Estimating the Effect of a Mis-measured, Endogenous, Binary Treatment

Francis J. DiTraglia

Camilo Garcia-Jimeno

University of Pennsylvania

March 5th, 2016

What is the causal effect of T^*?

$$y_i = h(T_i^*, \mathbf{x}_i) + \varepsilon_i$$

- y – Outcome of interest
- h – Unknown function that *does not depend on i*
- T^* – Unobserved, endogenous binary treatment
- T – Observed, mis-measured binary surrogate for T^*
- $\mathbf{x}$ – Exogenous covariates
- ε – Mean-zero error term
- z – Discrete instrumental variable

Example: Smoking and Birthweight (SNAP Trial)

Coleman et al. (N Engl J Med, 2012)

RCT with 1050 pregnant smokers in England: 521 given nicotine patches, the rest given placebo patches.

- y – Birthweight
- T^* – True smoking behavior
- T – Self-reported smoking behavior
- $\mathbf{x}$ – Mother characteristics
- z – Indicator of nicotine patch

Example: Returns to Schooling

Oreopoulos (2006, AER)

Fuzzy RD: minimum school-leaving age in UK increased from 14 to 15 in 1947 but some already stayed until 15 before the law and others failed to comply after it.

- y – Log wage
- T^* – School attendance at age 15
- T – Self-report of school attendance at age 15
- $\mathbf{x}$ – Individual characteristics
- z – Indicator: born in or after 1933

Related Literature

Continuous Treatment

Lewbel (1997, 2012), Schennach (2004, 2007), Chen et al. (2005), Hu & Schennach (2008), Song (2015), Hu et al. (2015). . .

Binary, Exogenous Treatment

Aigner (1973), Bollinger (1996), Kane et al. (1999), Black et al. (2000), Frazis & Loewenstein (2003), Mahajan (2006), Lewbel (2007)

Binary, Endogenous Treatment

Mahajan (2006), Shiu (2015), Ura (2015)

Model: $y = c + \beta T^* + \varepsilon$

First-stage

$p_k^* \equiv \mathbb{P}(T^* = 1 | z = z_k) \neq \mathbb{P}(T^* = 1 | z = z_\ell) \equiv p_\ell^*$, $k \neq \ell$

Measurement Error

- Non-differential: $\mathbb{E}[\varepsilon | T^*, T, z] = \mathbb{E}[\varepsilon | T^*, z]$
- Does not depend on z:

$$\alpha_0 = \mathbb{P}(T = 1 | T^* = 0, z)$$
$$\alpha_1 = \mathbb{P}(T = 0 | T^* = 1, z)$$

Notation

Define error term that absorbs constant: $u = c + \varepsilon$

Observable Moments: $y = \beta T^* + u$

	$z = 1$	$z = 2$	$\ldots$	$z = K$
$T = 0$	$\bar{y}_{01}$ $\qquad$ p_{01}	$\bar{y}_{02}$ $\qquad$ p_{02}	$\ldots$	$\bar{y}_{0K}$ $\qquad$ p_{0K}
$T = 1$	$\bar{y}_{11}$ $\qquad$ p_{11}	$\bar{y}_{12}$ $\qquad$ p_{12}	$\ldots$	$\bar{y}_{1K}$ $\qquad$ p_{1K}

$$\bar{y}_{tk} = \mathbb{E}[y \mid T = t, z = z_k], \quad p_{tk} = q_k p_k$$

$$q_k = \mathbb{P}(z = z_k), \quad p_k = \mathbb{P}(T = 1 \mid z = z_k)$$

Unobservable Moments: $y = \beta T^* + u$

$$m_{tk}^* = \mathbb{E}[u | T^* = t, z = z_k], \quad p_{tk}^* = q_k p_k^*$$

$$p_k^* = \mathbb{P}(T^* = 1 | z = z_k) = \frac{p_k - \alpha_0}{1 - \alpha_0 - \alpha_1}$$

Possible Assumptions On m_{tk}^*

Joint Exogeneity: $\mathbb{E}[\varepsilon \,|\, T^*, z] = 0$

$\implies m_{tk}^* = c \quad$ for all t, k

Exogenous Treatment: $\mathbb{E}[\varepsilon \,|\, T^*] = 0$

$\implies \dfrac{1}{\mathbb{P}(T^* = t)} \sum_k p_{tk}^* m_{tk}^* = c \quad$ for all t

Exogenous Instrument: $\mathbb{E}[\varepsilon \,|\, z] = 0$

$\implies (1 - p_k^*) m_{0k}^* + p_k^* m_{1k}^* = c \quad$ for all k

Theorem: β is unidentified regardless of K.

(For general case, see paper.)

Proof of special case: $\alpha_0 = 0$

1. System of equations simplifies to

$$\hat{y}_{0k} = c + p_k \beta \left(\frac{\alpha_1}{1 - \alpha_1} \right) - p_k m_{1k}^*$$

$$\hat{y}_{1k} = p_k \beta + p_k m_{1k}^*$$

2. $\beta/(1 - \alpha_1) \equiv \mathcal{W}$ is identified and imposing this, algebra gives $\beta \alpha_1/(1 - \alpha_1) = \mathcal{W} - \beta$.

Theorem: β is undentified regardless of K.

(For general case, see paper.)

Proof of special case: $\alpha_0 = 0$ continued. . .

3. Substituting:

$$(c + p_k \mathcal{W} - \hat{y}_{0k})/p_k = \beta + m_{1k}^*$$

$$\hat{y}_{1k}/p_k = \beta + m_{1k}^*$$

4. Linear system in (β, m_{1k}^*) – no solution or ∞ of solutions.

5. Sum original pair of equations $\implies c + p_k \mathcal{W} - \hat{y}_{0k} = \hat{y}_{1k}$ thus ∞ of solutions. The model is unidentified.

Conditional *Second* Moment Independence.

New Assumption

Homoskedastic errors w.r.t. the *instrument*: $E[\varepsilon^2|z] = E[\varepsilon^2]$

Reasonable?

Makes sense in an RCT or a natural experiment.

New Moment Conditions

Defining $\mu_{k\ell}^* = (p_k - \alpha_0)m_{1k}^* - (p_\ell - \alpha_0)m_{k\ell}^*$,

$$\mathbb{E}(y^2|z_k) - \mathbb{E}(y^2|z_\ell) \equiv \Delta\overline{y^2} = \beta\mathcal{W}(p_k - p_\ell) + 2\mathcal{W}\mu_{k\ell}^*$$

$$\mathbb{E}(yT|z_k) - \mathbb{E}(yT|z_\ell) \equiv \Delta\overline{yT} = (1 - \alpha_1)\mathcal{W}(p_k - p_\ell) + \mu_{k\ell}^*$$

Theorem: $(\alpha_1 - \alpha_0)$ is Identified if $E[\varepsilon^2|z] = E[\varepsilon^2]$

Requires only binary z

Solve for $\mu^*_{k\ell}$, substitute $\beta = \mathcal{W}(1 - \alpha_0 - \alpha_1)$, rearrange to find

$$\alpha_1 - \alpha_0 = 1 + \mathcal{R}/\mathcal{W}, \quad \text{where} \quad \mathcal{R} \equiv \frac{\Delta \overline{y^2} - 2\mathcal{W}\Delta\overline{yT}}{\mathcal{W}(p_k - p_\ell)}.$$

What good is $(\alpha_1 - \alpha_0)$?

- Test necessary condition for *no mis-classification*: $\alpha_0 = \alpha_1$

- Simple, tighter partial identification bounds for β

- If α_0 known, e.g. zero $\implies \beta$ point identified

Conditional *Third* Moment Independence

New Assumption

Third moment independence w.r.t instrument: $E[\varepsilon^3|z] = E[\varepsilon^3]$

New Moment Conditions

Define $\lambda_{k\ell}^* = (p_k - \alpha_0)v_{1k}^* - (p_\ell - \alpha_0)v_{1\ell}^*$

where $v_{tk}^* = \mathbb{E}(u^2|T^* = t, z_k)$. Then

$$
\begin{aligned}
\mathbb{E}(y^3|z_k) \quad &- \quad \mathbb{E}(y^3|z_\ell) \equiv \\
\Delta \overline{y^3} \quad &= \quad \beta^2 \mathcal{W}(p_k - p_\ell) + 3\beta\mathcal{W}\mu_{k\ell}^* + 3\mathcal{W}\lambda_{k\ell}^* \\
\mathbb{E}(y^2 T|z_k) \quad &- \quad \mathbb{E}(y^2 T|z_\ell) \equiv \\
\Delta \overline{y^2 T} \quad &= \quad \beta(1 - \alpha_1)\mathcal{W}(p_k - p_\ell) + 2(1 - \alpha_1)\mathcal{W}\mu_{k\ell}^* + \lambda_{k\ell}^*
\end{aligned}
$$

Theorem: β, α_0 and α_1 identified

Adding $E[\varepsilon^3|z] = E[\varepsilon^3]$, z need only be binary.

Solve for $\lambda_{k\ell}^*$, substitute and rearrange. After further substitutions:

$$2\mathcal{W}^2(1-\alpha_1)^2 + 2\mathcal{R}\mathcal{W}(1-\alpha_1) + (\mathcal{S} - \mathcal{R}^2) = 0$$

where

$$\mathcal{S} \equiv \frac{\Delta\overline{y^3} - 3\mathcal{W}\left[\Delta\overline{y^2T} + \mathcal{R}\Delta\overline{yT}\right]}{\mathcal{W}(p_k - p_\ell)}$$

- Quadratic in $(1-\alpha_1)$ and observables only
- Always two real roots: one is $(1-\alpha_1)$ and the other is α_0.
- To tell which is which, need $\alpha_0 + \alpha_1 < 1$.

Conclusion

- Effect of endogenous, mis-measured, binary treatment.

- Important in applied work but no solution in the literature.

- New partial and point identification results by exploiting higher moments of outcome variable.

- Test necessary condition for absence of measurement error.

- Next steps: use full independence of $z \rightarrow$ optimal estimator