Príliš silné predpoklady v ekonomických modeloch?

Analýza Parciálnej Identifikácie

Lukáš Lafférs

Department of Economics, NHH Norwegian School of Economics

http://sites.google.com/site/lukaslaffers

18.12.2012

This presentation is about Identification.

What can we learn from **data** and **assumptions**?

What drives our results?

Which assumptions are important?

Which assumptions are not important?

$$\text{Assumptions} + \text{Data} \rightarrow \text{Results}$$

Charles Manski: *The Law of Decreasing Credibility:* The credibility of inference decreases with the strength of the assumptions maintained.

Example - regression analysis

House prices:

y_i - price of a house

x_i - characteristics (area, number of bedrooms, location...)

A model (set of assumptions)

- $y_i = x_i'\theta + u_i$,
- (y_i, x_i) are iid,
- $E(y_i|x_i) = x_i'\theta$,

Suppose P is a distribution of (y_i, x_i).

The model is correctly specified if $\exists\theta$ such that $E_P(y_i|x_i) = x_i'\theta$ with P-probability 1.

A question of identification is a question about existence of such θ.

Object of interest

We are interested in an **identified set**: a set of all parameters that are compatible with observed data and assumptions that constitute the economic model.

- Koopmans and Reiersol (1950): "Scientific honesty demands that the specification of a model be based on prior knowledge of the phenomenon studied and possibly on criteria of simplicity, but not on the desire for identifiability of characteristics that the researcher happens to be interested in"
- Phillips (1989): "(it) seems important that we should understand the implications of identification failure for statistical inference. Yet, this is a subject that seems to be virtually untouched in the literature."

Applications

- Haile and Tamer (2003): "Inference with an incomplete model of English auctions." **Journal of Political Economy** 111(1):1–51,
- Blundell, Gosling, Ichimura and Meghir (2007): "Changes in the distribution of male and female wages accounting for employment composition." **Econometrica** 75:323–63,
- Kline, Tamer (2012): Bounds for best response functions in binary games **Journal of Econometrics** 166(1):92-105.

Setup and Notation

Potential outcomes framework of Rubin (1974).

Individual i: $y_i(.) : T \to Y$ (individuals do not interact)

(Potential) Treatment: $t \in T$ (mutually exclusive and exhaustive)

(Potential) Outcome: $y_i(t) \in Y$

Realized treatment: $z_i \in T$

Realized outcome: $y_i \equiv y_i(z_i) \in Y$

Instrument: $v_i \in V$

Fundamental problem: $y_i(t)$ is not observed for $t \neq z_i$

Distribution of (y_i, z_i, v_i) is observed.

Our Goal

Learn about the probability distribution of counter-factual outcomes
$P(y(t_1), y(t_2), \ldots, y(t_m))$

What are we interested in?

- $E[y(t_3)]$ - average treatment response
- $E[y(t_4)] - E[y(t_2)]$ - average treatment effect

Examples

- Effect of parental schooling on child's schooling
- Effectiveness of a labor participation program
- Effect of a medical intervention

Assumptions have to made in order to learn something about properties of an unobserved counter-factual distribution.

These assumption may or may not be strong enough be **point** identify the quantity of interest.

If only weak assumptions are made, the quantity of interest may be **partially** identified.

Examples (Manski)

Say we are interested in $E[y(t)]$

$$E[y(t)] = E[y|z = t].P(z = t) + E[y(t)|z \neq t].P(z \neq t)$$

Example

Say we are interested in $E[y(t)]$

$$E[y(t)] = E[y|z = t].P(z = t) + E[y(t)|z \neq t].P(z \neq t)$$

Observed quantities

Unobserved quantities

Example (exogenous selection)

If we assume that $E[y(t)|z = t] = E[y(t)|z \neq t]$

$$\begin{aligned} E[y(t)] &= E[y|z = t].P(z = t) + E[y(t)|z \neq t].P(z \neq t) \\ &= E[y|z = t].P(z = t) + E[y|z = t].P(z \neq t) \\ &= E[y|z = t] \end{aligned}$$

Under this assumption, $E[y(t)]$ is **point** identified.

Example (bounded support)

Suppose that $y_{min} \leq y_i(t) \leq y_{max}$

$$
\begin{array}{rcl}
LB_{E[y(t)]} & = & E[y|z=t].P(z=t) + y_{min}.P(z \neq t) \\
& \leq & \\
E[y(t)] & = & E[y|z=t].P(z=t) + E[y(t)|z \neq t].P(z \neq t) \\
& \leq & \\
UB_{E[y(t)]} & = & E[y|z=t].P(z=t) + y_{max}.P(z \neq t)
\end{array}
$$

Under this assumption, $E[y(t)]$ is **partially** identified and the interval $(LB_{E[y(t)]}, UB_{E[y(t)]})$ is called an **identified set**.

Different assumptions

Suppose we are interested in an effect of mother's education of child's education (deHaan 2011).

- Monotone Treatment Response (MTR) assumption
  $\forall i, t_2 \geq t_1 : y_i(t_2) \geq y_i(t_1)$
- Monotone Treatment Selection (MTS) assumption
  $\forall z_2 \geq z_1 : E[y(t)|z = z_2] \geq E[y(t)|z = z_1]$
- Monotone Instrumental Variable (MIV) assumption
  $\forall v_2 \geq v_1 : E[y(t)|v = v_2] \geq E[y(t)|v = v_1]$

Analytical bounds on $E[y(t)]$ under MTR, MTS, MIV, MTR+MTS are available. These then translate to bounds on $E[y(t)] - E[y(s)]$.

My contribution

- I introduce a method to trace the bounds on ATE in a straightforward manner.
- I can also replicate a few existing results.
- I can relax identifying assumptions and see how "important" they are. How robust are the results to a deviation from these identifying assumptions.
- I am flexible when modeling assumptions, adding or changing assumptions is trivial.

My contribution (2)

I replicated results of these papers:

- Shaikh and Vytlacil (2011): "Partial Identification in Triangular Systems of Equations With Binary Dependent Variables," **Econometrica**, 79, 949–955,
- Chesher (2010): "Instrumental Variable Models for Discrete Outcomes," **Econometrica** 78, 575–601,
- Manski and Pepper (2000): "Monotone Instrumental Variables, with an Application to the Returns to Schooling," **Econometrica**, 68, 997–1012.
- Balke and Pearl (1997): "Bounds on treatment effects from studies with imperfect compliance," **Journal of the American Statistical Association**, 439, 1172–1176,
- Manski (1997): "Monotone Treatment Response," **Econometrica**, 65, pp. 1311–1334,

Each of these papers provides a non-trivial proof of an identified set.

My identification method

- Builts upon ideas of Galichon and Henry (2009, 2010) and Ekeland, Galichon and Henry (2011).
- Each economic model involves observed and unobserved variables.
- The essence of my identification strategy is a search in a space of joint distributions of observed and unobserved component.
- Every assumption is basically a restriction on the space of joint distribution functions.
- Compatibility with observed data also translates into a set of restrictions on the space of joint dist. functions.

My identification method

- If for a particular set of assumptions and particular data there exist a joint distribution, then we cannot refute the economic model.
- On the other hand, if we cannot find such joint distribution, the economic model cannot be compatible with observed data.
- If the restrictions are linear in the joint distribution function (often) the search in space of joint distributions is a linear program.

Limitations

- Observed variables must be discrete
- Curse of dimensionality problem
- General statistical inference method not readily available

Now I demonstrate my identification method on a particular example.

Demonstration of my identification method

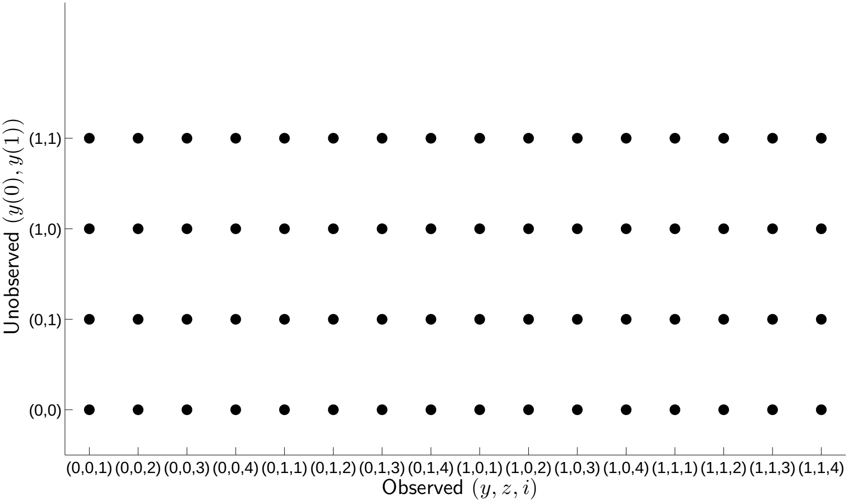
- $y_i \in Y = \{0, 1\}$ - child's college (0 - no college, 1 - college),
- $z_i \in Z = \{0, 1\}$ - mother's college (0 - no college, 1 - college)
- $v_i \in V = \{1, 2, 3, 4\}$ - father's schooling level (High school or less ($\leq$ 12 years), Some college (13-15 years), Bachelor degree (16 years), Master degree or more ($\geq$ 17 years)),

How does the search in the space of joint distributions of observed and unobserved component look like?

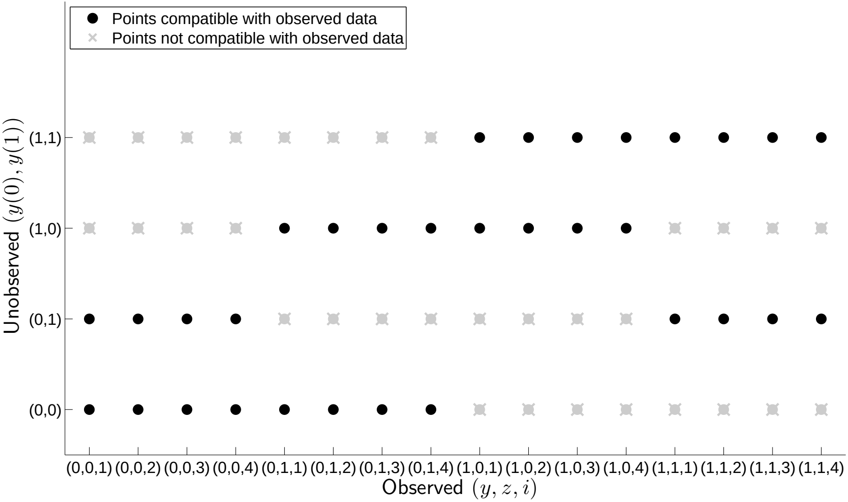
Suppose we are interested in an effect of mother's education of child's education (deHaan 2011).

- Monotone Treatment Response (MTR) assumption
  $\forall i, t_2 \geq t_1 : y_i(t_2) \geq y_i(t_1)$
- Monotone Treatment Selection (MTS) assumption
  $\forall z_2 \geq z_1 : E[y(t)|z = z_2] \geq E[y(t)|z = z_1]$
- Monotone Instrumental Variable (MIV) assumption
  $\forall v_2 \geq v_1 : E[y(t)|v = v_2] \geq E[y(t)|v = v_1]$

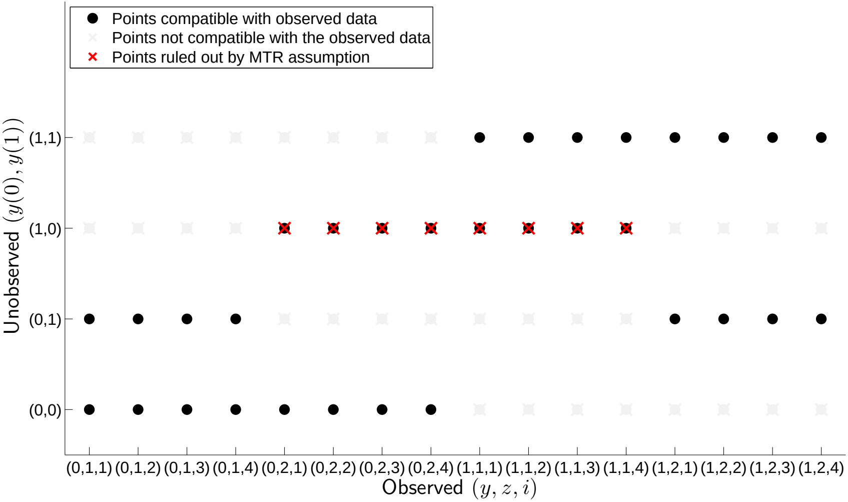
The Joint Support of $(y(0), y(1), y, z, v)$



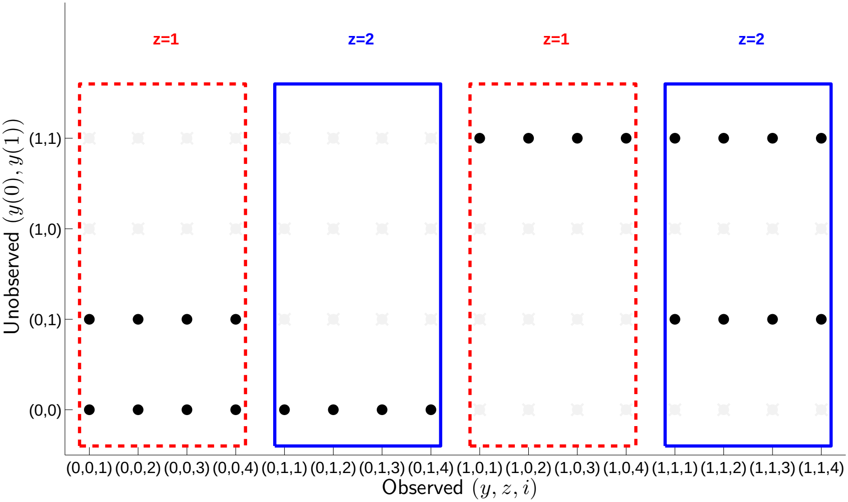
Compatibility with Observed Data: $\forall i, t : z_i = t \Rightarrow y_i = y_i(t)$



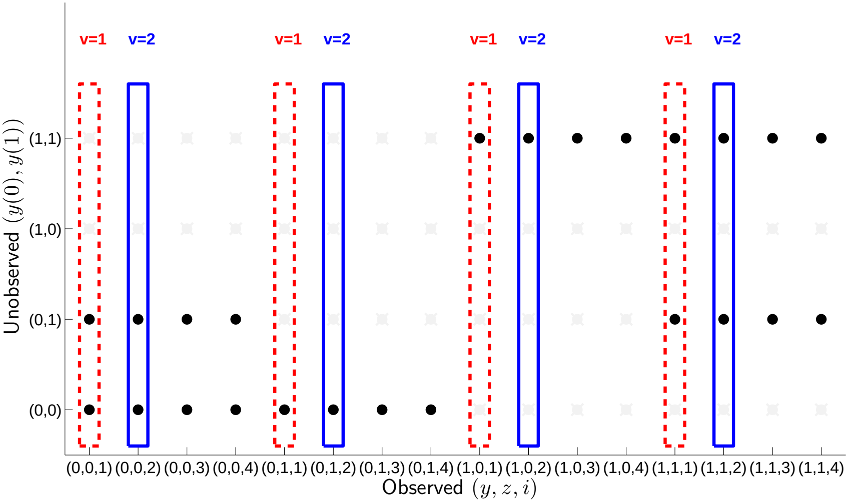
Monotone Treatment Response: $\forall i : y_i(0) \leq y_i(1)$



Monotone Treatment Selection: $E[y(t)|z=2] \geq E[y(t)|z=1]$



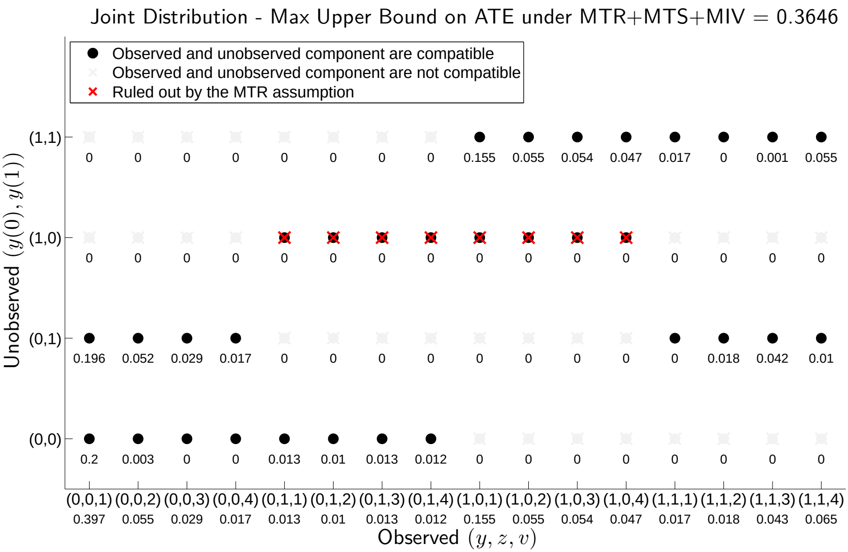
Monotone Instrumental Variable: $E[y(t)|v=1] \geq E[y(t)|v=0]$



Results

Bounds on Effect of Mother's College Increase on the Probability of Child Has College Degree		
Setup	Assumptions	[Lower Bound, Upper Bound]
LinProg (this paper)	No Assumptions	[-0.358, 0.641]
LinProg (this paper)	MTS	[-0.358, 0.365]
LinProg (this paper)	MTR	[0, 0.641]
LinProg (this paper)	MTR + MTS	[0, 0.365]
LinProg (this paper)	MTR + MTS + MIV	[0, 0.365]
Note: Estimates are not bias corrected, $n = 16912$		

In the empirical application



Corresponding Linear Program

$$\max_{\pi} \overbrace{\begin{bmatrix} 0&1&0&1&0&1&0&1&0&0&0&0&0&0&0&0&1&0&1&0&1&0&1&0 \end{bmatrix}}^{\text{Average Treatment Effect}} \times \pi$$

subject to

$$\text{DATA} \left\{ \begin{bmatrix}
1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1
\end{bmatrix} \times \pi = \begin{bmatrix} 0.397 \\ 0.055 \\ 0.029 \\ 0.017 \\ 0.013 \\ 0.01 \\ 0.013 \\ 0.012 \\ 0.155 \\ 0.055 \\ 0.054 \\ 0.047 \\ 0.017 \\ 0.018 \\ 0.043 \\ 0.065 \end{bmatrix} \right\} \begin{array}{c}\text{Observed} \\ \text{probabilities}\end{array}$$

$$\begin{array}{l}
\text{MTS} \left\{ \begin{bmatrix}
0&0&0&0&0&0&0&0&0&0&0&0&.19&.19&.19&.19&0&-.80&0&-.80&0&-.80&0&-.80 \\
0&.19&0&.19&0&.19&0&.19&0&0&0&0&.19&.19&.19&.19&-.80&-.80&-.80&-.80&-.80&-.80&-.80&-.80
\end{bmatrix} \right. \\
\text{MIV} \left\{ \begin{bmatrix}
0&0&0&0&0&0&0&0&0&0&0&0&.13&-.58&0&0&0&.13&0&-.58&0&0&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&.13&-.13&0&0&0&0&.13&0&-.13&0&0 \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&.14&-.13&0&0&0&0&0&.14&0&-.13 \\
0&.13&0&-.58&0&0&0&0&0&0&0&0&.13&-.58&0&0&.13&.13&-.58&-.58&0&0&0&0 \\
0&0&0&.13&0&-.13&0&0&0&0&0&0&0&.13&-.13&0&0&0&.13&.13&-.13&-.13&0&0 \\
0&0&0&0&0&.14&0&-.13&0&0&0&0&0&0&.14&-.13&0&0&0&0&.14&.14&-.13&-.13
\end{bmatrix} \right.
\end{array} \times \pi \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\pi \geq \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

Venues for further research

Relax different assumptions and see their identification strength.

- Mis-measurement of Outcomes or Treatments (MOT): $P[z_i = t \Rightarrow y_i = y_i(t)] \geq 1 - \alpha_{MOT}$,
- Relaxed Monotone Treatment Response (rMTR): $P[t_2 \geq t_1 \Rightarrow y_i(t_2) \geq y_i(t_1)] \geq 1 - \alpha_{MTR}$,
- Relaxed Monotone Treatment Selection (rMTS): $\forall z_2 \geq z_1 : E[y(t)|z = z_1] - E[y(t)|z = z_2] \leq \delta_{MTS}$,
- Relaxed Monotone Instrumental Variable (rMIV): $\forall v_2 \geq v_1 : E[y(t)|v = v_1] - E[y(t)|v = v_2] \leq \delta_{MIV}$.

Introduce new assumptions

- Monotone Selection Bias (MSB): $\forall z_2 \geq z_1, v_2 \geq v_1 : E[y(t)|z = z_1, v = v_2] - E[y(t)|z = z_2, v = v_2] \geq E[y(t)|z = z_1, v = v_1] - E[y(t)|z = z_2, v = v_1]$

Thank you for your attention!