Pakes Porter Ho Ishii (2011)

+

Moment Inequalities in Binary Choice Problems (Based on Dickstein and Morales (2013))

Michael J. Dickstein
Stanford University

Economics 258

March 2, 2014

Moment Inequalities in Binary Choice Problems

Based on Dickstein and Morales (2013)

Sources of Error in Empirical Models

A. Structural Error

- **Structural error**: component of the agent's payoff function known to the agent at the time of her decision but not observed by the econometrician.

Sources of Error in Empirical Models

A. Structural Error

- **Structural error**: component of the agent's payoff function known to the agent at the time of her decision but not observed by the econometrician.
    - Typically assumed to be the single source of error in structural models.
    - Economic theory does not generally place restrictions on its distribution. Therefore,

Sources of Error in Empirical Models

A. Structural Error

- **Structural error**: component of the agent's payoff function known to the agent at the time of her decision but not observed by the econometrician.
  - Typically assumed to be the single source of error in structural models.
  - Economic theory does not generally place restrictions on its distribution. Therefore,
    - In order to avoid **endogeneity**, we assume independence between structural error and observed covariates
    - In discrete choice models, the econometrician (arbitrarily) chooses its distribution (up to a finite parameter vector)

Sources of Error in Empirical Models

Expectational Error

- **Expectational error**: Mismatch between (a) the expectations of payoffs agents use when making decisions and (b) future realized payoffs.
- Standard datasets usually contain information on *ex post* realizations of payoff relevant variables but rarely incorporate information on expectations.
    - Expectational error often contributes to the error term in empirical models.

Sources of Error in Empirical Models

Expectational Error

- Economic theory places restrictions on the distribution of expectational error.
- The rational expectations assumption implies that:
    - Expectational error is mean independent of the unobserved expectation (analogous to the classical **error-in-variables**).
    - Expectational error is correlated with observed covariate (**endogeneity**)
    - Any variable in the agent's information set is a valid **instrumental variable**.

Example: Export Decision

Agents' Information and Expectations

Daniel owns a small winery in Chile and is wondering whether he should try to export to Venezuela. To access customers in this market, Daniel needs to participate in a wine trade fair to be organized in June 2001 in Caracas. Daniel needs to make a decision by December 31, 2000.

Daniel has information on shipping costs and does not face any other export costs.

Daniel does not know the exact sales revenues he will obtain if he attends the fair.

Example: Export Decision

Agents' Information and Expectations

Using the information available to him (sales revenue of competitors, demand predictions, current political unrest, nominal exchange rate, etc) Daniel must form an **expectation** about his potential export revenues.

Daniel's choice over whether to attend the fair (or not) is a **binary** choice. He will attend the fair if his **expected** sales revenues net of shipping costs are positive.

Example: Export Decision

Econometrician's Information

- The Chilean Customs Agency provides a dataset on the annual country-specific **realized** sales revenues obtained by each Chilean wine producer during 1995-2005
- The econometrician assumes that shipping costs are a linear function of the distance between Santiago de Chile and Caracas (distance known).
- The econometrician does not know: (a) the exact shipping cost per mile that Daniel will face if he chooses to export; (b) the expectation that Daniel had on December 31, 2000 about the revenue he would obtain in 2001 if he were to attend the fair.

Example: Export Decision

How shall the econometrician handle the unobserved shipping costs?

The shipping costs form the **structural error**; i.e. variables known to Daniel at the time of his decision, but unobserved by the econometrician.

The econometrician assumes that these shipping costs are normally distributed around a mean that he will estimate, and sets the variance to an arbitrary constant.

Example: Export Decision

How should the econometrician handle the unobserved expectations?

Options:

- 1. Compute Daniel's unobserved expectations. Assumptions:
  - Realized revenues are observed without error.
  - All information Daniel used to form his expectation is available to the econometrician (i.e. Daniel's information set is observable);
  - Daniel has rational expectations.

Example: Export Decision

How should the econometrician handle the unobserved expectations?

Options:

- 1. Compute Daniel's unobserved expectations. Assumptions:
  - Realized revenues are observed without error.
  - All information Daniel used to form his expectation is available to the econometrician (i.e. Daniel's information set is observable);
  - Daniel has rational expectations.
- 2. Assume that Daniel's expectations are identical to the observed measurement of realized revenues. Assumptions:
  - Realized revenues are observed without error.
  - Daniel has perfect foresight.

Example: Export Decision

How should the econometrician handle the unobserved expectations?

Options:

- 3. Use the ex post measurement of potential revenue as a proxy for the unobserved expectation. Assumptions:
  - Daniel has rational expectations.

Option 3 imposes fewer assumption but introduces **errors-in-variables**.

Goal of the paper

Illustrate how to use **moment inequalities** to **identify and estimate the index coefficients** of **binary choice models** allowing for **structural errors** and **errors-in-variables.**

Expectational and Measurement Error

- The payoff function of agent i for alternative j is

$$U_j = \beta \mathcal{E}[X_j | \mathcal{J}] + \nu_j = \beta X_j^* + \nu_j, \qquad j \in \{0, 1\}.$$

- We denote the expectational error as

$$\varepsilon_j = \beta X_j - \beta \mathcal{E}[X_j | \mathcal{J}],$$

and rewrite payoff function as

$$U_j = \beta X_j + \nu_j - \varepsilon_j.$$

- If $\mathcal{E}[\cdot] = \mathbb{E}[\cdot]$, then

$$\mathbb{E}[\varepsilon_j | X_j^*] = 0, \qquad \text{and} \qquad \mathbb{E}[\varepsilon_j | X_j] \neq 0.$$

Expectational and Measurement Error

- If $\mathcal{E}[\cdot] = \mathbb{E}[\cdot]$, then

$$\mathbb{E}[\varepsilon_j | X_j^*] = 0, \qquad \text{and} \qquad \mathbb{E}[\varepsilon_j | X_j] \neq 0.$$

- Rational expectations assumption $\Longrightarrow$ Errors-in-variables assumption.
- For any $Z \in \mathcal{J}$, $\mathbb{E}[\varepsilon_j | Z] = 0 \Longrightarrow$ Z is an IV.

Binary Choice Model

Agents' decisions

- Utility of individual i for any alternative $j \in \{0, 1\}$ is

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

- For any j, each dummy variable d_j is defined as

$$d_j = \mathbb{1}\{\Delta U_j \geq 0\}, \qquad \Delta U_j = U_j - U_{j'}$$

- Therefore, for any $j \in \{0, 1\}$, we can write the individual revealed preference inequality as

$$d_j \cdot (\beta \Delta X_j^* + \Delta \nu_j) \geq 0.$$

- The term $\beta \Delta X_j^*$ is the index function and β is the parameter we want to identify and estimate. We assume this index function is **linear in covariates**.

Binary Choice Model

Measurement model

- Without loss of generality, we can write

$$\beta \Delta X_j^* = \beta_1 \Delta X_{1j}^* + \beta_2 \Delta X_{2j}^*,$$

where $\beta_1 \Delta X_{1j}^*$ is measured without error,

$$\beta_1 \Delta Z_{1j} = \beta_1 \Delta X_{1j}^*,$$

and $\beta_2 \Delta X_{2j}^*$ is measured with error,

$$\beta_2 \Delta X_j = \beta_2 \Delta X_{2j}^* + \Delta \varepsilon_j.$$

Binary Choice Model

Measurement model

- Therefore, we can write the individual revealed preference inequality as

$$d_j \cdot (\beta_1 \Delta Z_{1j} + \beta_2 \Delta X_j + \Delta \nu_j + \Delta \varepsilon_j) \geq 0.$$

- In addition, we observe a vector ΔZ_{2j} and denote $\Delta Z_j = (\Delta Z_{1j}, \Delta Z_{2j})$.

Binary Choice Model

Assumptions

Assumption 1 *The random variable $\Delta\nu_j$ is independent of the random vector $(\Delta Z_j, \Delta X_j^*)$:*

$$F_\nu(\Delta\nu_j|(\Delta Z_j, \Delta X_j^*)) = F_\nu(\Delta\nu_j).$$

Assumption 2 *The marginal distribution function of $\Delta\nu_j$ is known up to a scale parameter, log concave, has mean zero, and, for any y in the support of $\Delta\nu_j$, verifies the following property:*

$$\frac{\partial^2 \mathbb{E}[\Delta\nu_j|\Delta\nu_j \geq y]}{\partial y^2} \geq 0.$$

Assumption 3 *The distribution of $\Delta\varepsilon_j$ conditional on $(\Delta X_j^*, \Delta Z_j, \Delta\nu_j)$ has support $(-\infty, \infty)$ and mean zero:*

$$\mathbb{E}[\Delta\varepsilon_j|\Delta X_j^*, \Delta Z_j, \Delta\nu_j] = 0.$$

Binary Choice Model

Comments on Assumptions

- Assumption 1 imposes that the endogeneity problem is exclusively due to measurement error.
  - It excludes models with random coefficients.
- Both the normal and the logistic distribution verify Assumptions 1 and 2.
  - Our model generalizes the probit and logit model to allow for classical measurement error in covariates.

Binary Choice Model

Comments on Assumptions

- Assumption 3 imposes the classical error-in-variables assumption.
  - It does not impose a parametric assumption on the marginal distribution of the measurement error.
  - It does not require full independence between the measurement error and the vector $(\Delta X_j^*, \Delta Z_j, \Delta \nu_j)$
- If $\Delta\varepsilon_j$ captures expectational error and we assume rational expectations, the only additional restriction that Assumption 3 imposes is that $\Delta Z_{2j} \in \mathcal{J}$.

Binary Choice Model

Identification

- Data are informative about the conditional density of $(d_j, \Delta X_j)$ given ΔZ_j.

$$\begin{aligned}
\mathcal{P}(d_j, \Delta X_j|\Delta Z_j) &= \int_{X^*} f(d_j, \Delta X_j|\Delta Z_j, \Delta X_j^*; \beta) f(\Delta X_j^*|\Delta Z_j) d\Delta X_j^*, \\
&= \int_{X^*} f(d_j|\Delta X_j^*; \beta) f(\Delta X_j|\Delta Z_j, \Delta X_j^*) f(\Delta X_j^*|\Delta Z_j) d\Delta
\end{aligned}$$

Binary Choice Model

Identification

- With

$$f(d_j|\Delta X_j^*;\beta) = \int_{\Delta\nu} \mathbb{1}\{\Delta\nu_j \geq -\beta\Delta X_j^*\} f(\Delta\nu_j) d\Delta\nu_j,$$

and

$$f(\Delta X_j^*|\Delta Z_j) = \frac{f(\Delta Z_j|\Delta X_j^*) f(\Delta X_j^*)}{\int_{\Delta X^*} f(\Delta Z_j|\Delta X_j^*) f(\Delta X_j^*) d\Delta X_j^*}.$$

- Assumptions 1 to 3 impose no restriction on $f(\Delta Z_j|\Delta X_j^*)$ and $f(\Delta X_j^*)$.
- Assumption 3 restricts the mean of $f(\Delta X_j|\Delta Z_j, \Delta X_j^*)$.

Binary Choice Model

Identification

- The parameter vector β is set-identified, not point-identified.
- Intuition:
  - Given a distribution $f(\Delta\nu_j)$, the parameter vector β captures choice sensitivity to differences in the covariates ΔX_j^*.
  - If ΔX_j^* was observed, this sensitivity could be measured directly and β would be point identified.
  - However, ΔX_{2j}^* is unobserved: the econometrician only observes the vector ΔX_j, which is known to have the same mean as ΔX_j^*.

- Intuition: (continued)
  - This mean restriction is not enough for point identification. Example: If choices are more sensitive to variation in $\Delta X_j^{(1)}$ than in $\Delta X_j^{(2)}$, it could be because: (a) the β coefficient on $\Delta X_j^{(1)}$ is higher than that on $\Delta X_j^{(2)}$; or (b) the variance of $\Delta X_j^{(1)}$ conditional on $\Delta X_j^{*(1)}$ is lower than the variance of $\Delta X_j^{(2)}$ conditional on $\Delta X_j^{*(2)}$.
  - Even if ΔZ_{2j} is fully independent of $\Delta \varepsilon_j$, the parameter β is still set identified (Chesher (2010))

Simulation Exercise

- The utility function is:

$$U_j = \beta^{(1)} X_j^{*(1)} + \beta^{(2)} X_j^{*(2)} + \beta^{(3)} X_j^{*(3)} + \nu_j, \quad j = \{0, 1\},$$

with $\beta = (\mathbf{0.5}, \mathbf{0.5}, \mathbf{0.25})$ and

$$\nu_j \backsim \mathbb{N}(0, \sigma_\nu^2), \qquad \sigma_\nu = 1.$$

Simulation Exercise

- The utility function is:

$$U_j = \beta^{(1)} X_j^{*(1)} + \beta^{(2)} X_j^{*(2)} + \beta^{(3)} X_j^{*(3)} + \nu_j, \quad j = \{0, 1\},$$

with $\beta = (\mathbf{0.5}, \mathbf{0.5}, \mathbf{0.25})$ and

$$\nu_j \backsim \mathbb{N}(0, \sigma_\nu^2), \qquad \sigma_\nu = 1.$$

- Both $X_j^{*(1)}$ and $X_j^{*(3)}$ are measured without error: $Z_{1j} = (X_j^{*(1)}, X_j^{*(3)})$.
- $X_j^{*(2)}$ is measured with error:

$$X_j = X_j^{*(2)} + \epsilon_j^x, \qquad \epsilon_j^x \backsim \mathbb{N}(0, \sigma_{\epsilon^x}^2), \qquad \sigma_{\epsilon^x}^2 = 0.2.$$

- We generate Z_{2j} as a second measurement of $X_j^{*(2)}$:

$$Z_{2j} = X_j^{*(2)} + \epsilon_j^z, \qquad \epsilon_j^z \backsim \mathbb{N}(0, \sigma_{\epsilon^z}^2), \qquad \sigma_{\epsilon^z}^2 = 0.02.$$

Conditional Moment Inequalities

- We first derive two types of moment inequalities conditional on the instrumental variable ΔZ_j.
- *Score Function Moment Inequalities*

$$\mathcal{M}_s(Z,j;\beta) = \mathbb{E}\left[d_j \frac{F_\nu\big(-(\beta_1\Delta Z_{1j} + \beta_2\Delta X_j)\big)}{1 - F_\nu\big(-(\beta_1\Delta Z_{1j} + \beta_2\Delta X_j)\big)} - d_{j'} \Big| Z \right] \geq 0.$$

- *Revealed Preference Moment Inequalities*

$$\mathcal{M}_r(Z,j;\beta) = \mathbb{E}\Big[d_j(\beta_1\Delta Z_{1j} + \beta_2\Delta X_j) + d_{j'}\mathbb{E}\big[\Delta\nu_{j'} | \Delta\nu_{j'} \geq -(\beta_1\Delta Z_{1j'} + \beta_2\Delta X_{j'})\big] \Big| Z \Big] \geq 0.$$

- For any Z and j, $\mathcal{M}_s(Z,j;\beta^*) \geq 0$ and $\mathcal{M}_r(Z,j;\beta^*) \geq 0$.

Unconditional Moment Inequalities

- The derive the same two types of unconditional inequalities.
- *Score Function Moment Inequalities*

$$\mathcal{M}_s^q(\beta) = \mathbb{E}\left[\sum_{j\in\{0,1\}} \left\{\Psi_q(\Delta Z_j)\left(d_j \frac{F_\nu\big(-(\beta_1\Delta Z_{1j} + \beta_2\Delta X_j)\big)}{1 - F_\nu\big(-(\beta_1\Delta Z_{1j} + \beta_2\Delta X_j)\big)} - d_{j'}\right)\right.\right.$$

- *Revealed Preference Moment Inequalities*

$$\mathcal{M}_r^q(\beta) = \mathbb{E}\left[\sum_{j\in\{0,1\}} \left\{\Psi_q(\Delta Z_j)\left(d_j(\beta_1\Delta Z_{1j} + \beta_2\Delta X_j)\right.\right.\right.$$

$$\left. + d_{j'}\mathbb{E}\big[\Delta\nu_{j'}|\Delta\nu_{j'} \geq -(\beta_1\Delta Z_{1j'} + \beta_2\Delta X_{j'})\big]\right)$$

- The set of functions $\{\Psi_q(\Delta Z_j),\ q \in Q\}$ groups different values of ΔZ_j into different unconditional moment inequalities. We call them *instrument functions*.

Instrument Functions

- As an example, in the particular case in which ΔZ_j is a 2×1 vector, the matrix Q defines 4 instrument functions:

$$\Psi_1(\Delta Z_j) = \mathbb{1}\{\Delta Z_{1j} \geq 0\}\mathbb{1}\{\Delta Z_{2j} \geq 0\},$$
$$\Psi_2(\Delta Z_j) = \mathbb{1}\{\Delta Z_{1j} \geq 0\}\mathbb{1}\{\Delta Z_{2j} < 0\},$$
$$\Psi_3(\Delta Z_j) = \mathbb{1}\{\Delta Z_{1j} < 0\}\mathbb{1}\{\Delta Z_{2j} \geq 0\},$$
$$\Psi_4(\Delta Z_j) = \mathbb{1}\{\Delta Z_{1j} < 0\}\mathbb{1}\{\Delta Z_{2j} < 0\}.$$

Identified Sets

- The identified sets defined by these sets of inequalities are

$$\Omega(\mathcal{M}_s) = \{\beta \in \Gamma_\beta : \mathcal{M}_s^q(\beta) \geq 0, q \in Q\},$$
$$\Omega(\mathcal{M}_r) = \{\beta \in \Gamma_\beta : \mathcal{M}_r^q(\beta) \geq 0, q \in Q\},$$
$$\Omega(\mathcal{M}) = \{\beta \in \Gamma_\beta : (\mathcal{M}_s^q(\beta); \mathcal{M}_r^q(\beta)) \geq 0, q \in Q\}.$$

Properties of the Identified Sets

All the following properties condition on Assumptions 1, 2, and 3.

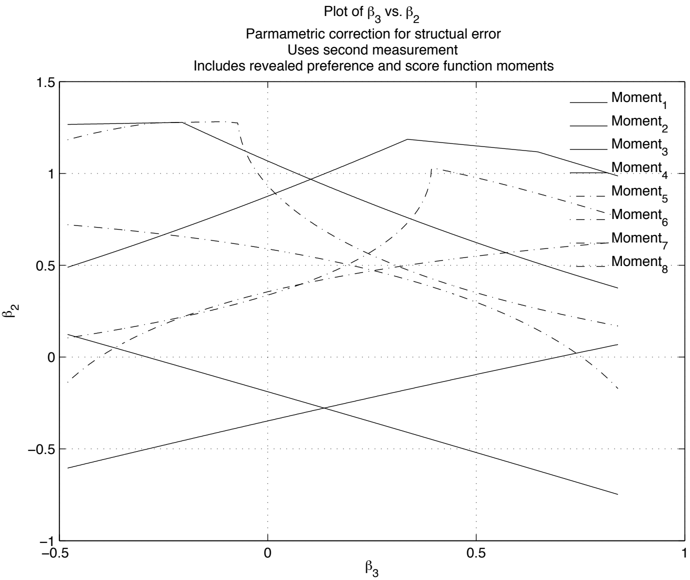
1. $\Omega(\mathcal{M}) \subseteq \Omega(\mathcal{M}_s)$ and $\Omega(\mathcal{M}) \subseteq \Omega(\mathcal{M}_r)$;
2. $\beta^* \in \Omega(\mathcal{M})$;
3. If $\bar{\bar{\sigma}}_\varepsilon^2 \geq \bar{\sigma}_\varepsilon^2$, then, $\Omega(\mathcal{M}_s|\bar{\sigma}_\varepsilon^2) \subseteq \Omega(\mathcal{M}_s|\bar{\bar{\sigma}}_\varepsilon^2)$, $\Omega(\mathcal{M}_r|\bar{\sigma}_\varepsilon^2) \subseteq \Omega(\mathcal{M}_r|\bar{\bar{\sigma}}_\varepsilon^2)$, and $\Omega(\mathcal{M}|\bar{\sigma}_\varepsilon^2) \subseteq \Omega(\mathcal{M}|\bar{\bar{\sigma}}_\varepsilon^2)$
4. $\Omega(\mathcal{M}_s)$, $\Omega(\mathcal{M}_r)$, and $\Omega(\mathcal{M})$ are convex.
5. If $\Delta X_j = \Delta X_{2j}^*$; and, for every $q \in Q$ and
$$\mathbb{E}\left[\sum_{j\in\{0,1\}} \Psi_q(\Delta Z_j)\right] \neq 0,$$
then $\Omega(\mathcal{M}_s) = \Omega(\mathcal{M}) = \beta^*$.
6. If $var(X - Z) \approx 0$, then $\Omega(\mathcal{M}_s)$, $\Omega(\mathcal{M}_r)$, and $\Omega(\mathcal{M})$ are all closed and bounded.

Simulation Exercise: Moment Inequalities Estimator

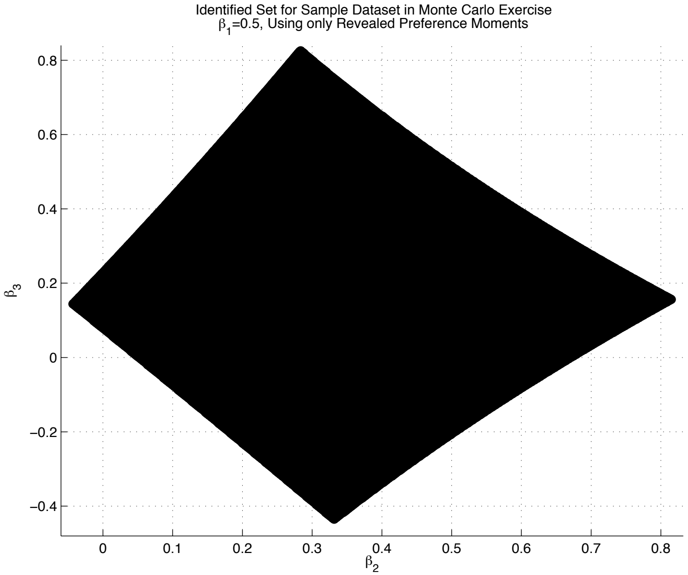
Observe X2 with Measurement Error, Use Z as IV

Case	min(b1)	max(b1)	min(b2)	max(b2)	min(b3)	max(b3)
No structural error correction	0.50	0.50	-0.11	0.24	-0.21	0.28
Parametric structural error, revealed pref moments	-0.47	1.17	-0.21	0.90	-0.61	0.95
Parametric structural error, score function moments	0.44	0.56	0.48	0.56	0.20	0.31
Parametric structural error, all moments	0.44	0.56	0.48	0.56	0.20	0.31
Non-parametric structural error	-1.65	5.37	-0.83	3.87	-2.28	4.14

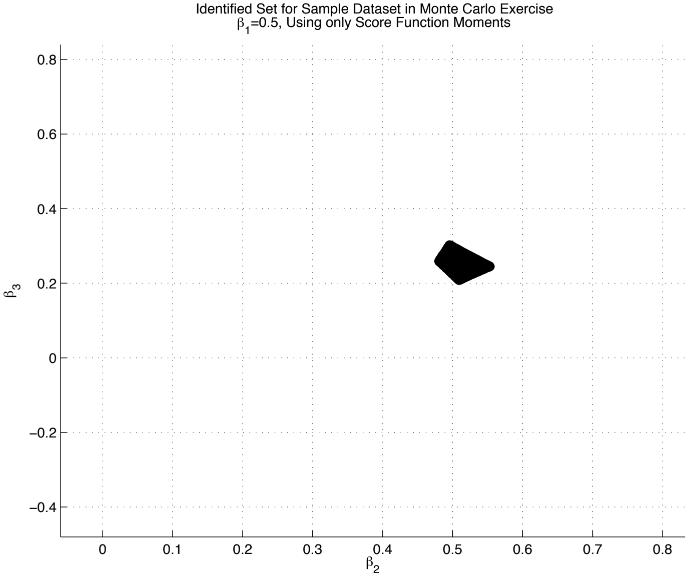
Simulation Exercise: Moment Inequalities Estimator



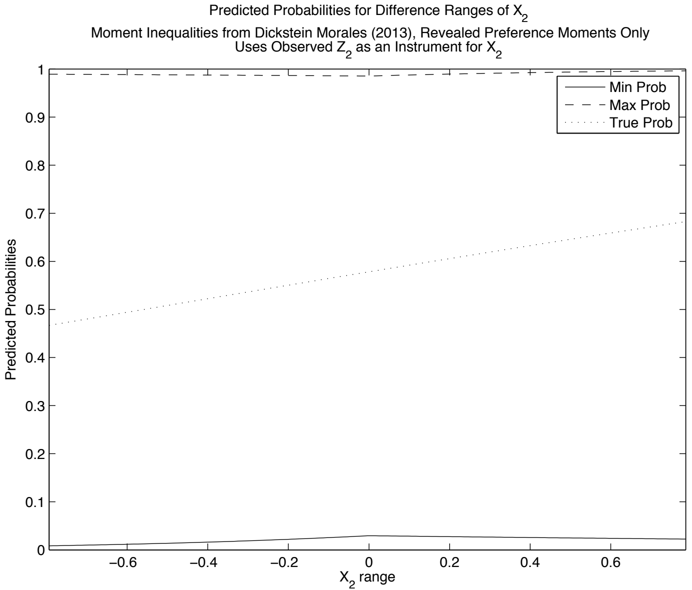
Simulation Exercise: Moment Inequalities Estimator



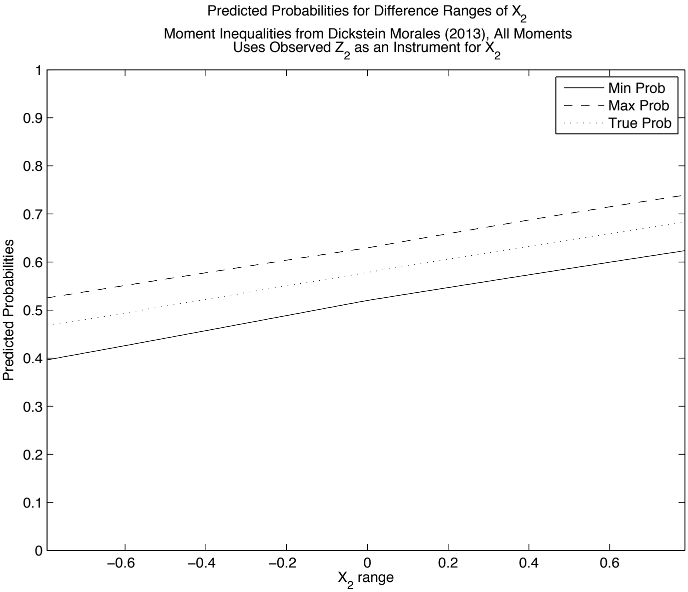
Simulation Exercise: Moment Inequalities Estimator



Simulation Exercise: Moment Inequalities Estimator



Simulation Exercise: Moment Inequalities Estimator



Misspecifications of the Model: Endogenous Instrument

- Substitute Assumption 3 by Assumption 3(b):

Assumption 3(b) *The distribution of* $\Delta\varepsilon_j$ *conditional on* $(\Delta X_j, \Delta\nu_j)$ *has support equal to* $(-\infty, \infty)$ *and expectation equal to 0:*

$$\mathbb{E}[\Delta\varepsilon_j|\Delta X_j, \Delta\nu_j] = 0.$$

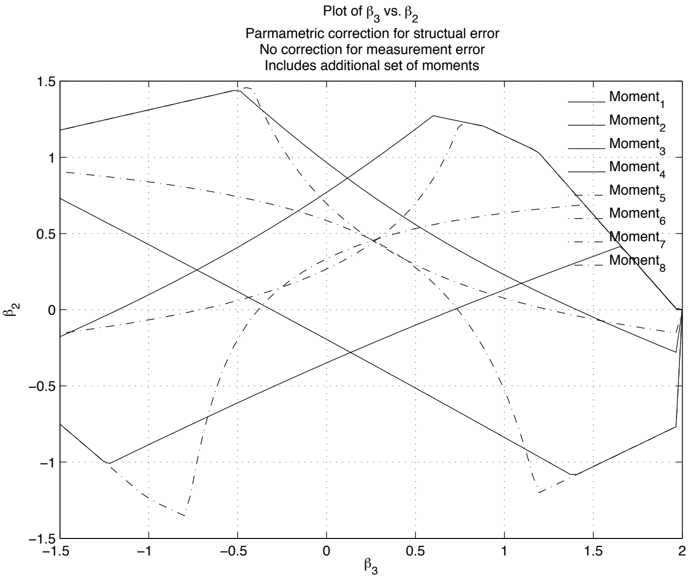
- Then we can derive the following inequalities:

$$\mathbb{M}_s^q(\beta) = \mathbb{E}\left[\sum_{j\in\{0,1\}}\left\{\Psi_q(\Delta X_j)\left(d_j\frac{F_\nu(-\beta\Delta X_j)}{1-F_\nu(-\beta\Delta X_j)} - d_{j'}\right)\right\}\right] \geq 0,$$

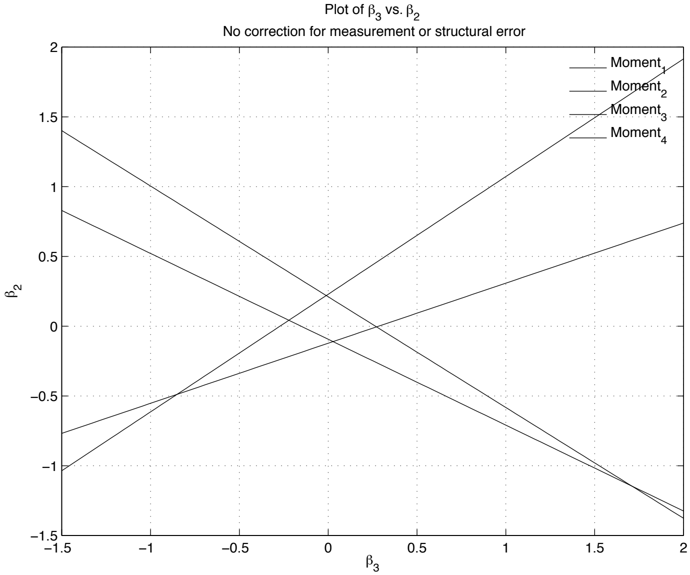
$$\mathbb{M}_r^q(\beta) = \mathbb{E}\left[\sum_{j\in\{0,1\}}\left\{\Psi_q(\Delta X_j)\left(d_j\beta\Delta X_j + d_{j'}\mathbb{E}[\Delta\nu_{j'}|\Delta\nu_{j'} \geq -\beta\Delta X_{j'}]\right)\right.\right.$$

- If Assumption 3(b) does not hold, what are the properties of these inequalities?

Misspecifications of the Model: Endogenous Instrument



Misspecifications of the Model: Endogenous Instrument



Misspecifications of the Model: No Structural Error

- Substitute Assumption 2 by Assumption 2(b):

  Assumption 2(b) *The marginal distribution function of $\Delta\nu_j$ has a degenerate distribution at $\Delta\nu_j = 0$.*

- If Assumptions 1, 2(b), and 3 hold, then the score function moment inequality is irrelevant.
- The revealed preference inequality becomes

$$\mathfrak{M}_r^q(\beta) = \mathbb{E}\left[\sum_{j\in\{0,1\}} \left\{\Psi_q(\Delta Z_j) d_j (\beta_1 \Delta Z_{1j} + \beta_2 \Delta X_j)\right\}\right] \geq 0,$$

Misspecifications of the Model: No Structural Error

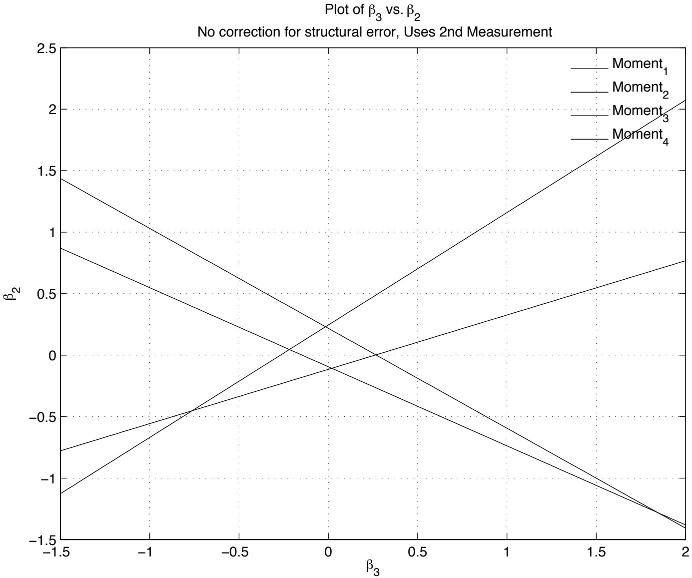
- Note that

$$\mathcal{M}_r^q(\beta) = \mathfrak{M}_r^q(\beta) +$$

$$\mathbb{E}\left[\sum_{j \in \{0,1\}} \left\{ \Psi_q d_{j'} \mathbb{E}\left[\Delta \nu_{j'} | \Delta \nu_{j'} \geq -(\beta_1 \Delta Z_{1j'} + \beta_2 \Delta X_{j'})\right] \right\}\right].$$

- If Assumption 2(b) does not hold, what are the properties of these inequalities?

Misspecifications of the Model: No Structural Error



Misspecifications of the Model: No Structural Error