Identifying effects of multivalued treatments

Sokbae Lee & Bernard Salanié (Econometrica, 2018)

Patrick Molligo

ECON 232M

Spring 2022

Overview

- **Main problem**:
  - Empirical models require monotonicity of instrument-treatment variation
  - What if there are many selection criteria for one treatment?
  - Can we relax the monotonicity assumption?

Overview

- **Main problem**:
  - Empirical models require monotonicity of instrument-treatment variation
  - What if there are many selection criteria for one treatment?
  - Can we relax the monotonicity assumption?
- **Approach**:
  - Think of discrete choice random utility model
  - Only observe partial set of alternatives
  - Vector of threshold-crossing rules

Overview

- **Main problem**:
  - Empirical models require monotonicity of instrument-treatment variation
  - What if there are many selection criteria for one treatment?
  - Can we relax the monotonicity assumption?
- **Approach**:
  - Think of discrete choice random utility model
  - Only observe partial set of alternatives
  - Vector of threshold-crossing rules
- **Identification**:
  1. Generalized MTE (ATE, ATT, PRTE)
  2. Identify threshold functions

Overview

- **Main problem**:
  - Empirical models require monotonicity of instrument-treatment variation
  - What if there are many selection criteria for one treatment?
  - Can we relax the monotonicity assumption?
- **Approach**:
  - Think of discrete choice random utility model
  - Only observe partial set of alternatives
  - Vector of threshold-crossing rules
- **Identification**:
  1. Generalized MTE (ATE, ATT, PRTE)
  2. Identify threshold functions

Contribution

- Foundational framework: Angrist and Imbens (1995)
- MTE, discrete choice: Heckman, Urzua, and Vytlacil (2006, 2008)
- Multi-valued, unordered treatments: Dahl (2002), Kirkeboen, Leuven, and Mogstad (2016), Kline and Walters (2016), Mountjoy (2022)
- **Unordered monotonicity**: Heckman and Pinto (2018), Pinto (2015)

Model overview

Sample: $\{(Y_i, D_i, Z_i) : i = 1, \ldots, N\}$

Potential outcome $Y_k : k \in \mathcal{K}$, for treatment k

Selection mechanism:

- Vector of J unobserved random variables: $\{V_j : j \in J\}$
- Vector of **known** functions $\{Q_j(Z) : j \in J\}$
- Treatment function:

$$D_k = d_k(\mathbf{V}, \mathbf{Q}(\mathbf{Z})) = \sum_{l \in \mathcal{L}} c_l^k \prod_{j \in l} S_j(\mathbf{V}, \mathbf{Q}(\mathbf{Z}))$$

Simple example

Case with $J = 2$ and $K = 3$:

$D = 0$ iff $V_1 < Q_1(Z)$ and $V_2 < Q_2(Z)$

$D = 1$ iff $V_1 > Q_1(Z)$ and $V_2 > Q_2(Z)$

$D = 2$ iff $(V_1 - Q_1(Z))$ and $(V_2 - Q_2(Z))$ opposite signs

Simple example

Case with $J = 2$ and $K = 3$:

$D = 0$ iff $V_1 < Q_1(Z)$ and $V_2 < Q_2(Z)$

$D = 1$ iff $V_1 > Q_1(Z)$ and $V_2 > Q_2(Z)$

$D = 2$ iff $(V_1 - Q_1(Z))$ and $(V_2 - Q_2(Z))$ opposite signs

Treatment function with "index of treatment" c_J^k:

$$D_k(S) = c^k_{\{\emptyset\}} + c^k_{\{1\}} S_1 + c^k_{\{2\}} S_2 + c^k_{\{1,2\}} S_1 S_2$$

$D_0 = S_1 S_2$
$D_1 = (1 - S_1)(1 - S_2)$
$D_2 = S_1 + S_2 - 2 S_1 S_2$

Treatment status visualized

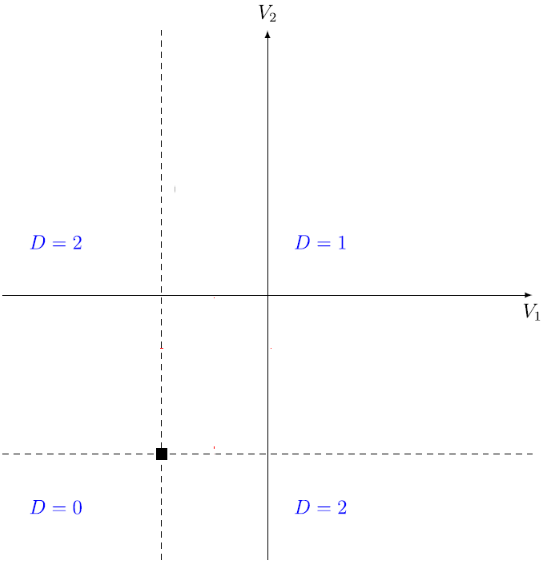

FIGURE 1.—Example 1.

Two-way flows

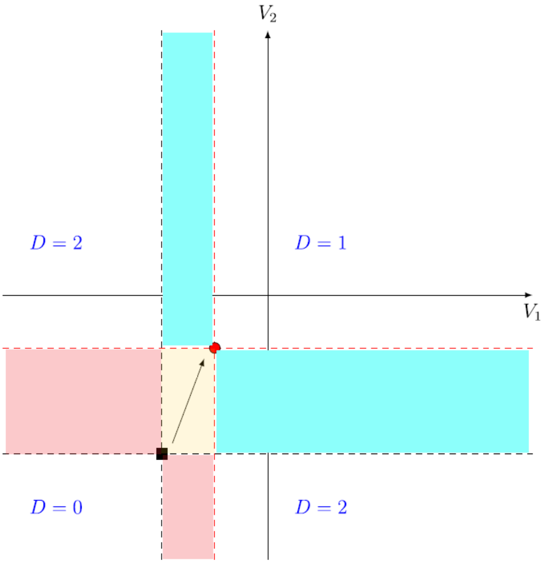

FIGURE 1.—Example 1.

New direction, same problem

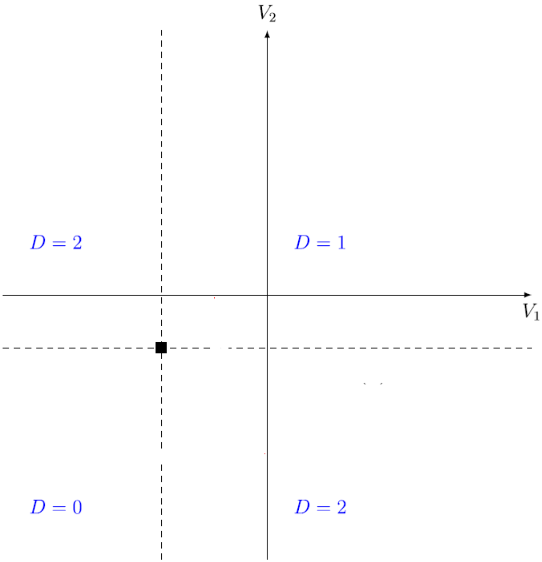

FIGURE 2.—Example 1 (continued).

New direction, same problem

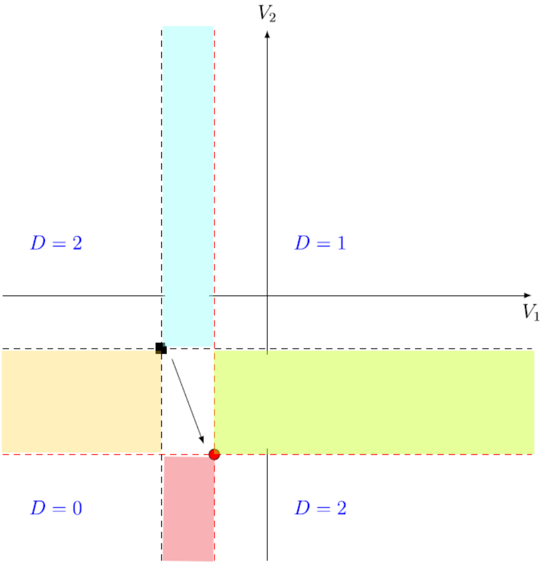

FIGURE 2.—Example 1 (continued).

Identification

Key Assumptions:

1. Selection mechanism
2. Conditional independence of instruments
3. Unobserved heterogeneity continuously distributed
4. Sufficient variation of $Q(z)$ when z varies

Identification

Given assumptions on previous slide:

$$f_V(q) = \frac{1}{c_J^k} TPr(D = k|Q(Z) = q)$$

$$E[G(Y_k)|V = q] = \frac{TE[G(Y)D_k|Q(Z) = q]}{TPr(D = k|Q(Z) = q)}$$

where $Th(q) \equiv \frac{\partial^J h}{\prod_{j=1}^J \partial q_j}(q)$

Thus the MTE between treatments k and l is given by:

$$E[G(Y_k)|V = v] - E[G(Y_l)|V = v]$$

Simple application: two-way flows

- Selection mechanism from motivating example
- Two instruments Z_1, Z_2
- Continuous density of (V_1, V_2)
- $Q_1(Z) \perp Z_2$ and $Q_2(Z) \perp Z_1$
- $k = 1, 2, 3$

$$f_{V_1,V_2}(q_1, q_2) = \frac{1}{c_J^k} \frac{\partial^2 Pr[D = k | Q_1(Z) = q_1, Q_2(Z) = q_2]}{\partial q_1 \partial q_2}$$

$$E[Y_k | V_1 = q_1, V_2 = q_2] = \frac{\frac{\partial^2 E[YD_k | Q_1(Z)=q_1, Q_2(Z)=q_2]}{\partial q_1 \partial q_2}}{\frac{\partial^2 Pr[D=k | Q_1(Z)=q_1, Q_2(Z)=q_2]}{\partial q_1 \partial q_2}}$$

Note: Q functions identified up to a constant $\implies$ identification up to location shift in (q_1, q_2)