# Inference for Partially Identified Models

Lukáš Lafférs

Department of Economics, NHH Norwegian School of Economics

April 9, 2013

Many interesting models are only partially identified, hence more than one model is supported by the data.

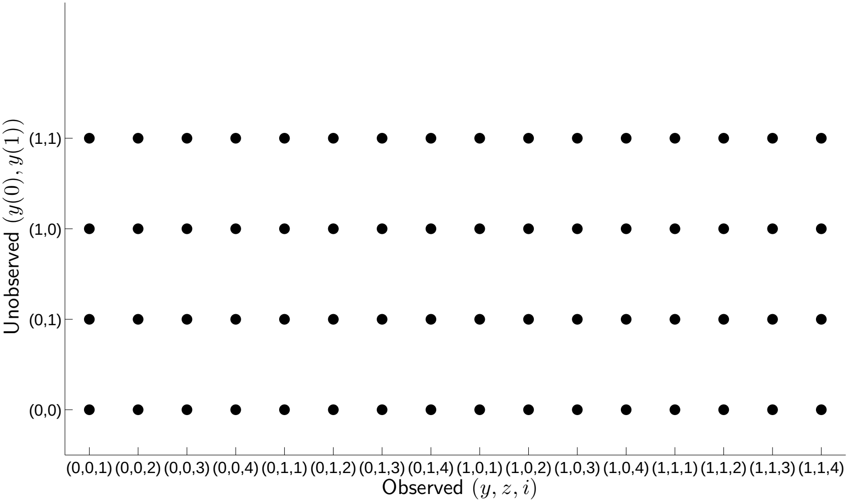
# Example: bounds on average treatment effects (Manski).

Under some reasonable non-parametric assumptions, the college increase of a mother increases the probability that child gets college degree by $0 - 36.5\%$.

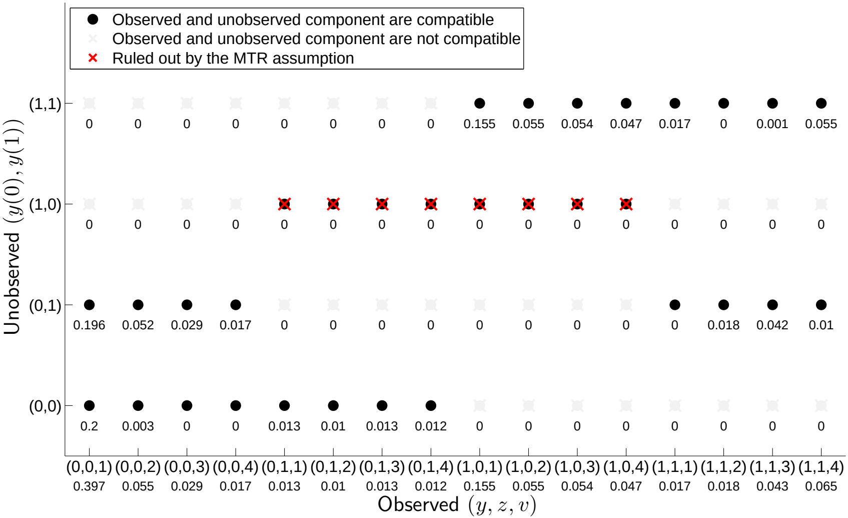
My previous work point out that in a rich class of problems that are partially identified, finding the tightest identified region is equivalent to solving a particular linear program.

The key is the search in the space of probability distribution functions.

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



# Joint Distribution - Max Upper Bound on ATE under MTR+MTS+MIV = 0.3646



# Example of the 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

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

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

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

# The Major Challenge

How to do statistical inference.

How to take into account the fact that we only have data sample of fixed length.

# What can be done

We wish to learn about $[L_{min}, L_{max}]$

- Confidence set for the true parameter
  $\inf_{\theta \in [L_{min}, L_{max}]} \liminf_{n \to \infty} Pr(\theta \subseteq C_n(1-\alpha)) \geq 1-\alpha$
- Confidence set for the whole identified set
  $\liminf_{n \to \infty} Pr([L_{min}, L_{max}] \subseteq C_n(1-\alpha)) \geq 1-\alpha$
- Bayes maximum posterior credible interval

# Setup

$[L_{min}, L_{max}]$ is the range of possible values of

$$(\min)\, \max_{\pi}\; c^T \pi$$

$$s.t$$

$$A_d \pi = p$$

$$A_s \pi \leq b_s$$

$$\pi \geq 0$$

depends on data

fixed

Original Form

$$(\min)\,\max_\pi\ c^T\pi$$
$$s.t$$
$$A_d\pi = p$$
$$A_s\pi \leq b_s$$
$$\pi \geq 0$$

depends on data

fixed

Standard Form

$$(\min)\,\max_\pi\ [c^T 0^T]\begin{bmatrix}\pi \\ v\end{bmatrix}$$
$$s.t$$
$$\begin{bmatrix}A_d & 0 \\ A_s & -I\end{bmatrix}\begin{bmatrix}\pi \\ v\end{bmatrix} = \begin{bmatrix}p \\ b_s\end{bmatrix}$$
$$\begin{bmatrix}\pi \\ v\end{bmatrix} \geq 0$$

depends on data

fixed

# Different Methods

This paper: compare the performance of different inferential methods and give some practical advice.

$$\max_{\pi} \bar{c}^T \bar{\pi}$$
$$s.t$$
$$\bar{A}\bar{\pi} = \bar{m}$$
$$\bar{\pi} \geq 0$$

- Bootstrap (canonical, bias corrected, accelerated, percentile)
- Subsampling (Romano and Shaikh 2012)
- Modified Bootstrap (Freyberger and Horowitz 2012)
- Intersection bounds (Chernozhukov, Lee and Rosen 2012)
- Robust Bayes (Kitagawa 2011)
- Projection (Laffers 2015?)

# My contribution

Given some recent advances in the development of inferential schemes, what inferential scheme should we use if we study bounds?

In other words: Can we do better than the regular bootstrap?

# Bootstrap

(introduced by Efron 1979)

Pros:

- very easy to implement

Cons:

- in some cases theoretically not justified
- e.g. when parameter is on the boundary of the identified set

# Subsampling

(Romano and Shaikh 2012)

Pros:

- works under minimal assumptions

Cons:

- less accurate than bootstrap (when it is consistent)
- very "data hungry"
- too many tuning parameters involve
  - subsample size
  - scaling sequence
  - number of subsamples

# Modified bootstrap

(Freyberger and Horowitz 2012)

Pros:

- designed for random linear program
- theoretically justified

Cons:

- relatively easy to implement

# Intersection Bounds

(Chernozhukov, Lee and Rosen 2012)

The identification region is defined as
$\left[\sup_{v \in V} \theta^l(v), \inf_{v \in V} \theta^u(v)\right]$

Pros:

- everything is data driven, no tuning parameters

Cons:

- how to transform my LP into this setup(?)
- implementation

## Robust Bayes

(Kitagawa 2011)

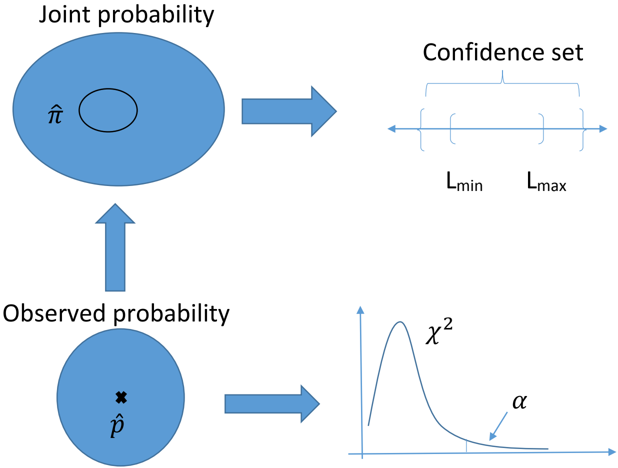
Pros:

- coherent decision-theoretic framework

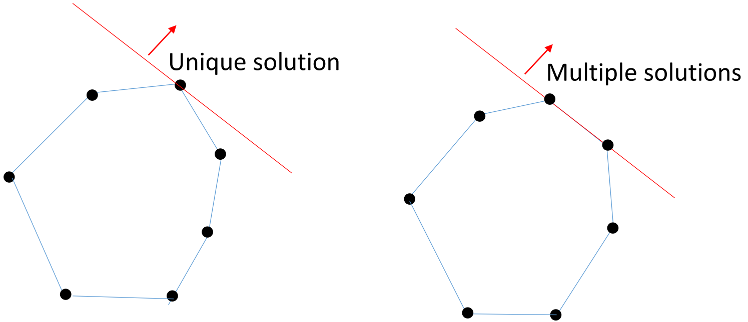
Cons:

- priors
- complex method to explain and sell further
- implementation
- does not address non-differentiability

# Projection (Illustration)



# A bit more on the Modified Bootstrap



To check for empirical coverage, I have to sample from the "true" model.

Thank you for your attention. Any comments are very welcome.

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

lukas.laffers@nhh.no