-\beta_1 dist_j\big) + (1-d_{ijt})\sigma\frac{\phi\big(\sigma^{-1}(\eta^{-1}\alpha_{jt}r_{iht}-\beta_0-\beta_1 dist_j)\big)}{1-\Phi\big(\sigma^{-1}(\eta^{-1}\alpha_{jt}r_{iht}-\beta_0-\beta_1 dist_j)\big)}|Z_{ijt}]\geq 0.$$

- Identical process for $m_l^r(\cdot)$ but starting from

$$\mathbb{E}[(1-d_{ijt})(-(\eta^{-1}\mathbb{E}[r_{ijt}|\mathcal{J}_{ijt}]-\beta_0-\beta_1 dist_j-\nu_{ijt}))|\mathcal{J}_{ijt}, dist_j]\geq 0.$$

From Conditional to Unconditional Moments

- In order to apply these moment inequalities, need to derive unconditional moments that are consistent with the conditional moments:

$$\mathbb{E}\left[\left\{\begin{array}{c} m_l(d_{ijt}, r_{ijt}, dist_j; (\beta_0, \beta_1, \sigma)) \\ m_u(d_{ijt}, r_{ijt}, dist_j; (\beta_0, \beta_1, \sigma)) \\ m_l^r(d_{ijt}, r_{ijt}, dist_j; (\beta_0, \beta_1, \sigma)) \\ m_u^r(d_{ijt}, r_{ijt}, dist_j; (\beta_0, \beta_1, \sigma)) \end{array}\right\} \times g(Z_{ijt})\right] \geq 0,$$

  where $g(Z_{ijt})$ is the instrument function.
- Suggested $g(\cdot)$ functions in Andrews and Shi (2013) or Armstrong (2014) are computationally expensive here.
- For each $a \in \{0.5, 1, 1.5, 2\}$, DM estimate an identified set with

$$g_a(Z_{kjt}) = \left\{\begin{array}{c} \mathbb{1}\{Z_{kjt} > med(Z_{kjt})\} \\ \mathbb{1}\{Z_{kjt} \leq med(Z_{kjt})\} \end{array}\right\} \times (|Z_{kjt} - med(Z_{kjt})|)^a$$

Inference Procedure

- DM combine the odds-based and revealed-preference moment inequalities to compute a single confidence set.
  - They show that the resulting confidence set is tighter than that they'd get if they had used only the odds-based inequalities or only the revealed-preference inequalities.
- Inference is based on a finite set of unconditional moment inequalities. In spite of the loss of information this entails, the resulting confidence set is still sufficiently tight to yield economically meaningful results.
- To compute the confidence set for the true parameter θ^*, DM apply the procedure in Andrews and Soares (2010) with the Modified Method of Moments (MMM) test statistic.

Parameter Estimates

Table 2: Parameter estimates

	Chemicals			Food		
Estimator	σ	β_0	β_1	σ	β_0	β_1
Perfect Foresight (MLE)	1,038.6 (11.7)	745.2 (8.9)	1,087.8 (12.9)	1,578.1 (16.9)	2,025.1 (3.7)	214.5 (23.6)
Minimal Information (MLE)	395.5 (2.6)	298.3 (2.2)	447.1 (6.1)	959.9 (8.1)	1,259.3 (2.2)	129.4 (18.1)
Moment Inequality	[85.1, 117.6]	[62.8, 82.4]	[142.6, 197.1]	[114.9, 160.0]	[167.1, 264.0]	[36.4, 81.3]

Notes: All parameters are reported in thousands of year 2000 USD and are conditional on the assumption that $\eta = 5$. For the two ML estimators, standard errors are reported in parentheses. For the moment inequality estimates, extreme points of the 95% confidence set are reported in square brackets. These confidence sets are projections of a confidence set for $(\beta_0, \beta_1, \sigma)$ computed according to the procedure described in Appendix A.5.

Average (Across Destinations) Fixed Export Costs

Table 3: Average fixed export costs

Estimator	Chemicals Argentina	Chemicals Japan	Chemicals United States	Food Argentina	Food Japan	Food United States
Perfect Foresight (MLE)	868.0 (51.7)	2,621.4 (159.4)	1,645.0 (97.6)	2,049.3 (87.2)	2,395.1 (103.9)	2,202.5 (93.5)
Minimal Information (MLE)	348.7 (12.9)	1,069.4 (40.9)	668.1 (24.2)	1,273.9 (43.1)	1,482.4 (50.3)	1,366.3 (45.5)
Moment Inequality	[79.1, 104.1]	[309.2, 420.5]	[181.3, 243.6]	[175.6, 270.1]	[269.1, 361.0]	[227.3, 308.9]

Notes: All parameters are reported in thousands of year 2000 USD and are conditional on the assumption that $\eta = 5$. For the two ML estimators, standard errors are reported in parentheses. For the moment inequality estimates, extreme points of the 95% confidence set are reported in square brackets. These confidence sets are projections of a confidence set for $(\beta_0, \beta_1, \sigma)$ computed according to the procedure described in Appendix A.5.

Testing Content of Firms' Information Sets: "What do exporters know?"

- Given a set of unconditional moment inequalities, Bugni et al. (2015) describe a procedure to test the null hypothesis that there exists at least one value of the parameter vector θ consistent with all inequalities.
- This is a test of joint validity of all the inequalities used for identification.
- Each moment inequality is implied by: (a) the assumptions embedded in the theoretical model; (b) the assumption that the corresponding $Z_{ijt} \subseteq \mathcal{J}_{ijt}$.
- Therefore, rejecting the null could mean that either the theoretical model is inconsistent with the data or our vector of instruments Z_{ijt} is invalid. DM therefore show that, for some (simple) vector Z_{ijt} the model is not rejectedl.

Testing Content of Firms' Information Sets: "What do exporters know?"

Table 5: Testing Content of Information Sets

Set of Firms	Set of Export Destinations	Variable Tested	Chemicals Reject at 5%	Chemicals p-value RC	Food Reject at 5%	Food p-value RC
All	All	$(dist_j, r_{iht-1}, R_{jt-1})$	No	0.140	No	0.975
All	All	$(\alpha_{jt} r_{iht})$	Yes	0.005	Yes	0.005
Large	Popular	$(dist_j, r_{iht-1}, R_{jt-1}, \alpha_{jt-1})$	No	0.110	No	0.940
Large	Unpopular	$(dist_j, r_{iht-1}, R_{jt-1}, \alpha_{jt-1})$	No	0.110	No	0.970
Small	Popular	$(dist_j, r_{iht-1}, R_{jt-1}, \alpha_{jt-1})$	Yes	0.005	Yes	0.005
Small	Unpopular	$(dist_j, r_{iht-1}, R_{jt-1}, \alpha_{jt-1})$	Yes	0.020	Yes	0.005
Small & Exporter$_{t-1}$	All	$(dist_j, r_{iht-1}, R_{jt-1}, \alpha_{jt-1})$	Yes	0.005	Yes	0.005
Large & Non-exporter$_{t-1}$	All	$(dist_j, r_{iht-1}, R_{jt-1}, \alpha_{jt-1})$	No	0.145	No	0.990
Small & Non-Exporter$_{t-1}$	All	$(dist_j, r_{iht-1}, R_{jt-1}, \alpha_{jt-1})$	Yes	0.005	Yes	0.005
Large & Exporter$_{t-1}$	All	$(dist_j, r_{iht-1}, R_{jt-1}, \alpha_{jt-1})$	No	0.105	No	0.985

Notes: *Large* firms are those with above median domestic sales in the previous year. Conversely, firm i at period t is defined as *Small* if its domestic sales fall below the median. *Popular* export destinations are those with above median number of exporters in the previous year. We define a firm i at period t as *Exporter*$_{t-1}$ with respect to a country j if $d_{ijt-1} = 1$ and as a *Non-exporter*$_{t-1}$ if $d_{ijt-1} = 0$. For details on how to compute these p-values, see Bugni et al. (2015). All numbers reported in this table are independent of the value of η chosen as the normalizing constant.

What did we Learn about Exporters' Information Sets?

- We can reject that firms have perfect foresight.
- We cannot reject that potential exporters know, at least, their own lagged domestic sales, lagged Chilean aggregate exports and distance to each potential destination market.
- Large firms have relevant information about potential export revenues that small firms do not have.
- There is no evidence that:
  - firms have information about destination markets that have been popular in the past that they do not have for nonpopular markets;

    Learning from other exporters?
  - previous exporters have more information than previously non-exporting firms;

    Learning from previous export experience?

Counterfactuals

Table 6: Impact of 40% Reduction in Fixed Costs in Chemicals

	1996			2005		
Estimator	Argentina	Japan	United States	Argentina	Japan	United States
			% Change in Number of Exporters			
Perfect Foresight	52.6	663.7	201.1	51.6	632.7	201.9
Minimal Info.	54.9	486.2	125.6	53.5	755.1	135.8
Moment Inequality	[54.9, 64.5]	[135.7, 1796.7]	[433.1, 521.1]	[45.1, 56.6]	[0,1678.2]	[444.1, 534.6]
			Counterfactual Number of Exporters			
Perfect Foresight	67	38	51	70	37	72
Minimal Info.	68	29	38	71	43	56
Moment Inequality	[68, 72]	[12, 95]	[91, 106]	[68, 72]	[5, 89]	[131, 152]

Notes: For the moment inequality estimates, the minimum and maximum predicted values obtained by projecting the 95% confidence set for θ are reported in squared brackets. Counterfactual numbers of exporters are computed by rounding the outcome of multiplying the observed number of exporters by the counterfactual changes predicted by each of the three models. For the chemicals sector, observed number of exporters to Argentina, Japan and United States in 2005 are 46, 5 and 24, respectively. Analogous numbers for 1996 are 44, 5, 17. All numbers reported in this table are independent of the value of η chosen as normalizing constant.

Extension to Study Sunk Costs/Dynamics

Table 7: Export fixed and sunk costs: firm average

			Chemicals	
Estimator	Cost	Argentina	Japan	United States
Benchmark	Fixed	[79.1, 104.1]	[309.2, 420.5]	[181.3, 243.6]
Dynamics	Fixed	[55.8, 109.3]	[853.3, 1,670.0]	[409.2, 800.8]
	Sunk	[384,2, 734,3]	[5,874.4, 11,224.5]	[2,816.6, 5,382.7]
Selection	Fixed	[67.7, 135.1]	[1,033.9, 2,064.3]	[495.8, 989.9]

Notes: All variables are reported in thousands of year 2000 USD and are conditional on the assumption that $\eta = 5$. Extreme points of 95% confidence sets computed according to the procedure described in Appendix A.5 are reported in square brackets.