Identification of Discrete Choice Models Using Moment Inequalities: Theory and Application

Eduardo Morales

Columbia University & Princeton University

April 24, 2012

Topics

- Mapping between statistical and behavioral discrete choice models.
  - Pakes (2010); Pakes et al. (2011); Dickstein and Morales (2012).
- Applications.
  - Ho (2009), Holmes (2011), Morales et al. (2011), Ho and Pakes (2011), Crawford and Yurukoglu (2011), Ishii (2012).

MAPPING BETWEEN STATISTICAL AND BEHAVIORAL MODELS

Discrete Choice Problem

- Utility of agent i for alternative j is:

$$U_{ij} = \beta x_{ij}^* + \nu_{ij}, \qquad j = 1, \ldots, J, \quad x_{ij}^* \in \mathcal{X} \in R^2.$$

- Choice of individual i is captured in d_i and

$$d_{ij} = \mathbb{1}[U_{ij} \geq U_{ij'}, \text{ for } j' = 1 \ldots, J], \qquad d_i = d(x_i^*, \nu_i)$$

- We observe a random sample of vectors

$$(d_i, x_i, z_i) \sim \mathcal{P}, \qquad i = 1, \ldots, N,$$

with $x_{i1} = x_{i1}^* + \epsilon_{i1}^x$, $z_{i1} = x_{i1}^* + \epsilon_{i1}^z$, $x_{i2} = x_{i2}^*$.

- Assumptions on $(\nu_i, \epsilon_{i1}^x, \epsilon_{i1}^z)$:
  - structural error: $\nu_i \sim F_\nu(\nu_i | x_i^*, \epsilon_i^z, \epsilon_i^x) = F_\nu(\nu_i | x_i^*)$
  - measurement errors: $\mathbb{E}[\epsilon_i^x | x_i^*, \nu_i, \epsilon_i^z] = \mathbb{E}[\epsilon_i^x | x_i^*] = 0$
- Note: we omit subindex i from here onwards.

Example I: Estimation of Structural Parameters

Consumer Choice Problem

- Each individual maximizes a utility function:

$$U_j = \beta_1 \mathbb{E}[x_j | \mathcal{J}] + \beta_2 p_j + \nu_j, \qquad j = 1, \ldots, J.$$

- Instead of agent's i subjective expectations, the econometrician observes both the realized values,

$$x_j = \mathbb{E}[x_j | \mathcal{J}] + \epsilon_j^x, \; \mathbb{E}[\epsilon_j^x | \mathcal{J}] = 0,$$

  and an additional shifter of agent's i expectations:

$$z_j = \mathbb{E}[x_j | \mathcal{J}] + \epsilon_j^z.$$

- The expectational error corresponds to the classical measurement error in an explanatory variable (errors-in-variables).

Example II: Estimation of Structural Parameters

Two-period Entry Problem

- Each firm decides where to locate a plant:

$$U_j = \mathbb{E}[R_j|\mathcal{J}] + \beta_2 F_j + \nu_j, \qquad j = 1, \ldots, J.$$

- Instead of agent's i subjective expectations, the econometrician observes realized values:

$$R_j = \mathbb{E}[R_j|\mathcal{J}] + \epsilon_j^R.$$

  The econometrician also observes Z such that

$$Z_j = \mathbb{E}[R_j|\mathcal{J}] + \epsilon_j^Z,$$

- Similar to previous example but with structural restriction: $\beta_1 = 1$.
- Standard application of the moment inequalities estimator:
  - Ho (2009), Holmes (2011), Morales et al. (2011), Ishii (2012).

Example III: Estimation of Reduced Form Parameters

General Entry Problem

- Each firm decides where to locate a plant. Period t profits of locating plant in location j are:

$$\pi_{jt} = R_{jt} - \sum_{j'=1}^{J} F_{jj'} d_{ij't-1}$$

- The expected present value of locating plant in location j is:

$$U_{jt} = \mathbb{E}[(\pi_{jt} + \sum_{s=t+1}^{\infty} \delta^{s-t} d_{js}\pi_{js})|\mathcal{J}_t, d_{jt} = 1],$$
$$= V_{jt} - \sum_{j'=1}^{J} F_{jj'} d_{j't-1},$$

with

$$V_{jt} = R_{jt} + \mathbb{E}[(\sum_{s=t+1}^{\infty} \delta^{s-t} d_{js}\pi_{js})|\mathcal{J}_t, d_{jt} = 1].$$

Example III: Estimation of Reduced Form Parameters

General Entry Problem (cont.)

- Assume a projection of V_{jt} on a set of observable covariates:

$$V_{jt} = \beta x_{jt} + \nu_{jt}, \qquad (x_t, \nu_t) \in \mathcal{J}_{it}.$$

- Assume a structural form for $F_{jj'}$:

$$F_{jj'} = \theta \| L_j - L_{j'} \|,$$

  with $\| \cdot \|$ some measure of distance, and L_j an indicator of location j.

- Therefore:

$$U_{jt} = \beta x_{jt} + \theta \sum_{j'=1}^{J} \| L_j - L_{j'} \| d_{j't-1} + \nu_{jt}.$$

- Additionally, we can allow for measurement error in x_{jt}.

Maximum Likelihood Estimation

- Model 1:
  - Assume (up to a finite parameter vector) the distributions:

$$\{F_\nu(\nu|x^*), F_\epsilon(\epsilon|x^*), \mathcal{P}_{x^*}(x^*)\}$$

  - The individual i likelihood function is:

$$\begin{aligned}
\mathcal{L}(d|x,z) &= \mathbb{P}(d_j = 1|x,z) \\
&= \mathbb{E}(\mathbb{1}\{d_j = 1\}|x,z) \\
&= \mathbb{E}(\mathbb{E}(\mathbb{1}\{d_j = 1\}|x,z,x^*)|x,z) \\
&= \mathbb{E}(\mathbb{E}(\mathbb{1}\{d_j = 1\}|x^*)|x,z) \\
&= \int_x \left[\int_\nu \mathbb{1}\{U_j \geq \max_{j' \in J}\{U_{j'}\}\} dF_\nu(\nu|x^*)\right] dF_{x^*}(x^*|x,z),
\end{aligned}$$

    with

$$U_j = \beta x_j^* + \nu_j.$$

Maximum Likelihood Estimation

- Model 2:
  - Assume (up to a finite parameter vector) the distributions:

$$\{F_\nu(\nu|x^*), F_\epsilon(\epsilon^x|x^*), \mathcal{P}_{x^*}(x^*)\}$$

such that:

$$F_\nu(\nu|x^*) = F_\nu(\nu), \quad \text{and} \quad F_\epsilon(\epsilon|x) = F_\epsilon(\epsilon)$$

  - The individual i likelihood function is:

$$\begin{aligned}\mathcal{L}(d|x) &= \mathbb{P}(d_j = 1|x) \\ &= \mathbb{E}(\mathbb{1}\{d_j = 1\}|x) \\ &= \int_{\nu+\epsilon} \mathbb{1}\{U_j \geq \max_{j' \in J}\{U_{j'}\}\} dF_{\nu+\epsilon}(\nu + \epsilon^x),\end{aligned}$$

with

$$U_j = \beta x_j + \nu_j - \beta \epsilon_j^x.$$

Maximum Likelihood Estimation

- Model 3:
  - Assume (up to a finite parameter vector) the distribution

$$F_\nu(\nu|x^*),$$

    and assume

$$x = x^*.$$

  - The individual i likelihood function is:

$$\begin{aligned}\mathcal{L}(d|x) &= \mathbb{P}(d_j = 1|x)\\ &= \mathbb{E}(\mathbb{1}\{d_j = 1\}|x)\\ &= \mathbb{E}(\mathbb{1}\{d_j = 1\}|x^*)\\ &= \int_\nu \mathbb{1}\{U_j \geq \max_{j' \in J}\{U_{j'}\}\} dF_\nu(\nu|x^*),\end{aligned}$$

    with

$$U_j = \beta x_j^* + \nu_j.$$

Summary of MLE (similar in GMM)

- Dealing with structural and measurement error in MLE or GMM requires:
  - Assuming both the marginal distribution of the unobserved true covariates, and the distribution of both measurement and structural error conditional on these covariates (Model 1); or,
  - Assuming the structural error is independent of the true covariates, and the measurement error is independent of the *observed* covariates (Model 2):
    - Only in this case the difference between structural and measurement error is irrelevant!. We can think of a single error, η, such that:

      $$\eta = \nu + \epsilon$$

      and assume a single distribution $F_\eta(\eta)$; or,
  - Assuming the distribution of the structural error conditional on the true covariates, and assuming that these are measured without error (Model 3).

Moment Inequalities

Introduction

- Given our discrete choice problem, we can derive conditional moment inequalities. A conditional moment inequality is:

$$\mathbb{E}[m(d, x, z, j, j'; \beta_0)|x, z] \geq 0, \text{ where } d_j = 1, \text{ and } j' \neq j.$$

- For simplicity, in these slides we base identification on a set of unconditional moment inequalities:

$$\mathbb{E}[m_s(d, x, z, j, j'; \beta_0)] \geq 0, \quad s = 1, \ldots, S.$$

- Moment inequalities will generically lead to set identification. The identified set is:

$$\{\beta \in \mathcal{B} : \int \sum_{s \in S} \sum_{j \in J} \sum_{j' \neq j} (\min\{m_s(d, x, z, j, j'; \beta), 0\}^2) d\mathcal{P}(d, x, z) = 0\}$$

- We denote the identified set as $\mathcal{B}_{\mathcal{M}}(\mathcal{P})$.

Moment Inequalities

Deriving Moment Inequalities from our Discrete Choice Problem: Model 1

- Assumptions:

$$(\epsilon_j^x, \nu_j) = (0, 0), \quad \text{for every } j \in J.$$

- Moment inequalities.

$$m_s(d, x, z, j, j'; \beta_0) = \mathbb{1}\{d_j = 1\}\beta_0 \Delta x_{jj'} \geq 0,$$

- Normalization by scale: (1) $\beta_{01} = 1$; (2) $\|\beta_0\| = 1$.
- Completely deterministic model; very likely rejected by the data.
  - Contradictory inequalities derived from the model.

Moment Inequalities

Deriving Moment Inequalities from our Discrete Choice Problem: Model 2

- Assumptions:
  - No structural error: $\nu_j = 0$, for every $j \in J$;
  - Measurement error indep. of true covariate: $F_\epsilon[\epsilon_1^x | x^*] = F_\epsilon[\epsilon_1^x]$;
    - No parametric assumption on $F_\epsilon[\epsilon_1^x]$ needed.
  - Structural restriction: $\beta_{01} = 1$.
- Moment inequalities.

$$\mathbb{E}[\mathbb{1}\{d_j = 1\}\chi_s(\Delta x_{2jj'})\beta_0 \Delta x^*_{jj'}] \geq 0,$$
$$\mathbb{E}[\mathbb{1}\{d_j = 1\}\chi_s(\Delta x_{2jj'})(\beta_0 \Delta x_{jj'} - \beta_0 \Delta \epsilon^x_{1jj'})] \geq 0,$$
$$\mathbb{E}[\mathbb{1}\{d_j = 1\}\chi_s(\Delta x_{2jj'})\beta_0 \Delta x_{jj'}] \geq 0,$$

with

$$\chi_1(\Delta x_{2jj'}) = \mathbb{1}\{\Delta x_{2jj'} \geq 0\},$$
$$\chi_2(\Delta x_{2jj'}) = \mathbb{1}\{\Delta x_{2jj'} < 0\}.$$

- No need to normalize by scale.

Moment Inequalities

Deriving Moment Inequalities from our Discrete Choice Problem: Model 3

- Assumptions:
  - No structural error: $\nu_j = 0$, for every $j \in J$;
  - Measurement error indep. of true covariate: $F_\epsilon[\epsilon_1^x | x^*] = F_\epsilon[\epsilon_1^x]$
    - No parametric assumption on $F_\epsilon[\epsilon_1^x]$ needed.
  - Additional indicator of the covariate measured with error: z_1.
- Moment inequalities.

$$\mathbb{E}[\mathbb{1}\{d_j = 1\}\chi_s(\Delta z_{1jj'}, \Delta x_{2jj'})\beta_0 \Delta x^*_{jj'}] \geq 0,$$

$$\mathbb{E}[\mathbb{1}\{d_j = 1\}\chi_s(\Delta z_{1jj'}, \Delta x_{2jj'})(\beta_0 \Delta x_{jj'} - \beta_0 \Delta \epsilon^x_{1jj'})] \geq 0,$$

$$\mathbb{E}[\mathbb{1}\{d_j = 1\}\chi_s(\Delta z_{1jj'}, \Delta x_{2jj'})\beta_0 \Delta x_{jj'}] \geq 0,$$

with

$$\chi_1(\Delta z_{1jj'}, \Delta x_{2jj'}) = \mathbb{1}\{\Delta z_{1jj'} \geq 0\}\mathbb{1}\{\Delta x_{2jj'} \geq 0\},$$

$$\chi_2(\Delta z_{1jj'}, \Delta x_{2jj'}) = \mathbb{1}\{\Delta z_{1jj'} \geq 0\}\mathbb{1}\{\Delta x_{2jj'} < 0\},$$

$$\chi_3(\Delta z_{1jj'}, \Delta x_{2jj'}) = \mathbb{1}\{\Delta z_{1jj'} < 0\}\mathbb{1}\{\Delta x_{2jj'} \geq 0\},$$

$$\chi_4(\Delta z_{1jj'}, \Delta x_{2jj'}) = \mathbb{1}\{\Delta z_{1jj'} < 0\}\mathbb{1}\{\Delta x_{2jj'} < 0\}.$$

Moment Inequalities

Deriving Moment Inequalities from our Discrete Choice Problem: Model 4

- Assumptions:
  - Structural error organized in nests:

$$\nu_j - \nu_{j'} = \begin{cases} 0 & \text{if } G(j) = G(j'), \\ \mathbb{R} & \text{if } G(j) \neq G(j'), \end{cases}$$

    where $G(j)$ denotes a particular subset of J.
    - No parametric assumption on $F_\nu[\nu|x]$ needed.
  - No measurement error.
- Moment inequalities.

$$\mathbb{E}[\mathbb{1}\{d_j = 1\}\mathbb{1}\{G(j) = G(j')\}\chi_s(\Delta x_{jj'})(\beta_0 \Delta x_{jj'} + \Delta\nu_{jj'})] \geq 0,$$
$$\mathbb{E}[\mathbb{1}\{d_j = 1\}\mathbb{1}\{G(j) = G(j')\}\chi_s(\Delta x_{jj'})\beta_0 \Delta x_{jj'}] \geq 0.$$

- Trivial to allow for measurement error indep. of the true covariate.
  - impose $\beta_{01} = 1$ as a structural assumption (as in Model 2); or,
  - use an additional indicator z_1 (as in Model 3).

Moment Inequalities

Deriving Moment Inequalities from our Discrete Choice Problem: Additional Models

See Dickstein and Morales (2012) for guidance on how to build moment inequalities for the following models:

- Model 5:
  - Assume a parametric distribution for ν_i:
    - $F_{\nu}[\nu|x;\sigma]$
  - Allow for distribution free measurement error in two cases:
    - multiple indicator assumption;
    - fixed parameter on the variable measured with error.
- Model 6:
  - Assume ν_i is distributed independently of x.
    - No parametric assumption on $F_{\nu}[\nu|x]$ needed.
  - Allow for distribution free measurement error in two cases:
    - multiple indicator assumption;
    - fixed parameter on the variable measured with error.

Summary of Moment Inequalities

- In words, dealing with both structural and measurement error using moment inequalities requires:
  - In order to deal with measurement error, we may apply any of the usual IV solutions to measurement error problems in linear regression models. In particular, no parametric assumption on its distribution function is needed.
  - In order to deal with structural error, one needs to:
    - Assume it away; or,
    - Assume that it is an individual effect common to a subset of choices; or,
    - Assume a parametric distribution on it; or,
    - Assume that it is distributed independently of x (no parametric assumption needed).

APPLICATION OF MOMENT INEQUALITIES:

RATIONAL EXPECTATIONS SINGLE AGENT DYNAMIC DISCRETE CHOICE MODELS.

Gravity and Extended Gravity:
Estimating a Structural Model of Export Entry.

Eduardo Morales
Gloria Sheu
Andrés Zahler

Motivation

Firm 1 (Compañía Manufacturera de Aconcagua): standard gravity factors

Trade models based on gravity imply that firms should enter first bordering countries. . .

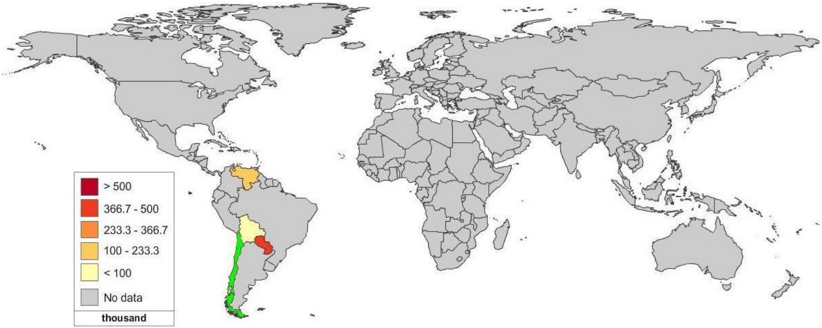


Figure: Year 1995

Motivation

Firm 1 (Compañía Manufacturera de Aconcagua): standard gravity factors

. . . or markets that are geographically close and share the same official language as the domestic country. . .

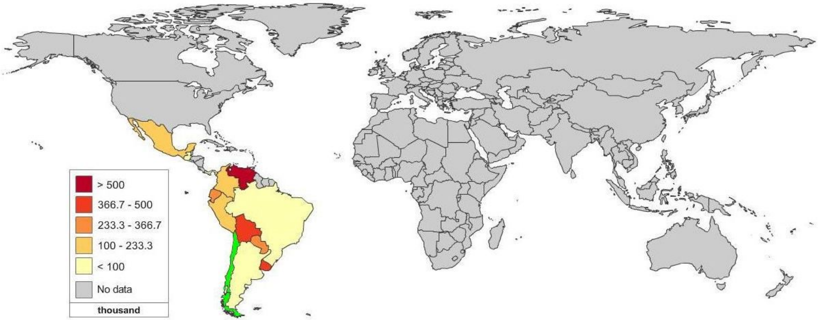


Figure: Year 2000

Motivation

Firm 1 (Compañía Manufacturera de Aconcagua): standard gravity factors

. . . and only later they overcome larger entry costs and access countries that are further away and in which different languages are spoken . . .

Figure: Year 2005

Motivation

Firm 2 (Industrias Cerecita SA): Extended gravity factors

. . . but, some firms, even though they follow this rule. . .

Figure: Year 1995

Motivation

Firm 2 (Industrias Cerecita SA): Extended gravity factors

. . . at the early stages of their export history. . .

Figure: Year 2000

Motivation

Firm 2 (Industrias Cerecita SA): Extended gravity equation

. . . suddenly access countries that are very different from the country of origin but similar to previous export destinations . . .

Figure: Year 2002

Motivation

Firm 2 (Industrias Cerecita SA): Extended gravity equation

. . . and they don't stay for a long time in these countries . . .

Figure: Year 2003

Motivation

Firm 2 (Industrias Cerecita SA): Extended gravity equation

...and they never export high volumes to them.

Figure: Year 2004

Motivation

Firm 2 (Industrias Cerecita SA): Extended gravity equation

Are firm 2's entry costs in these "far away" countries as high as standard gravity factors would predict?

Figure: Year 2005

Aim of the paper

- Empirical analysis of country-specific firm export dynamics.
  - Analyse the determinants of firm entry and exit into spatially related markets.
- Entry and exit into each potential destination country may depend on:
  - Similarity between the importing country and firm's home country.
    - *Gravity*: closeness between home and destination markets.
  - Similarity between the current importing country and *prior* destinations of the firm's exports.
    - *Extended gravity*: similarities between two receiving countries.
- Quantify how strong *gravity* and *extended gravity* effects are in determining firms' country-specific entry and exit decisions.
  - Structural estimation of destination-specific costs of exporting that depend on *gravity* and *extended gravity* factors.
  - Measurement paper.

Quick summary of model and estimation

- Multi-period, multi-country firm-level trade model.
  - Single-agent dynamic model with multiple spatially related markets.
- Four different type of costs:

	Gravity	Extended Gravity
Marginal Costs	✓	
Fixed Costs	✓	
Sunk Costs	✓	✓

- Estimation based on a moment inequalities.
  - Application of an analogue of Euler's perturbation method.

Overview

Overview

Data sources

- Data sources:
  - Chilean customs database: revenue at the country-firm-year level.
  - Chilean industrial survey: sector (4 digit ISIC), domestic sales, value added, proportion of skilled workers, average wage at the firm-year level.
  - CEPII: geographical location and official language at the country level.
  - WDI: GDP and GDP per capita at the country-year level.
- Unbalanced panel of firms for the sample period: 1995-2005.
- Sector:
  - Manufacture of chemicals and chemical products (18% of manufacturing exports).

Transition matrices

Sector: Manufacture of chemicals and chemical products

- Strong effect of export history:

		Conditional on $t-1$		Conditional on $t-2$	
All Countries but South American		Entry	No Entry	Entry	No Entry
Options Relative	Shares Continent	0.035	0.965	0.010	0.990
to Prior Bundle	Does Not	0.006	0.994	0.005	0.995
All Countries but Spanish Speaking		Entry	No Entry	Entry	No Entry
Options Relative	Shares Language	0.024	0.976	0.010	0.990
to Prior Bundle	Does Not	0.009	0.991	0.008	0.992
All Countries but U Middle Income		Entry	No Entry	Entry	No Entry
Options Relative	Shares Income Group	0.021	0.979	0.006	0.994
to Prior Bundle	Does Not	0.005	0.995	0.004	0.996
All Countries but Bordering Chile		Entry	No Entry	Entry	No Entry
Options Relative	Shares Border	0.075	0.925	0.032	0.968
to Prior Bundle	Does Not	0.009	0.991	0.007	0.993

Overview

Model

Demand

- Utility function for the representative consumer of every country j in every sector:

$$Q_{jt} = \left[\int_{i \in A} q_{ijt}^{\frac{\eta-1}{\eta}} di \right]^{\frac{\eta}{\eta-1}}$$

- Budget constraint is:

$$\int_{i \in A} p_{ijt} q_{ijt} di = C_{jt}$$

Model

Supply

- Every firm i is the single supplier of a variety in any country.
- Every firm i faces a constant marginal cost of supplying any market j.

$$mc_{ijt} = f_{it} g_{jt}^{M} \epsilon_{ijt}^{M}$$

- Fixed costs of exporting:

$$fc_{ijt} = g_{j}^{F} + \epsilon_{ijt}^{F}$$

- Sunk costs of exporting:

$$sc_{ijb_{t-1}t} = g_{j}^{S} - e_{jb_{t-1}}^{S} + \epsilon_{ijt}^{S}$$

- Basic costs of exporting:

$$bc_{it} = \mu_{0}^{B} + \epsilon_{it}^{B}$$

- Terms:
  - g_{jt}^{M}, g_{j}^{F}, g_{j}^{S}: gravity equation terms
  - $e_{jb_{t-1}}^{S}$: EXTENDED gravity equation terms
  - $(\epsilon_{ijt}^{M}, \epsilon_{ijt}^{F}, \epsilon_{ijt}^{S}, \epsilon_{it}^{B})$: unobservable variables.

Interactions between destination countries

- How do we measure those interactions between destination countries?
  - Dummy variable for exporting at $t-1$ to at least one country that:
    - shares a border with j (and not with Chile).
    - has the same official language as j (except Spanish).
    - belongs to the same continent as j (except South America).
    - belongs to the same GDPpc bracket as j (except upper middle income countries).
- Therefore, a firm will not benefit from those interactions when entering j if:
  - It did not export in the previous period.
  - It exported in the previous period to countries that are very different from j.
  - Country j is very similar to Chile.

Static problem of the firm

- Given the demand and supply assumptions indicated above, the *potential* revenue from exporting to country j for firm i at period t (expressed in logs) is:

$$\log(r_{ijt}) = c + (1-\eta)\ln(mc_{ijt}) + \ln(C_{jt}) - (1-\eta)\ln(P_{jt})$$

  where P_j is the price index of the sector in the destination country j.

- The expression for the (log of) *gross* profits from exporting is:

$$\log(v_{ijt}) = -\log(\eta) + \log(r_{ijt})$$

Dynamic problem of the firm

- Possibly nonzero sunk costs of exporting make the entry and exit in each country j the outcome of a dynamic optimization problem.
- Possibly nonzero extended gravity effects in sunk costs make the entry and exit in each country j dependent on the entry and exit decision in every other country j'.
- In every period t, firm j is going to pick a *bundle* of export destinations taking into account the effect on future export profits.
- The size of the choice set ($2^{79} \sim 10^{23}$) makes it impossible to use any standard discrete choice model in the estimation.
- We base our estimation strategy on moment inequalities derived from the previous model.
- We apply Euler's perturbation method (one period deviations) to build these moment inequalities.

Dynamic problem of the firm

- The *net* profits from exporting are:

$$\pi_{ib_t b_{t-1} t} = \sum_{j \in b_t} \pi_{ijb_{t-1}t} - \mathbb{1}\{b_{t-1} = \emptyset\} bc_{it}$$

$$\pi_{ijb_{t-1}t} = v_{ijt} - fc_{ijt} - \mathbb{1}\{j \notin b_{t-1}\} sc_{ijb_{t-1}t}$$

Assumption 1: Firms Maximize Profits

Let's denote by $o_1^T = \{o_1, o_2, \ldots, o_T\}$ *an observed sequence of bundles, then:*

$$o_t = \underset{b_t \in \mathcal{B}_{io_{t-1}t}}{\text{argmax}} \; \mathcal{E}_i \left[\Pi_{ib_t o_{t-1} t} | \mathcal{J}_{it} \right] \qquad \forall \quad t = 1, 2, \ldots, T$$

where

$$\Pi_{ib_t o_{t-1} t} = \pi_{ib_t o_{t-1} t} + \delta \pi_{i\mathbf{b}_{t+1} b_t t+1} + \omega_{i\mathbf{b}_{t+1} t+2},$$

$$\omega_{i\mathbf{b}_{t+1} t+2} = \omega_{it+2}(\delta, \pi_{i\mathbf{b}_{t+2}\mathbf{b}_{t+1} t+2}, \pi_{i\mathbf{b}_{t+3}\mathbf{b}_{t+2} t+3}, \ldots),$$

$$\mathbf{b}_{t+s} = \underset{b_{t+s} \in \mathcal{B}_{ib_{t+s-1} t+s}}{\text{argmax}} \; \mathcal{E}_i \left[\Pi_{ib_{t+s} b_{t+s-1} t+s} | \mathcal{J}_{it+s} \right], \quad \forall s \geq 1. \qquad \blacksquare$$

Dynamic problem of the firm

- Assumption 1 is compatible with firms being forward-looking in many different degrees:
  - Firms may take into account only one period ahead:

$$\omega_{i\mathbf{b}_{t+1}t+2} = 0.$$

  - Firms may consider any finite number p of periods ahead:

$$\omega_{i\mathbf{b}_{t+1}t+2} = \delta^2 \pi_{i\mathbf{b}_{t+2}\mathbf{b}_{t+1}t+2} + \delta^3 \pi_{i\mathbf{b}_{t+3}\mathbf{b}_{t+2}t+3} + \cdots + \delta^p \pi_{i\mathbf{b}_{t+p}\mathbf{b}_{t+p-1}t+p};$$

  - Firms may consider an infinite number of period ahead (perfectly forward looking firms):

$$\omega_{i\mathbf{b}_{t+1}t+2} = \mathcal{E}_i \left[\Pi_{i\mathbf{b}_{t+2}\mathbf{b}_{t+1}t} \middle| \mathcal{J}_{it+2} \right].$$

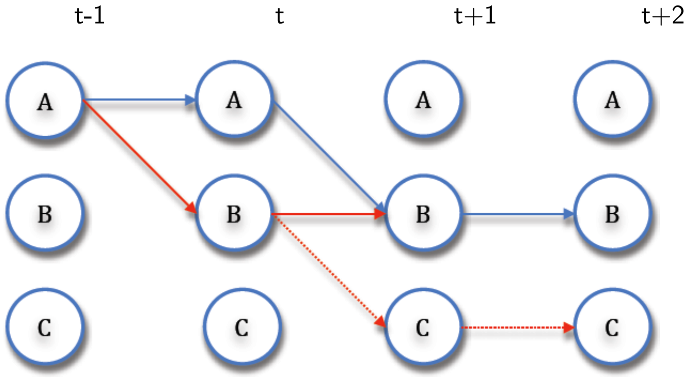
Instead of deriving a likelihood function from the model...

- Deriving choice probabilities from Assumption 1 requires:
  - Specify firms' choice sets: $\{\mathcal{B}_{io_{t-1}t}\}_{i=1,t=1}^{I,T}$.
  - Specify firms' conditional expectation functions: $\{\mathcal{E}_i[\cdot|\mathcal{J}_{it}]\}_{i=1}^{I}$.
  - Specify firms' information sets: $\{\mathcal{J}_{it}\}_{i=1,t=1}^{I,T}$.
  - Specify function $\omega_{i\mathbf{b}_{t+1}t+2}$.
  - Specify distribution function for the unobservable vector: $(\epsilon_{ijt}^{M}, \epsilon_{ijt}^{F}, \epsilon_{ijt}^{S}, \epsilon_{ijt}^{B})$.
  - Specify stochastic process for the state variables.
  - Solve the dynamic optimization problem:

$$o_t = \underset{b_t \in \mathcal{B}_{io_{t-1}t}}{\text{argmax}} \; \mathcal{E}_i\left[\Pi_{ib_t o_{t-1}t} | \mathcal{J}_{it}\right].$$

- Problems with this approach:
  - We impose assumptions on objects on which we have no information.
  - Solving the dynamic optimization problem is computationally very complicated.

Deriving moment inequalities: one-period deviations



0 $\mathcal{E}_i[(\pi_{iA_tA_{t-1}} - \pi_{iB_tA_{t-1}}) + \dots$ 0

$$\dots + \delta(\pi_{iB_{t+1}A_t} - \pi_{iB_{t+1}B_t})|\mathcal{J}_{it}]$$

Deriving moment inequalities: one-period deviations

Proof

- Given Assumption 1 and assuming $o'_t \in \mathcal{B}_{io_{t-1}t}$:

$$\mathcal{E}_i\left[\pi_{io_to_{t-1}t} + \delta\pi_{i\mathbf{o}_{t+1}o_tt}|\mathcal{J}_{it}\right] \geq \mathcal{E}_i\left[\pi_{io'_to_{t-1}t} + \delta\pi_{i\mathbf{o}_{t+1}o'_tt}|\mathcal{J}_{it}\right]$$

$$\mathbf{o}_{t+1} = \underset{b_{t+1}\in\mathcal{B}_{io_tt+1}}{\text{argmax}}\ \mathcal{E}_i\left[\Pi_{ib_{t+1}o_tt+1}|\mathcal{J}_{it+1}\right]$$

- Simplifying notation:

$$\mathcal{E}_i\left[\pi_{id\mathbf{o}_{t+1}t}|\mathcal{J}_{it}\right] \geq 0$$

$$\pi_{id\mathbf{o}_{t+1}t} = (\pi_{io_to_{t-1}t} - \pi_{io'_to_{t-1}t}) + \delta(\pi_{i\mathbf{o}_{t+1}o_tt+1} - \pi_{i\mathbf{o}_{t+1}o'_tt+1})$$

- Given that this inequality holds for every deviation, firm, and time period, the model predicts that:

$$\frac{1}{D^k}\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{d=1}^{D^k_{it}}\mathcal{E}_i[\pi_{id\mathbf{o}_{t+1}t}|\mathcal{J}_{it}] \geq 0.$$

Deriving moment inequalities: one-period deviations

- Finally, we impose some restrictions on the expectations of the agents:

Assumption 2: Firms are Right on Average

There is a positive valued function $g_{k_l}(\cdot)$ and an $x_{it} \in \mathcal{J}_{it}$ such that:

$$\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{d=1}^{D_{it}^{k}} \mathcal{E}_i[\pi_{id\mathbf{o}_{t+1}t}|\mathcal{J}_{it}] \geq 0 \quad \Rightarrow \quad \mathbb{E}\Big[\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{d=1}^{D_{it}^{k}} \pi_{id\mathbf{o}_{t+1}t}\, g_{k_l}(x_{idt})\Big] \geq 0$$

where $\mathbb{E}[\cdot]$ denotes the statistical expectation or expectation with respect to the data generation process.

■

Overview

List of parameters

- The list of parameters to estimate through moment inequalities is:
  - Fixed cost parameters:

$$\overbrace{\mu_0^F}^{\text{constant}}, \quad \underbrace{\mu_{cont}^F, \quad \mu_{lan}^F, \quad \mu_{gdppc}^F}_{\text{gravity effect}}$$

  - Basic cost parameters:

$$\overbrace{\mu_0^B}^{\text{constant}}$$

  - Sunk cost parameters:

$$\overbrace{\mu_0^S}^{\text{constant}}, \quad \underbrace{\mu_{cont}^S, \quad \mu_{lan}^S, \quad \mu_{gdppc}^S}_{\text{gravity effect}}, \quad \underbrace{\zeta_{bord}^S, \quad \zeta_{cont}^S, \quad \zeta_{lan}^S, \quad \zeta_{gdppc}^S}_{\text{extended gravity effect}}$$

Overview

An Example: μ_0^F

- Imagine we observe the following export history for firm i in country j:

Year	1	2	3	4	5
Profits	v_{ij1}	v_{ij2}	v_{ij3}	v_{ij4}	v_{ij5}
Exports	0	1	1	0	0

- Using the following counterfactual strategy:

Year	1	2	3	4	5
Actual	0	1	1	0	0
Counterfactual	0	1	1	1	0

we get the following difference in profits:

$$\pi_{ido_5 4} = -v_{ij4} + fc_{ij4}$$

- If j is in South America, speaks Spanish, and has an upper middle GDPpc:

$$\pi_{ido_5 4} = -v_{ij4} + \mu_0^F + \epsilon_{ij4}^F$$

An Example: ζ^S_{lan}

- Assume the following stream of profits and export strategies in countries j and j':

	Year	7	8	9
Actual	Country j	0	1	0
	Country j'	0	0	0
Counterfactual	Country j	0	0	0
	Country j'	0	1	0

- We get the following difference in profits:

$$\pi_{ido_9 8} = v_{ij8} - v_{ij'8} - (fc_{ij8} - fc_{ij'8}) - (sc_{ij8} - sc_{ij'8})$$

- If i exports at year 7 to a country that shares official language with j (and nothing with j') and neither j nor j' are in South America, speak Spanish or have an upper middle GDPpc, then the moment inequality simplifies to:

$$\pi_{ido_9 8} = v_{ij8} - v_{ij'8} + \zeta^S_{lan}$$

Overview

Estimation strategy: linear moment inequalities estimation

Mom.Ineq.Framework

- Our model provides S moment inequalities, each indexed by a pair (k, l):

$$m_s(\theta) = \mathbb{E}\Big[\frac{1}{D_k}\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{d=1}^{D_{it}^k}\pi_{id\mathbf{o}_{t+1}t}\, g_{k_l}(x_{idt})\Big] \geq 0$$

with $\theta = (\beta, \mu^F, \mu^S, \mu^B, \zeta^S)$, and:

$$\pi_{id\mathbf{o}_{t+1}t} = v_{idt}(\beta) - g_d^F(\mu_F) - (g^S_{d\mathbf{o}_{t+1}}(\mu_S) - e^S_{d\mathbf{o}_{t+1}}(\zeta_S)) - bc_{d\mathbf{o}_{t+1}}(\mu_B) - \epsilon_{2idt}$$

with $g^F(\cdot)$, $g^S(\cdot)$, $e^S(\cdot)$, and $bc(\cdot)$ linear functions and $v(\cdot)$ is *loglinear*.

- We follow a two-stage estimation strategy:
  - First stage: linear panel data estimates of β.
  - Second stage: linear moment inequality estimates of $(\mu^F, \mu^S, \mu^B, \zeta^S)$ are obtained conditional on the first stage estimates $\hat{\beta}$.

First stage estimation

- In order to estimate β we use the revenue equation that arises from solving the static problem of the firm.

$$\log(r_{ijt}) = \beta Z_{ijt} + (1-\eta)\epsilon_{ijt}^{M}$$

Results

and we define an approximation to gross profits from exporting as:

$$\hat{v}_{ijt} = \frac{1}{\hat{\eta}}\hat{\alpha}\exp(\hat{\theta}_1 Z_{ijt})$$

where we borrow the estimate $\hat{\eta}$ from Broda, Greenfield, and Weinstein (2006), and $\hat{\alpha}$ accounts for the effect of higher order moments of ϵ_{ij}^{M} on the expected value of v_{ij}.

- We define the first stage error as:

$$\epsilon_{1idt} = v_{idt} - \hat{v}_{idt}$$

Second stage estimation

- In order to obtain estimates for $(\mu^F, \mu^S, \mu^B, \zeta^S)$ we build sample analogues of the moment inequalities derived from the model.

$$\tilde{m}_s(\mu^F, \mu^S, \mu^B, \zeta^S) = \frac{1}{D_k} \sum_{i=1}^{I} \sum_{t=1}^{T} \sum_{d=1}^{D_{it}^k} \tilde{\pi}_{ido_{t+1}t} g_{k_l}(x_{idt}) \geq 0$$

and

$$\begin{aligned} \tilde{\pi}_{ido_{t+1}t} &= \hat{v}_{idt} - g_d^F(\mu_F) - (g^S_{d\mathbf{o}_{t+1}}(\mu_S) - e^S_{d\mathbf{o}_{t+1}}(\zeta_S)) - bc_{d\mathbf{o}_{t+1}}(\mu_B) \\ &= \pi_{ido_{t+1}t}(\theta) - \epsilon_{idt} \end{aligned}$$

with $\epsilon_{idt} = \epsilon_{1idt} + \epsilon_{2idt}$

Estimation strategy

Second stage estimation

- Our linear moment inequality estimate is consistent as long as the following assumption holds:

Assumption 3: Orthogonality

$$\frac{1}{D_k}\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{d=1}^{D_{it}^k}\epsilon_{idt}g_{k_l}(x_{idt}) \xrightarrow{D^k\to\infty} 0.$$

■

- Caveat: Heterogeneity vs. state-dependence?
  - Correcting for selection in ϵ_{idt} (see Dickstein and Morales (2012))

Overview

Combining parameters

Manufacture of chemicals and chemical products: narrow version

If the firm at t exports to...	If the firm at t-1 exported to...	Estimates Lower	Estimates Upper	Estimates Midpoint	Mean profits/Cost Exporters	Mean profits/Cost All
US	US	223,708	242,627	233,170	1.95	0.72
	Canada	319,578	520,165	419,870	1.08	0.40
	Mexico	379,653	538,986	459,320	0.99	0.36
	UK	319,578	520,165	419,870	1.08	0.40
	Colombia	387,527	538,986	463,260	0.98	0.36
	Chile	387,527	650,634	519,080	0.88	0.32
Brazil	Brazil	219,751	242,528	231,140	1.47	0.75
	Argentina	277,823	421,634	349,730	0.97	0.49
	Portugal	223,374	409,745	316,560	1.07	0.55
	Ecuador	283,323	421,640	352,480	0.96	0.49
	Chile	283,323	570,953	427,138	0.79	0.41
Colombia	Colombia	215,785	235,578	225,680	1.42	0.68
	Venezuela	264,520	313,216	288,868	1.11	0.53
	Argentina	277,303	313,216	295,260	1.08	0.52
	Chile	277,303	481,768	379,540	0.84	0.41

Summary

- Moment inequalities allow for identification of structural parameters in single agent dynamic models that are computationally impossible to estimate using standard estimation methods.
  - Possibility of dealing very large discrete choice sets.
- This is possible because we do not need to solve the model for identification and estimation.
- Besides, we incorporate distribution-free measurement error.
- Drawbacks:
  - Strong assumptions on structural errors.
  - Impossibility of performing counterfactuals without previously solving the model.

Proof

Back

- From Assumption 1 we know that:

$$\mathcal{E}_i[\pi_{io_t o_{t-1} t} + \delta\pi_{i\mathbf{o}_{t+1} o_t t+1} + \omega_{i\mathbf{o}_{t+1} t+2} | \mathcal{J}_{it}] \geq \mathcal{E}_i[\pi_{io'_t o_{t-1} t} + \delta\pi_{i\mathbf{o}'_{t+1} o'_t t+1} + \omega_{i\mathbf{o}'_{t+1} t+2} | \mathcal{J}_{it}],$$

with

$$\mathbf{o}_{t+1} = \underset{b_{t+1} \in \mathcal{B}_{io_t t+1}}{\operatorname{argmax}} \; \mathcal{E}_i \left[\Pi_{ib_{t+1} o_t t+1} | \mathcal{J}_{it+1} \right],$$

and

$$\mathbf{o}'_{t+1} = \underset{b_{t+1} \in \mathcal{B}_{io'_t t+1}}{\operatorname{argmax}} \; \mathcal{E}_i \left[\Pi_{ib_{t+1} o'_t t+1} | \mathcal{J}_{it+1} \right].$$

By transitivity of preferences,

$$\mathcal{E}_i[\pi_{io_t o_{t-1} t} + \delta\pi_{i\mathbf{o}_{t+1} o_t t+1} + \omega_{i\mathbf{o}_{t+1} t+2} | \mathcal{J}_{it}] \geq \mathcal{E}_i[\pi_{io'_t o_{t-1} t} + \delta\pi_{i\mathbf{o}_{t+1} o'_t t+1} + \omega_{i\mathbf{o}_{t+1} t+2} | \mathcal{J}_{it}],$$

Canceling terms on both sides:

$$\mathcal{E}_i[\pi_{io_t o_{t-1} t} + \delta\pi_{i\mathbf{o}_{t+1} o_t t+1} | \mathcal{J}_{it}] \geq \mathcal{E}_i[\pi_{io'_t o_{t-1} t} + \delta\pi_{i\mathbf{o}_{t+1} o'_t t+1} | \mathcal{J}_{it}]. \quad \text{Q.E.D.}$$

Estimation strategy

First stage results: $\hat{\theta}_1$

Independent Variable	Log Revenue Regression
$domsales_{it}$	0.431***
	(0.040)
$avgskill_{it}$	-0.427***
	(0.150)
$lavgwage_{it}$	0.322***
	(0.077)
$lavgvalueadd_{it}$	-0.140***
	(0.057)
$border_{jt}$	0.409***
	(0.084)
$cont_{jt}$	0.134
	(0.084)
$lang_{jt}$	-0.198*
	(0.103)
$gdppc_{jt}$	0.179***
	(0.068)
$lgdp_{jt}$	0.250***
	(0.022)
$avglrer_{j}$	-0.103***
	(0.012)
$devlrer_{jt}$	0.610***
	(0.259)
$legal_{j}$	-0.069***
	(0.020)
N Observations	6,253

Notes: * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance. Robust standard errors are in parentheses. The dependent variable is log revenue. Year fixed effects are included