EC309: Class 10

MT 2022

Pinjas Albagli

December 9, 2022

A simple model of demand for differentiated products

Model: Demand side

- A continuous measure I of consumers indexed by $i \in \mathcal{I}$ with unit demands choose among a set of differentiated products indexed by $j \in \mathcal{J} = \{1, \ldots, J\}$ and an outside option (think not purchasing) represented by $j = 0$.

- Preferences:

$$U_i(c, j) = \tilde{\alpha} c + \tilde{x}_j' \tilde{\beta} + \varepsilon_{ij}$$

  - $c \in \mathbb{R}_+$: consumption of other goods with price normalized to 1.
  - $\tilde{x}_j \in \mathbb{R}^{\tilde{K}}$: vector of utility-relevant characteristics of alternative $j \in \mathcal{J} \cup \{0\}$ with $\tilde{x}_0 \equiv 0$.
  - $\varepsilon_{ij} \in \mathbb{R}$: idiosyncratic taste shock of consumer i for alternative j (think of match value).

Model: Demand side

- Consumer problem:

$$\max_{c\in\mathbb{R}_+,\, j\in\mathcal{J}_0} U_i\left(c, j\right)$$

$$\text{s.t. } p_j + c \leq y_i$$

  - $\mathcal{J}_0 \equiv \mathcal{J} \cup \{0\}$.
  - $p_j \in \mathbb{R}_{++}$: price of alternative j with $p_0 \equiv 0$.
  - $y_i \in \mathbb{R}_{++}$: consumer i's income.

Model: Demand side

- Under local nonsatiation, budget constraint is binding $\implies$ can solve for c and substitute into utility function:

$$\begin{aligned}
\max_{j \in \mathcal{J}_0} U_i\left(y_i - p_j, j\right) &= \max_{j \in \mathcal{J}_0} \tilde{\alpha}\left(y_i - p_j\right) + \tilde{x}_j' \tilde{\beta} + \varepsilon_{ij} \\
&= \max_{j \in \mathcal{J}_0} \alpha_i + \alpha p_j + \tilde{x}_j' \tilde{\beta} + \varepsilon_{ij} \\
&\equiv \max_{j \in \mathcal{J}_0} \tilde{u}_{ij}
\end{aligned}$$

where

- $\alpha_i \equiv \tilde{\alpha} y_i$
- $\alpha \equiv -\tilde{\alpha}$

Model: Demand side

- We can normalize by subtracting α_i from $\tilde{u}_{ij}$ $\forall j \in \mathcal{J}_0$:

$$\arg\max_{j \in \mathcal{J}_0} \tilde{u}_{ij} = \arg\max_{j \in \mathcal{J}_0} u_{ij}$$

where

  - $u_{ij} \equiv \tilde{u}_{ij} - \alpha_i = \alpha p_j + \tilde{x}_j' \tilde{\beta} + \varepsilon_{ij}$
  - $u_{i0} = \varepsilon_{i0}$

Model: Demand side

- Assume a specific distribution for the idiosyncratic tastes:

$$\varepsilon_{ij} \underset{(i,j)}{\overset{iid}{\sim}} \text{Extreme Value Type I}$$

with pdf $f(\varepsilon) = \exp(-\varepsilon) \exp(-\exp(-\varepsilon))$ and cdf $F(\varepsilon) = \exp(-\exp(-\varepsilon))$

Model: Demand side

- Let $\delta_j \equiv \alpha p_j + \tilde{x}_j' \tilde{\beta} \implies u_{ij} = \delta_j + \varepsilon_{ij}$ for $j \in \mathcal{J}_0$, with $\delta_0 = 0$ and $\delta \equiv (\delta_0, \ldots, \delta_J)$.
- It can be shown that the market share of good j is given by

$$s_j(\delta) = \mathbb{P}\left(u_{ij} > u_{ik} \; \forall k \in \mathcal{J}_0 \setminus \{j\}\right)$$

$$= \frac{\exp(\delta_j)}{\sum_{k \in \mathcal{J}_0} \exp(\delta_k)}.$$

- Therefore, total demand for product j is

$$q_j = I \, s_j(\delta).$$

Model: Demand side

- Let $\{s_j\}_{j\in\mathcal{J}_0}$ represent observed market shares and notice that (at the true δ)

$$s_j = \mathscr{s}_j\left(\delta\right).$$

- Total demand for product j is then $s_j\, I$ and it can be shown that the price elasticities are given by

$$\eta_{jk} \equiv \frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j}$$

$$= \begin{cases} \alpha\left(1 - s_j\right) p_j & \text{for } k = j \\ -\,\alpha\, s_k\, p_k & \text{for } k \neq j \end{cases}$$

Model: Supply side

- There are J firms, with firm j producing good $j \in \mathcal{J}$ at constant marginal cost c_j.
- Firms are profit-maximizing, price-setting oligopolists:

$$\max_{p_j \in \mathbb{R}_{++}} \Pi_j(p, \tilde{x}) = \max_{p_j \in \mathbb{R}_{++}} [p_j - c_j] \, I \, s_j(\delta(p, \tilde{x}))$$

where

- $p \equiv (p_1, \ldots, p_J)$.
- $\tilde{x} \equiv \text{vec}\left(\begin{bmatrix} \tilde{x}_1 & \cdots & \tilde{x}_J \end{bmatrix}\right)$

Model: Supply side

- Assuming the existence of a pure strategy Nash equilibrium, the first order condition for p_j can be written as

$$p_j = c_j - \frac{s_j\left(\delta\left(p, \tilde{x}\right)\right)}{\frac{\partial s_j(\delta(p,\tilde{x}))}{\partial \delta_j}\frac{\partial \delta_j(p_j,\tilde{x}_j)}{\partial p_j}}$$

$$= c_j - \frac{s_j\left(\delta\right)}{s_j\left(\delta\right)\left[1 - s_j\left(\delta\right)\right]\alpha}$$

Model: Supply side

- If the J FOCs define a unique Nash equilibrium for all $\tilde{x}$ and $(\alpha, \beta')'$, then they implicitly define a reduced-form price equation

$$p_j = p_j(\tilde{x}).$$

- Then, the implied reduced-form quantity equation is

$$q_j = I\, s_j(\delta(p(\tilde{x}), \tilde{x}))$$

$$\equiv q_j(\tilde{x})$$

Research design

Data

- Suppose we do **not** have access to micro-level data on consumer purchases.We have aggregate data instead, including:
  - product identifiers (i.e., $\{j\}_{j \in \mathcal{J}}$)
  - market shares (i.e., $\{s_j\}_{j \in \mathcal{J}_0}$)
  - prices (i.e., $\{p_j\}_{j \in \mathcal{J}}$)
  - product characteristics (i.e., $\{\tilde{x}_j\}_{j \in \mathcal{J}}$; initially assume we observe the entire vector $\tilde{x}_j$)

Identification: Inverting the market share equation[1]

- Note that for $j \in \mathcal{J}$

$$\begin{aligned} \ln(s_j) - \ln(s_0) &= \ln(\exp(\delta_j)) - \ln\left(\sum_{k \in \mathcal{J}_0} \exp(\delta_k)\right) - \left[\ln(1) - \ln\left(\sum_{k \in \mathcal{J}_0} \exp(\delta_k)\right)\right] \\ &= \delta_j \end{aligned}$$

- Therefore, $\{\delta_j\}_{j \in \mathcal{J}}$ are identified from market shares (i.e., aggregate data).

[1] See Steven Berry (1994), "Estimating Discrete-Choice Models of Product Differentiation", *The RAND Journal of Economics*, 25(2): 242–262 for details.

Identification: Ideal case

- If all utility-relevant product characteristics $\tilde{x}_j$ are observed, then $(\alpha, \beta')'$ is identified and can be recovered from an OLS regression (linear projection) of δ_j on $\left(p_j, \tilde{x}_j'\right)'$ (note that we get a perfect fit).

$$\ln\left(s_j\right) - \ln\left(s_0\right) = \alpha p + \tilde{x}_j' \beta$$

Identification: Uncorrelated unobserved characteristics

- Now, partition $\tilde{x}_j = \left(x_j', w_j'\right)$ and $\tilde{\beta} = (\beta', \gamma')'$ with $x_j, \beta \in \mathbb{R}^K$, $w_j, \gamma \in \mathbb{R}^M$, and $K + M = \tilde{K}$.

- Suppose we only observe product characteristics x_j and let $\xi_j \equiv w_j'\gamma \in \mathbb{R}$.Then,

$$\delta_j = \alpha p_j + x_j'\beta + \xi_j .$$

- If $\mathbb{E}\left[\left(p_j, x_j'\right)' \xi_j\right] = 0$, $(\alpha, \beta')'$ is identified and can be consistently estimated by OLS regression of δ_j on $\left(p_j, x_j'\right)'$ (and, obviously, we no longer get a perfect fit).

Identification: Correlated unobserved characteristics

- However, $\mathbb{E}\left[\left(p_j, x_j'\right)' \xi_j\right] = 0$ is unlikely to hold in typical applications. In particular:
    - It is typically assumed that the location of products in characteristic space is exogenous, i.e., $\mathbb{E}\left[x_j \xi_j\right] = 0$.
    - But price is most likely endogenous: $\mathbb{E}\left[p_j \tilde{x}_j\right] \neq 0$ in general and $\mathbb{E}\left[p_j \xi_j\right] \neq 0$ is particularly concerning in our case.
- If we can find instruments $z_j \in \mathbb{R}^L$ with $L \geq K+1$, $\mathbb{E}\left[z_j \xi_j\right] = 0$, and rank $\left(\mathbb{E}\left[z_j z_j'\right]^{-1} \mathbb{E}\left[z_j \left(p_j, x_j'\right)\right]\right) = K+1$, then $(\alpha, \beta')'$ is identified and can be consistently estimated by 2SLS or GMM.

Identification: BLP instruments[2]

- Notice that under the assumption of exogenous characteristics, $\{x_k\}_{k\in\mathcal{J}\setminus\{j\}}$ provide valid instruments:

  1. they are excluded from the utility functionin the sense that u_{ij} does not depend on $\tilde{x}_{ik}$ for $k \neq j$, and
  2. market equilibrium implies that they correlate with pricethrough the reduced form price function implicitly defined by FOC system.

[2] See Steven Berry, James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium", *Econometrica*, 63(4): 841–90 for details.

Identification: BLP instruments

- Conditional moment restrictions $\mathbb{E}[\xi_j \mid z_j] = 0$ imply the existence of optimal instruments $q(z_j) \in \mathbb{R}^Q$ in a GMM framework:[3]
  - LIE $\implies \mathbb{E}[q(z_j)\xi_j] = \mathbb{E}[q(z_j)\mathbb{E}[\xi_j \mid z_j]] = 0 \; \forall q : \mathbb{R}^L \to \mathbb{R}^Q$ such that $\mathbb{E}\left[\|q(z_j)\|^2\right] < \infty$.
  - We can choose $q(z_j)$ to minimize the asymptotic variance of the GMM estimator based on (unconditional) moment restrictions $\mathbb{E}[g(z_j, \xi_j)] = 0$ where $g(z_j, \xi_j) \equiv q(z_j)\xi_j$.
  - BLP argue that the optimal instruments based on $\{x_k\}_{k \in \mathcal{J} \setminus \{j\}}$ are hard to compute and suggest estimating them with

$$\left(x_j', \sum_{k \in \mathcal{J} \setminus \{j\}} x_k' \right)'$$

[3] See, e.g., Hansen (1985) and Chamberlain (1987).

Simulation

Simulation setup

- We will assume

- $J = 50$ firms/products
- $K = 6$ product characteristics
- $M = 1,000$ independent markets
- $I = 1,000$ consumers per market

- A random subset $\mathcal{F}_m \subset \mathcal{J}$ of firms compete in market m
- $J_m \equiv |\mathcal{F}_m| \in \{3, \ldots, 30\}$
- Firms j' marginal cost c_j is homogeneous across markets
- Product characteristics $\tilde{x}_j$ are homogeneous across markets

Simulation setup

- We will begin by drawing

- $\boldsymbol{X} = \begin{bmatrix} \tilde{x}_1 & \cdots & \tilde{x}_J \end{bmatrix}'_{J \times K}$ from $\tilde{x}_{jk} \underset{(j,k)}{\overset{iid}{\sim}} \mathsf{U}(0,1)$
- $c = \begin{pmatrix} c_1 & \cdots & c_J \end{pmatrix}'_{J \times 1}$ from $c_j \underset{j}{\overset{iid}{\sim}} \mathsf{U}(0,1)$
- $\alpha = -|u|$ with $u \sim \mathsf{N}(0,1)$
- $\beta_k = |u_k|$ with $u_k \underset{k}{\overset{iid}{\sim}} \mathsf{N}(0,1)$

- We will simulate 3 different scenarios.

Simulation setup

1. Uncorrelated product characteristics:
   - For each market $m \in \{1, \ldots, M\}$:
     - Completely disregard the supply side of the model.
     - Draw $p_m = \begin{pmatrix} p_{j_1} & \cdots & p_{j_{J_m}} \end{pmatrix}'_{J_m \times 1}$ from $p_{j_f} \overset{iid}{\underset{j_f}{\sim}} \text{U}(0, 1) \implies p_m$ is exogenous.
     - Compute the implied market shares and construct an aggregate dataset.

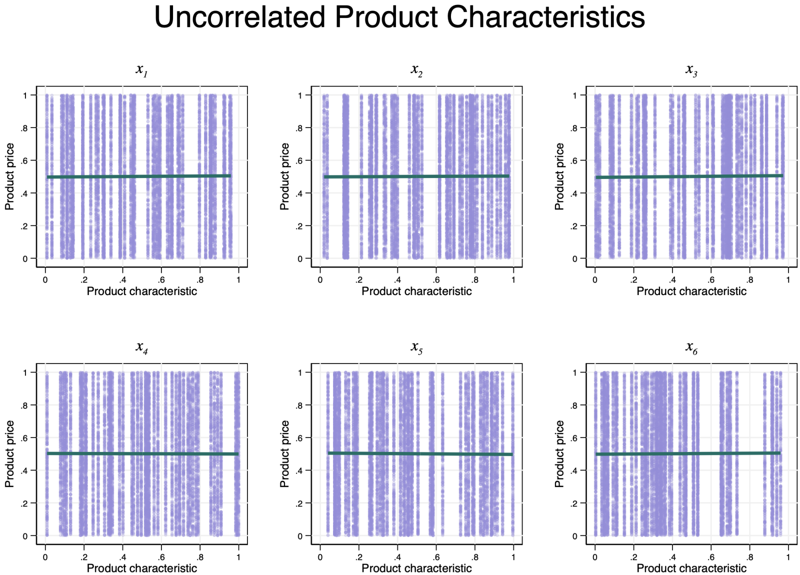
Simulation setup

2. Full model (aggregate):
   - For each market $m \in \{1, \dots, M\}$:
     - Solve the oligopoly model (numerically) given $\boldsymbol{X}(\mathcal{F}_m)$ and $c(\mathcal{F}_m)$.
     - Compute the implied market shares and construct an aggregate dataset.
     - p_j is now endogenous and will generally correlate with $\tilde{x}$.

Simulation setup

3. Full model (micro):
   - Take a random subsample of (already simulated) markets $\mathcal{M} \subseteq \{1, \ldots, M\}$.
   - For each market $m \in \mathcal{M}$:
     - Simulate I consumers by drawing $\varepsilon_{ij} \underset{(i,j)}{\sim}$ Extreme Value Type I.
     - Solve their optimization problem $j_i = \underset{j \in \mathcal{F}_j}{\arg\max}\ u_{ij}$
     - Collect choices from consumers in all markets and construct a micro level dataset.

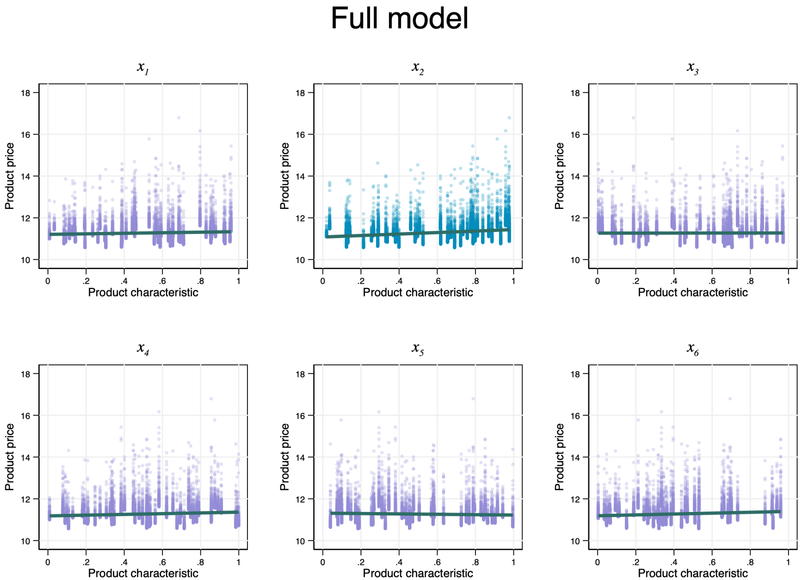
Results: Simulation 1



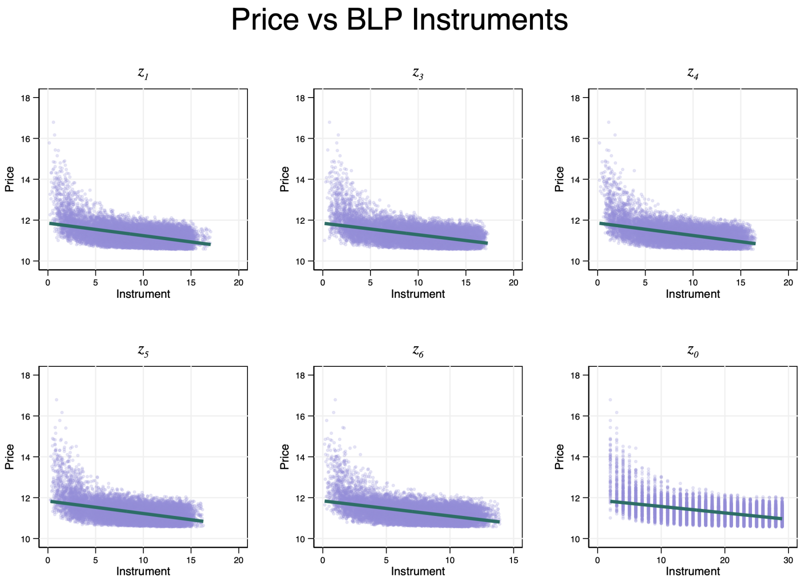
Results: Simulation 1

	True value (1)	Ideal case: 1 market (2)	Ideal case: all (3)	Ideal case: all (4)	Unobserved chars. (5)	Unobserved chars. (6)
α	-.09600331	-.09600317	-.09600331	-.1037554	-.1006866	-.0938619
		(.)	(2.122e-09)	(.0122413)	(.01304542)	(.00507127)
β_1	.6435798	.6435799	.6435798	.6004406	.7282101	.7752318
		(.)	(2.349e-09)	(.01372512)	(.01451106)	(.00559429)
β_2	1.098282	1.098282	1.098282	1.048032	.9590136	1.005762
		(.)	(2.244e-09)	(.01310407)	(.01406804)	(.00571655)
β_3	.01442098	.01442109	.01442099	.01346644	-.00428069	-.0011558
		(.)	(2.271e-09)	(.0132637)	(.01383497)	(.00504379)
β_4	.353057	.3530568	.353057	.3329975	.3216736	.3390694
		(.)	(2.259e-09)	(.01321704)	(.01418514)	(.00534151)
β_5	.09964361	.09964366	.09964361	.07577854		
		(.)	(2.127e-09)	(.01242897)		
β_6	.6959321	.6959323	.6959321	.6706914		
		(.)	(2.371e-09)	(.01384664)		
Observations		8	16,646	16,646	16,646	16,646
Market FE		no	yes	no	no	yes
R squared		1	1	.47	.39	.92

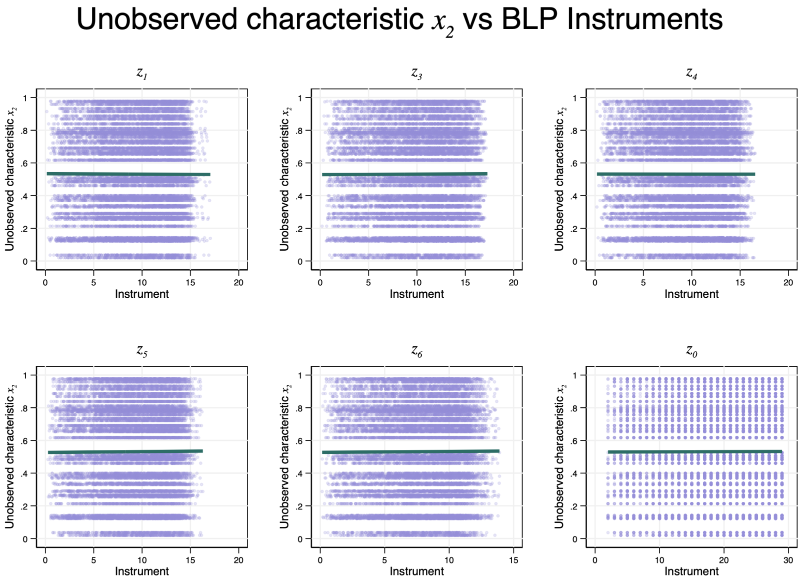
Results: Simulation 2



Results: Simulation 2



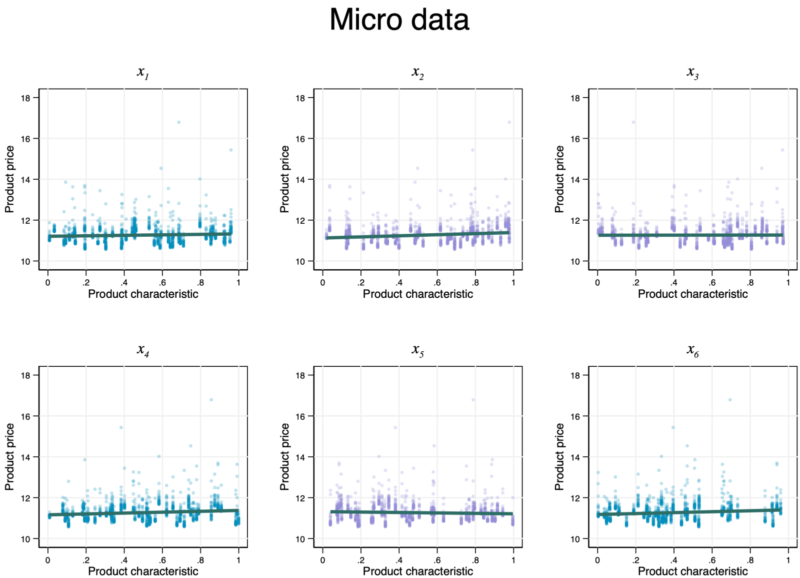
Results: Simulation 2



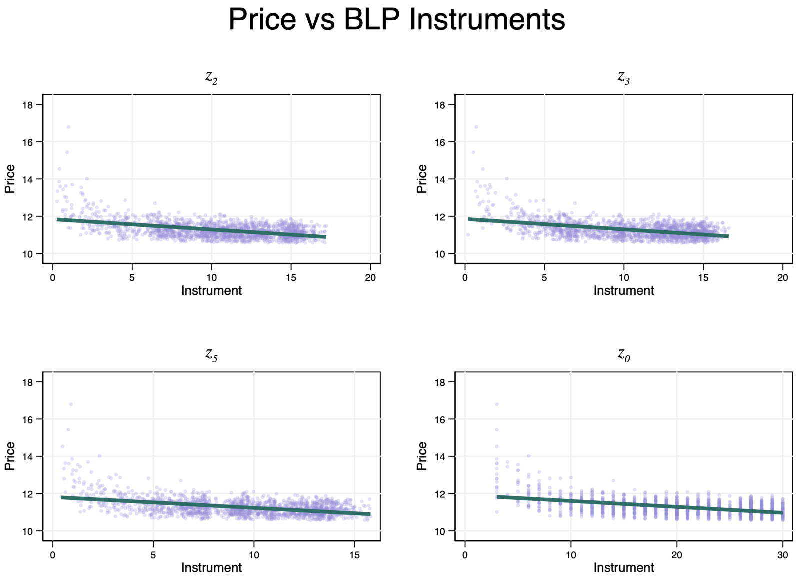
Results: Simulation 2

	True value (1)	Ideal case: all (2)	Ideal case: 1 mkt (3)	OLS (4)	2SLS (5)	GMM (6)
α	-.09600331	-.09600331	-.09600331	.05029967	-.09668759	-.09665311
		(2.943e-10)	(.)	(.00502223)	(.01031643)	(.01031525)
β_1	.6435798	.6435798	.6435801	.9132266	.9257918	.9257825
		(2.982e-09)	(.)	(.00910179)	(.00901818)	(.00901799)
β_2	1.098282	1.098282	1.098282			
		(2.888e-09)	(.)			
β_3	.01442098	.01442098	.01442087	-.09624888	-.09453851	-.09460775
		(2.882e-09)	(.)	(.00858381)	(.00881397)	(.00881318)
β_4	.353057	.353057	.353057	.404345	.4287442	.4287129
		(2.883e-09)	(.)	(.00853423)	(.00890241)	(.00890192)
β_5	.09964361	.09964361	.09964354	.08872654	.0787545	.07881458
		(2.687e-09)	(.)	(.0082242)	(.00821652)	(.00821565)
β_6	.6959321	.6959321	.6959322	.5087821	.5370063	.537018
		(3.030e-09)	(.)	(.00941232)	(.00980163)	(.00980147)
Observations	-0.0960033	16,646	8	16,646	16,646	16,646
R squared	0.6435798	1	1	.57	.55	.55

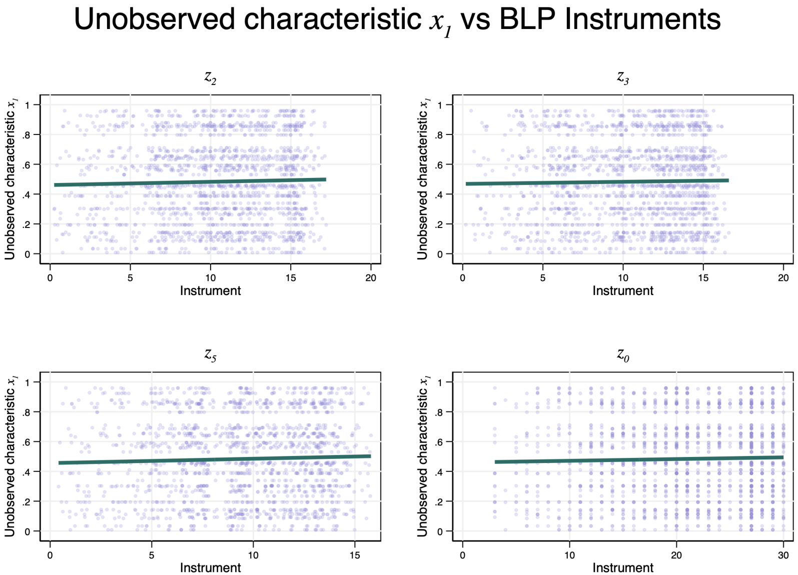
Results: Simulation 3



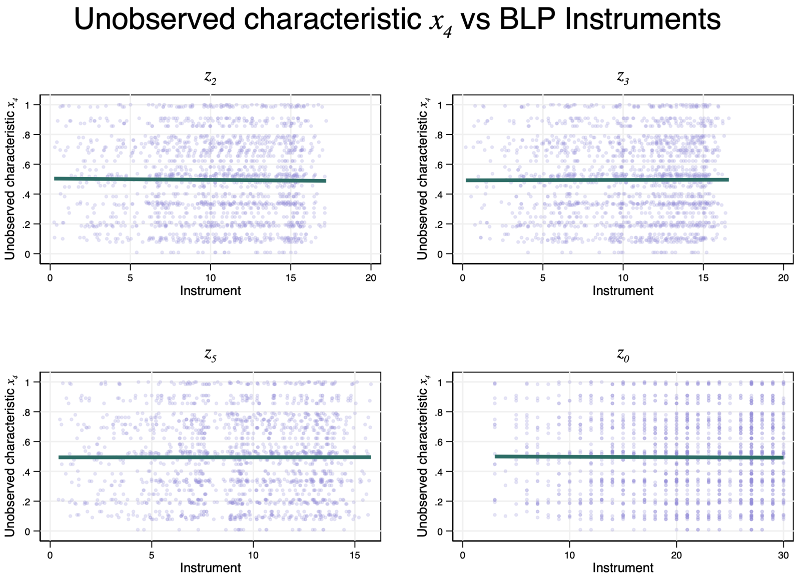
Results: Simulation 3



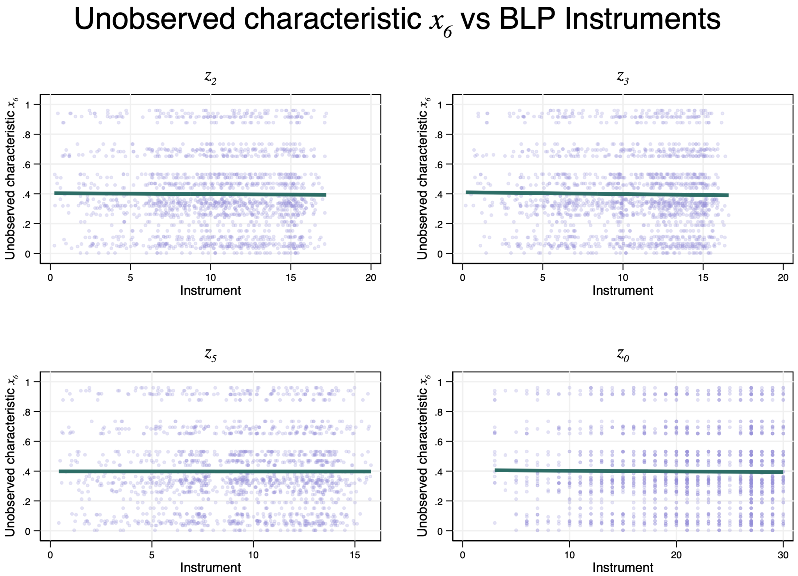
Results: Simulation 3



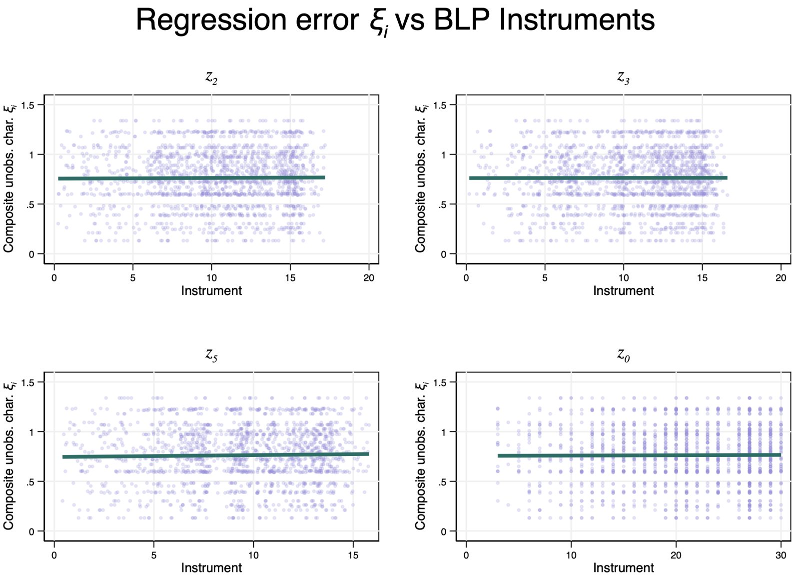
Results: Simulation 3



Results: Simulation 3



Results: Simulation 3



Results: Simulation 3

	True value (1)	MLE: full (2)	MLE: 1 mkt (3)	MLE: unobs (4)	2S: full (5)	2S: OLS (6)	2S: 2SLS (7)
α	-.09600331	-.09562517	-.1111048	-.03607351	-.09624459	-.03640158	-.0861858
		(.00173241)	(.02150302)	(.00146551)	(.00234374)	(.00260278)	(.03691428)
β_1	.6435798	.6338653	.4001693		.6245034		
		(.01392492)	(.1568457)		(.02193014)		
β_2	1.098282	1.099261	1.276831	1.192569	1.121005	1.219285	1.376285
		(.01279931)	(.1441864)	(.01279111)	(.02155154)	(.03374266)	(.1194792)
β_3	.01442098	.01774573	.17592	.1453956	.0228891	.1621521	.2928563
		(.01246104)	(.172369)	(.01168156)	(.02083269)	(.02957332)	(.1009159)
β_4	.353057	.329237	.4790873		.3248394		
		(.0134101)	(.1926015)		(.02244393)		
β_5	.09964361	.08720957	.06435913	.02598797	.08565779	-.01853273	.08073281
		(.01185686)	(.1386335)	(.01183394)	(.02021623)	(.03169017)	(.07951131)
β_6	.6959321	.6819313	.8032762		.7022982		
		(.01360263)	(.1571375)		(.02274619)		
Observations		1,772,000	20,000	1,772,000	1,772	1,772	1,772
R squared					.79	.48	.43