Instrumental Variables, and the Sign of the Average Treatment Effect

Cecilia Machado, Columbia University,
Azeem Shaikh, University of Chicago,
and
Edward Vytlacil, Yale University

Cowles Lunch
February 25

Question of Interest

Under what conditions can we use instrumental variables to infer the sign of the Average Treatment Effect?

- We address this question while imposing alternative monotonicity/weak separability conditions.
- For exposition, we focus on very simple context:
  - binary outcome,
  - binary endogenous regressor (treatment)
  - binary instrument
  - no other covariates.

Notation

Notation:

- $Y \in \{0, 1\}$ outcome of interest.
- $D \in \{0, 1\}$ endogenous regressor (treatment) of interest.
- $Z \in \{0, 1\}$ instrument.
- Order Z such that $\Pr[D = 1|Z = 1] > \Pr[D = 1|Z = 0]$. (identified from data)

Examples

Example:

- Y mortality 12 months after start of study.
- D medical intervention.
- Z random assignment with imperfect compliance.

Example:

- Y employment at 12 months
- D receipt of job training at date 0.
- Z local labor market conditions at date 0.

Example:

- Y employment at age 30
- D college attendance.
- Z nearest university near or far at age 18.

Potential Outcomes, Choices

Define Potential Outcomes, Choices:

- Y_1 potential outcome if treated
- Y_0 potential outcome if not treated
- D_1 potential choice if instrument externally set to 1.
- D_0 potential choice if instrument externally set to 0.

Thus

$$Y = DY_1 + (1 - D)Y_0$$

$$D = ZD_1 + (1 - Z)D_0$$

Restrictions

Will maintain independence assumption

$$Z \perp\!\!\!\perp (Y_0, Y_1, D_0, D_1)$$

We additionally consider imposing each of following alternative conditions

(I) $D_1 \geq D_0$ w.p.1 or $D_1 \leq D_0$ w.p.1;

(II) $(D_1 \geq D_0$ or $D_1 \leq D_0)$ and $(Y_1 \geq Y_0$ or $Y_1 \leq Y_0)$

(III) $Y_1 \geq Y_0$ w.p.1 or $Y_1 \leq Y_0$ w.p.1

Restrictions

(I) $D_1 \geq D_0$ or $D_1 \leq D_0$;
(II) $(D_1 \geq D_0$ or $D_1 \leq D_0)$ and $(Y_1 \geq Y_0$ or $Y_1 \leq Y_0)$;
(III) $Y_1 \geq Y_0$ or $Y_1 \leq Y_0$

Restriction on treatment equation versus restriction on outcome equation?

Not imposing common treatment effect model, not imposing linear model.

Alternative Notation

Without loss of generality:

$$\begin{aligned} D &= \mathbf{1}[\delta_0 + \delta_1 Z + \epsilon_1 \geq 0] \\ Y &= \mathbf{1}[\alpha_0 + \alpha_1 D + \epsilon_2 \geq 0] \end{aligned}$$

with δ_1, α_1 possibly random. Thus,

$$\begin{aligned} D_z &= \mathbf{1}[\delta_0 + \delta_1 z + \epsilon_1 \geq 0] \\ Y_d &= \mathbf{1}[\alpha_0 + \alpha_1 d + \epsilon_2 \geq 0] \end{aligned}$$

$$Z \perp\!\!\!\perp (Y_0, Y_1, D_0, D_1) \Leftrightarrow Z \perp\!\!\!\perp (\epsilon_1, \epsilon_2, \delta_1, \alpha_1)$$

Alternative Notation (contd.)

Without loss of generality:

$$\begin{aligned} D &= \mathbf{1}[\delta_0 + \delta_1 Z + \epsilon_1 \geq 0] \\ Y &= \mathbf{1}[\alpha_0 + \alpha_1 D + \epsilon_2 \geq 0] \end{aligned}$$

- Restriction that $D_1 \geq D_0$ w.p.1 or $D_1 \leq D_0$ w.p.1 equivalent to δ_1 nonrandom coefficient.
- Restriction that $Y_1 \geq Y_0$ w.p.1 or $Y_1 \leq Y_0$ w.p.1 equivalent to α_1 nonrandom coefficient.
- More general notion: weak separability.
  Appropriate equivalence results developed by Vytlacil (2002, 2006) and Vytlacil and Yildiz (2006).

Parameter of Interest: $\Delta = E(Y_1 - Y_0)$

Define Δ as the average treatment effect:

$$\Delta = E(Y_1 - Y_0) = \Pr[Y_1 = 1] - \Pr[Y_0 = 1]$$

Our goal is to use IV to identify sign of Δ.

Reduced Form Parameter: $\Delta_{YZ} = E(Y|Z=1) - E(Y|Z=0)$

Define Δ_{YZ} as the reduced form paramter:

$$\begin{aligned}\Delta_{YZ} &= E(Y|Z=1) - E(Y|Z=0) \\ &= \Pr[Y=1|Z=1] - \Pr[Y=1|Z=0]\end{aligned}$$

- Is the numerator of TSLS, rescaled TSLS
- $\Pr[D=1|Z=1] > \Pr[D=1|Z=0]$
  $\Rightarrow \text{Sgn}[\Delta_{YZ}] = \text{Sgn}[\Delta]$.
- We will establish conditions under which the value of Δ_{YZ} identifies the sign of Δ.

Reduced Form Parameter: $\Delta_{YZ} = E(Y|Z=1) - E(Y|Z=0)$ (contd.)

Using that $Z \perp\!\!\!\perp (D_0, D_1, Y_0, Y_1)$, we have

$$\begin{aligned}\Delta_{YZ} = &\Big\{\Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0] \\ &- \Pr[Y_1 = 0, Y_0 = 1, D_1 = 1, D_0 = 0]\Big\} \\ &- \Big\{\Pr[Y_1 = 1, Y_0 = 0, D_1 = 0, D_0 = 1] \\ &- \Pr[Y_1 = 0, Y_0 = 1, D_1 = 0, D_0 = 1]\Big\}\end{aligned}$$

Question of Interest

Under what conditions will $\Delta_{YZ} = E(Y|Z=1) - E(Y|Z=0)$ allow us to infer the sign of $\Delta = E(Y_1 - Y_0)$?

- We consider imposing each of following alternative conditions
  - (I) $D_1 \geq D_0$ or $D_1 \leq D_0$;
  - (II) $(D_1 \geq D_0$ or $D_1 \leq D_0)$ and $(Y_1 \geq Y_0$ or $Y_1 \leq Y_0)$
  - (III) $Y_1 \geq Y_0$ or $Y_1 \leq Y_0$

Summary of Results

Under what conditions will Δ_{YZ} allow us to infer sign of Δ?

Imposing

(I) $D_1 \geq D_0$ or $D_1 \leq D_0$:

- Sign of Δ identified if Δ_{YZ} far enough from zero;
- How far is far enough depends on strength of instrument.

(II) $(D_1 \geq D_0$ or $D_1 \leq D_0)$ and $(Y_1 \geq Y_0$ or $Y_1 \leq Y_0)$:

- Sign of Δ is identified by sign of Δ_{YZ}.

(III) $Y_1 \geq Y_0$ or $Y_1 \leq Y_0$

- More nuanced results.

Drawing on Sharp Bounds

In order to establish necessary and sufficient conditions to be able to use Δ_{YZ} to identify the sign of Δ, we draw upon sharp bounds on Δ from the previous literature when possible, and develop the sharp bounds on Δ when necessary.

(I) $D_1 \geq D_0$ or $D_1 \leq D_0$:
- Sharp bounds by Balke and Pearl (1997), Heckman and Vytlacil (2001)

(II) $(D_1 \geq D_0$ or $D_1 \leq D_0)$ and $(Y_1 \geq Y_0$ or $Y_1 \leq Y_0)$:
- Sharp bounds by Bhattacharya, Shaikh and Vytlacil (2008), Shaikh and Vytlacil (2007). Extended by Chiburis (2008).

(III) $Y_1 \geq Y_0$ or $Y_1 \leq Y_0$
- developed by us

Relationship between IV and Sharp Bounds to Identify Sign of Δ

- We develop conditions on Δ_{YZ} that are necessary and sufficient to identify the sign of Δ under alternative restrictions.
- Alternatively, we could construct sharp bounds on Δ and ask under which conditions will the bounds be strictly positive or strictly negative.
- Equivalence, in each case we consider, bounds will be strictly positive or strictly negative iff the corresponding restriction on Δ_{YZ} holds.
- Additionally, in each case we consider, necessary and sufficient testable restrictions can be stated as restrictions on Δ_{YZ}.

(I) Imposing Weak Separability on D equation

Impose $D_1 \geq D_0$ w.p.1 or $D_1 \leq D_0$ w.p.1.

- Using $Z \perp\!\!\!\perp (D_0, D_1)$, we have $\Pr[D = 1|Z = 1] \geq \Pr[D = 1|Z = 0] \Rightarrow D_1 \geq D_0$ w.p.1.
- Sharp bounds on Δ developed by Balke and Pearl (1997), extended to general Y, Z by Heckman and Vytlacil (2001).
- Width of sharp bounds on Δ depends linearly on $1 - \Pr[D = 1|Z = 1]$ and on $\Pr[D = 1|Z = 0]$.
- Relation to Manski (1990) mean independence IV bounds? Relation to Balke and Pearl bounds that don't impose $D_1 \geq D_0$?

(I) Imposing Weak Separability on D equation (contd.)

Using sharp bounds of B-P/H-V, we can

- Identify that $\Delta > 0$ iff
$$\Delta_{YZ} \geq \Pr[Y=1, D=0|Z=1] + \Pr[Y=0, D=1|Z=0]$$
- Identify that $\Delta < 0$ iff
$$\Delta_{YZ} \leq -\Pr[Y=0, D=0|Z=1] - \Pr[Y=1, D=1|Z=0]$$
- Cannot identify sign of Δ iff
$$-\Pr[Y=0, D=0|Z=1] - \Pr[Y=1, D=1|Z=0] < \Delta_{YZ} < \Pr[Y=1, D=0|Z=1] + \Pr[Y=0, D=1|Z=0]$$

(I) Imposing Weak Separability on D equation (contd.)

Can show that necessary and sufficient restrictions on observed data such that there will exist a model satisfying the assumptions and consistent with the observed data are that:

$$\max\begin{pmatrix}\Pr(Y=1,D=0|Z=1)-\Pr(Y=1,D=0|Z=0)\\ \Pr(Y=0,D=1|Z=0)-\Pr(Y=0,D=1|Z=1)\end{pmatrix}$$
$$\leq \Delta_{YZ} \leq$$
$$\min\begin{pmatrix}\Pr(Y=1,D=1|Z=1)-\Pr(Y=1,D=1|Z=0)\\ \Pr(Y=0,D=0|Z=0)-\Pr(Y=0,D=0|Z=1)\end{pmatrix}$$

(I) Imposing Weak Separability on D equation summary

Thus:

- Identify that $\Delta > 0$ if Δ_{YZ} is sufficiently positive.
- Identify that $\Delta < 0$ if Δ_{YZ} is sufficiently negative.
- Fail to identify sign of Δ if Δ_{YZ} is close enough to zero.
- How far from zero is required to identify the sign of Δ depends on strength of instrument and the strength of the effect of D.
- Reject model if Δ_{YZ} far enough from zero.

(II) Impose Weak Separability on Y and D equations

Impose $\big(D_1 \geq D_0$ w.p.1 or $D_1 \leq D_0$ w.p.1$\big)$,
and $\big(Y_1 \geq Y_0$ w.p.1 or $Y_1 \leq Y_0$ w.p.1$\big)$

- Using $Z \perp\!\!\!\perp (D_0, D_1)$, we have $\Pr[D = 1|Z = 1] \geq \Pr[D = 1|Z = 0] \Rightarrow D_1 \geq D_0$ w.p.1..
- D is potentially endogenous, so we do not immediately know whether $Y_1 \geq Y_0$ or $Y_1 \leq Y_0$

(II) Impose Weak Separability on Y and D equations (contd.)

Sharp bounds on Δ developed in Bhattacharya, Shaikh and Vytlacil (2008). For more general model with covariates and general Z, sharp bounds in Shaikh and Vytlacil (2007), Chiburis (2008).

- Use modified-IV strategy that simultaneously shifts exogenous covariates for the Y equation and the instruments Z.
- Recover what shifts in exogenous covariates over- or under-compensates for shifts in D, use this information to construct sharp bounds on Δ.
- Worst case, no exogenous covariates, still uncover with IV whether no shift in exogenous covariates over- or under-compensates for shifts in D.
- Relationship with Manski and Pepper (2000) developed by Bhattacharya, Shaikh and Vytlacil.

(II) Impose Weak Separability on Y and D equations (contd.)

Using that $D_1 \geq D_0$ w.p.1,

$$\Delta_{YZ} = Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0] - \Pr[Y_1 = 0, Y_0 = 1, D_1 = 1, D_0 = 0]$$

Using that $D_1 \geq D_0$ w.p.1, and $Y_1 \geq Y_0$ w.p.1 or $Y_1 \leq Y_0$ w.p.1.,

$$\Delta_{YZ} = \begin{cases} \Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0] & \text{if } Y_1 \geq Y_0 \text{ w.p.1} \\ -\Pr[Y_1 = 0, Y_0 = 1, D_1 = 1, D_0 = 0] & \text{if } Y_1 \leq Y_0 \text{ w.p.1} \end{cases}$$

Thus, sign of Δ_{YZ} equals sign of Δ.

(II) Impose Weak Separability on Y and D equations (contd.)

Additionally, can show necessary and sufficient testable restrictions on Δ_{YZ}.

Reject model if

$$\Delta_{YZ} < \max\begin{pmatrix} \Pr(Y=1, D=0|Z=1) - \Pr(Y=1, D=0|Z=0) \\ \Pr(Y=0, D=1|Z=0) - \Pr(Y=0, D=1|Z=1) \end{pmatrix}$$

Data consistent with model with $\Delta < 0$ if

$$\max\begin{pmatrix} \Pr(Y=1, D=0|Z=1) - \Pr(Y=1, D=0|Z=0) \\ \Pr(Y=0, D=1|Z=0) - \Pr(Y=0, D=1|Z=1) \end{pmatrix} \leq \Delta_{YZ} \leq 0$$

(II) Impose Weak Separability on Y and D equations (contd.)

Data consistent with model with $\Delta > 0$ if

$$0 \leq \Delta_{YZ} \leq \min \begin{pmatrix} \Pr(Y=1, D=1|Z=1) - \Pr(Y=1, D=1|Z=0) \\ \Pr(Y=0, D=0|Z=0) - \Pr(Y=0, D=0|Z=1) \end{pmatrix}$$

Reject model if

$$\Delta_{YZ} > \min \begin{pmatrix} \Pr(Y=1, D=1|Z=1) - \Pr(Y=1, D=1|Z=0) \\ \Pr(Y=0, D=0|Z=0) - \Pr(Y=0, D=0|Z=1) \end{pmatrix}$$

(III) Impose Weak Separability on Y equation

Impose $Y_1 \geq Y_0$ w.p.1 or $Y_1 \leq Y_0$ w.p.1., do not impose monotonicity on D equation

- D is potentially endogenous, so we do not immediately know whether $Y_1 \geq Y_0$ or $Y_1 \leq Y_0$
- Sharp bounds on Δ imposing a priori the direction of effect developed by Manski and Pepper (2000) under weaker independence assumptions.
- Sharp bounds on Δ without imposing a priori the direction of effect previously unknown.
    - Chiburis (2008) sharp bounds under alternative independence assumption

(III) Impose Weak Separability on Y: LP Problem

Following Balke and Pearl, represent the sharp bounds on Δ as the solution to a linear programming problem.

Let

$$q_{ijkl} = \Pr[D_0 = i, D_1 = j, Y_0 = k, Y_1 = l]\ , \quad i, j, k, l \in \{0, 1\}$$

Then

$$\begin{aligned} \Delta &= \Pr[Y_1 = 1] - \Pr[Y_0 = 1] \\ &= \Pr[Y_0 = 0, Y_1 = 1] - \Pr[Y_0 = 1, Y_1 = 0] \\ &= \sum_{i,j} (q_{ij01} - q_{ij10}) \end{aligned}$$

(III) Impose Weak Separability on Y: LP Problem (contd.)

$q_{ijkl} = \Pr[D_0 = i, D_1 = j, Y_0 = k, Y_1 = l]$

Will maximize and minimize $\Delta = \sum_{i,j}(q_{ij01} - q_{ij10})$ subject to restrictions.

Restrictions from latent probabilities being non-negative and summing to one:

$$\begin{aligned} q_{ijkl} &\geq 0 \quad i, j, k, l \in \{0, 1\} \\ \sum_{i,j,k,l} q_{ijkl} &= 1 \end{aligned} \tag{1}$$

(III) Impose Weak Separability on Y: LP Problem (contd.)

$q_{ijkl} = \Pr[D_0 = i, D_1 = j, Y_0 = k, Y_1 = l]$

Let $p_{ij.k} = \Pr(D = i, Y = j | Z = k)$

Restrictions relating observed probabilities to latent probabilities, using $Z \perp\!\!\!\perp (D_0, D_1, Y_0, Y_1)$,

$$\begin{aligned} p_{0y.0} &= \sum_{i,j} q_{0iyj} \\ p_{1y.0} &= \sum_{i,j} q_{1ijy} \\ p_{0y.1} &= \sum_{i,j} q_{i0yj} \\ p_{1y.1} &= \sum_{i,j} q_{i1jy} \end{aligned} \quad (2)$$

for $y \in \{0, 1\}$.

(III) Impose Weak Separability on Y: LP Problem (contd.)

$q_{ijkl} = \Pr[D_0 = i, D_1 = j, Y_0 = k, Y_1 = l]$

$Y_1 \geq Y_0$ w.p.1 $\Rightarrow \Pr[Y_0 = 1, Y_1 = 0] = 0$
$Y_1 \leq Y_0$ w.p.1 $\Rightarrow \Pr[Y_0 = 0, Y_1 = 1] = 0$

Thus, restrictions imposed by $Y_1 \geq Y_0$ w.p.1 or $Y_1 \leq Y_0$ w.p.1

$$q_{ij10} = 0 \qquad (3)$$

or

$$q_{ij01} = 0 \qquad (4)$$

for $i, j \in \{0, 1\}$

(III) Impose Weak Separability on Y: LP Problem (contd.)

Bounds on Δ found by maximizing and minimizing $\Delta = \sum_{i,j}(q_{ij01} - q_{ij10})$ subject to restrictions:

1. latent probabilities nonnegative and sum to one
2. Relationship holds between observed probabilities and latent probabilities implied by independence of Z.
3. $Y_1 \geq Y_0 \Rightarrow q_{ij10} = 0$, or
4. $Y_1 \leq Y_0 \Rightarrow q_{ij01} = 0$.

(III) Impose Weak Separability on Y: LP Problem (contd.)

Restrictions on q_{ijkl}:

1. latent probabilities nonnegative, sum to one
2. Relationship between observed probabilities and latent probabilities implied by independence of Z.
3. $Y_1 \geq Y_0 \Rightarrow q_{ij10} = 0$, or
4. $Y_1 \leq Y_0 \Rightarrow q_{ij01} = 0$.

Four cases:

(i) $\exists\ q_{ijkl}$ satisfying (1), (2), (3) and (1), (2), (4).

(ii) $\exists\ q_{ijkl}$ satisfying (1), (2), (3) but not (1), (2), (4).

(iii) $\exists\ q_{ijkl}$ satisfying (1), (2), (4) but not (1), (2), (3).

(iv) $\nexists\ q_{ijkl}$ satisfying (1), (2), (3) or (1), (2), (4).

(III) Impose Weak Separability on Y: Restrictions from $Y_1 \geq Y_0$

Necessary and sufficient conditions for there to exist q_{ijkl} satisfying (1), (2), (3):

$$\max\begin{pmatrix} -\Pr(Y=1, D=1 \mid Z=0) \\ -\Pr(Y=0, D=0 \mid Z=1) \end{pmatrix} \leq \Delta_{YZ} \leq \min\begin{pmatrix} \Pr(Y=1, D=1 \mid Z=1) \\ \Pr(Y=0, D=0 \mid Z=0) \end{pmatrix}$$

We now show necessity.

(III) Impose Weak Separability on Y: Restrictions from $Y_1 \geq Y_0$

Proof of necessity. Recall that

$$\Delta_{YZ} = \Big\{ \Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0] - \Pr[Y_1 = 0, Y_0 = 1, D_1 = 1, D_0 = 0] \Big\} - \Big\{ \Pr[Y_1 = 1, Y_0 = 0, D_1 = 0, D_0 = 1] - \Pr[Y_1 = 0, Y_0 = 1, D_1 = 0, D_0 = 1] \Big\}$$

Imposing $Y_1 \geq Y_0$, we have

$$\Delta_{YZ} = \Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0] - \Pr[Y_1 = 1, Y_0 = 0, D_1 = 0, D_0 = 1].$$

(III) Impose Weak Separability on Y: Restrictions from $Y_1 \geq Y_0$

$$\Delta_{YZ} = \Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0] - \Pr[Y_1 = 1, Y_0 = 0, D_1 = 0, D_0 = 1].$$

Thus

$$-\Pr[Y_1 = 1, Y_0 = 0, D_1 = 0, D_0 = 1] \leq \Delta_{YZ} \leq \Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0]$$

Deriving upper bound on Δ_{YZ},

$$Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0] \leq \min\begin{pmatrix} \Pr(Y_1 = 1, D_1 = 1) \\ \Pr(Y_0 = 0, D_0 = 0) \end{pmatrix} = \min\begin{pmatrix} \Pr(Y = 1, D = 1 \mid Z = 1) \\ \Pr(Y = 0, D = 0 \mid Z = 0) \end{pmatrix}$$

(III) Impose Weak Separability on Y: Restrictions from $Y_1 \geq Y_0$

$$\Delta_{YZ} = \Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0] \\ - \Pr[Y_1 = 1, Y_0 = 0, D_1 = 0, D_0 = 1].$$

Thus

$$- \Pr[Y_1 = 1, Y_0 = 0, D_1 = 0, D_0 = 1] \\ \leq \Delta_{YZ} \leq \Pr[Y_1 = 1, Y_0 = 0, D_1 = 1, D_0 = 0]$$

Deriving lower bound on Δ_{YZ},

$$\begin{aligned} - \Pr[Y_1 = 1, Y_0 = 0, D_1 = 0, D_0 = 1] & \\ \geq \max & \begin{pmatrix} -\Pr(Y_1 = 1, D_0 = 1) \\ -\Pr(Y_0 = 0, D_1 = 0) \end{pmatrix} \\ = \max & \begin{pmatrix} -\Pr(Y = 1, D = 1 \mid Z = 0) \\ -\Pr(Y = 0, D = 0 \mid Z = 1) \end{pmatrix} \end{aligned}$$

(III) Impose Weak Separability on Y: Restrictions from $Y_1 \geq Y_0$

Thus, (1)-(3) imply

$$B_L^I \leq \Delta_{YZ} \leq B_U^I$$

where

$$\begin{aligned} B_L^I &= \max\begin{pmatrix} -\Pr(Y=1, D=1 \mid Z=0) \\ -\Pr(Y=0, D=0 \mid Z=1) \end{pmatrix} \\ B_U^I &= \min\begin{pmatrix} \Pr(Y=1, D=1 \mid Z=1) \\ \Pr(Y=0, D=0 \mid Z=0) \end{pmatrix} \end{aligned}$$

Can show this condition is also sufficient, only testable restriction from (1)-(3).

(III) Impose Weak Separability on Y: Restrictions from $Y_1 \leq Y_0$

Necessary and sufficient conditions for there to exist q_{ijkl} satisfying (1), (2), (4):

$$B_L^D \leq \Delta_{YZ} \leq B_U^D$$

where

$$\begin{aligned} B_L^D &= \max\begin{pmatrix} -\Pr(Y=0, D=1 \mid Z=1) \\ -\Pr(Y=1, D=0 \mid Z=0) \end{pmatrix} \\ B_U^D &= \min\begin{pmatrix} \Pr(Y=0, D=1 \mid Z=0) \\ \Pr(Y=1, D=0 \mid Z=1) \end{pmatrix} \end{aligned}$$

Can show this condition is also sufficient, only testable restriction from (1), (2), (4).

(III) Impose Weak Separability on Y: Determining Sign of Δ

Case (i):
There exists q_{ijkl} satisfying IV assumptions and $Y_1 \geq Y_0$ and exists q_{ijkl} satisfying IV assumptions and $Y_1 \leq Y_0$ if

$$\Delta_{YZ} \in [B_L^I, B_U^I] \cap [B_L^D, B_U^D]$$

i.e., if

$$\max\{B_L^D, B_L^I\} \leq \Delta_{YZ} \leq \min\{B_U^D, B_U^I\}.$$

In this case, cannot determine sign of Δ

(III) Impose Weak Separability on Y: Determining Sign of Δ

Case (ii):

There exists q_{ijkl} satisfying IV assumptions and $Y_1 \geq Y_0$ but does not exist q_{ijkl} satisfying IV assumptions and $Y_1 \leq Y_0$ if

$$\Delta_{YZ} \in [B_L^I, B_U^I], \qquad \Delta_{YZ} \notin [B_L^D, B_U^D]$$

In this case, Δ_{YZ} identifies that $\Delta \geq 0$.

And likewise with cases (iii) and (iv)

(III) Impose Weak Separability on Y: Determining Sign of Δ

- if Δ_{YZ} close to zero, cannot determine sign of Δ.
- if Δ_{YZ} far enough from zero, but not too far from zero, can determine sign of Δ.
- if Δ_{YZ} too far from zero, reject assumptions.

- Counter-intuitive:

  possible to have $B_U^I < B_U^D$, possible to have $B_L^I < B_L^D$.
  Thus, possible to have Δ_{YZ} so positive that conclude $\Delta < 0$, or to have Δ_{YZ} so negative that conclude $\Delta > 0$.

(III) Impose Weak Separability on Y: Determining Sign of Δ

If $\Pr[D = 1|Z = 0] = 0$, or if $\Pr[D = 1|Z = 1] = 1$,

- $B_L^I = 0$, so $Y_1 \geq Y_0$ feasible iff $\Delta_{YZ} \in [0, B_U^I]$.
- $B_U^D = 0$, so $Y_1 \leq Y_0$ feasible iff $\Delta_{YZ} \in [B_L^D, 0]$.
- So sign of Δ equals sign of Δ_{YZ}.

Relevant for randomized experiments with imperfect compliance.

Relationship to Shaikh and Vytlacil?

Inference on the Sign of Δ

Consider testing the family of (two) null hypotheses

$$H_I : \exists\ q_{i,j,k,l} \text{ satisfying (1), (2), (3)}$$

$$H_D : \exists\ q_{i,j,k,l} \text{ satisfying (1), (2), (4)}$$

in a way that controls the familywise error rate ($FWER$) at level α, where

$$FWER = P\{\text{even one false rejection}\}\ .$$

Inference on the Sign of Δ (cont.)

$$H_I : \exists\ q_{i,j,k,l} \text{ satisfying (1), (2), (3)}$$

$$H_D : \exists\ q_{i,j,k,l} \text{ satisfying (1), (2), (4)}$$

If we

1. fail to reject H_I and reject H_D, then we conclude that $\Delta > 0$,
2. fail to reject H_D and reject H_I, then we conclude that $\Delta < 0$,
3. fail to reject both H_I and H_D, then we can't determine the sign of Δ,
4. reject both H_I and H_D, then we conclude that the model is incorrect.

Inference on the Sign of Δ (cont.)

What is the probability of inferring the sign of Δ incorrectly in this way?

There are three cases to consider:

1. $\Delta > 0$ (H_I true and H_D false). Infer wrong sign only if we reject H_I.
2. $\Delta < 0$ (H_D true and H_I false). Infer wrong sign only if we reject H_D.
3. $\Delta = 0$ (both H_I and H_D true). Infer wrong sign only if we reject H_I or H_D.

Hence, this probability is bounded above by the *FWER*, which is $\leq \alpha$.

Inference on the Sign of Δ (cont.)

Tests of H_I and H_D that control the usual probability of a false rejection at level α are readily available.

p-values for tests of H_I and H_D can be combined in a variety of ways to produce tests of the family of null hypotheses that control the *FWER* at level α.

e.g.,

1. Single-step Bonferroni correction
2. Stepwise method of Holm (1979)
3. Stepwise method of Romano and Wolf (2005)

The last of these methods differs from the first two in that it incorporates information about the dependence structure between the *p*-values.

Example: Vitamin A Supplements

From Sommer and Zeger (1991):

- D indicator for taking vitamin A supplement
- Y indicator for survival 12 months after start of study.
- Z indicator for random assignment to receive supplement.

Example: Vitamin A Supplements

Estimated Conditional Probabilities $P(y, d|z)$
Vitamin A Supplements

	$Z = 0$		$Z = 1$	
	$Y = 0$	Y=1	Y=0	Y=1
D=0	.0064	.9936	.0028	.1972
D=1	0	0	0	.7990

Thus,

$$\begin{aligned} \Delta_{YZ} &= .0026 \\ P(1) - P(0) &= .799 \\ \Delta_{YZ}/[P(1) - P(0)] &= .00325 \end{aligned}$$

Example: Vitamin A Supplements (contd.)

$\Delta_{YZ} = .0026$

- Testable Restriction for Y increasing in D: $0 \leq \Delta_{YZ} \leq .0064$, do not reject. (ignoring sample variability)
- Testable Restriction for Y decreasing in D: $\Delta_{YZ} = 0$, reject. (ignoring sample variability)
- If imposing Y increasing or decreasing in D, then identify increasing, and sharp bounds

$$.0026 \leq \Delta \leq .0064$$

- In comparison, Balke-Pearl Bounds, not imposing any form of monotonicity:

$$-.1946 \leq \Delta \leq .0064$$

Summary of Results

Given alternative monotonicity restrictions, under what conditions will Δ_{YZ} allow us to infer sign of Δ?

Imposing

(I) $D_1 \geq D_0$ or $D_1 \leq D_0$:

- Apply sharp bounds of Balke-Pearl/Heckman-Vytlacil
- Identify that $\Delta > 0$ if Δ_{YZ} is sufficiently positive.
- Identify that $\Delta < 0$ if Δ_{YZ} is sufficiently negative.
- Fail to identify sign of Δ if Δ_{YZ} is close enough to zero.
- How far from zero is required to identify the sign of Δ depends on strength of instrument and the strength of the effect of D.

Summary of Results (contd.)

Given alternative monotonicity restrictions, under what conditions will Δ_{YZ} allow us to infer sign of Δ?

Imposing

(II) $(D_1 \geq D_0$ or $D_1 \leq D_0)$ and $(Y_1 \geq Y_0$ or $Y_1 \leq Y_0)$:

- Apply sharp bounds of Bhattacharya-Shaikh-Vytlacil/Shaikh-Vytlacil
- $\text{Sgn}[\Delta] = \text{Sgn}[\Delta_{YZ}]$.

Summary of Results (contd.)

Given alternative monotonicity restrictions, under what conditions will Δ_{YZ} allow us to infer sign of Δ?

Imposing

(III) $Y_1 \geq Y_0$ or $Y_1 \leq Y_0$

- We develop sharp bounds, using linear programming formulation.
- if Δ_{YZ} close to zero, cannot determine sign of Δ.
- if Δ_{YZ} far enough from zero, but not too far from zero, can determine sign of Δ.
- if Δ_{YZ} too far from zero, reject assumptions.
- possible to have Δ_{YZ} so positive that conclude $\Delta < 0$, or to have Δ_{YZ} so negative that conclude $\Delta > 0$.