Econometrics of Policy Evaluation

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Estimation Strategies

Philipp Eisenhauer

Setup

The Generalized Roy Model

Potential Outcomes

$$Y_1 = \mu_1(X) + U_1$$

$$Y_0 = \mu_0(X) + U_0$$

Observed Outcome

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

Choice

$$D = \mathrm{I}[\mu_D(X, Z) - V > 0]$$

Key Concept

Definition: Individuals select their treatment status based on gains unobservable by the econometrician. More formally,

$$Y_1 - Y_0 \not\perp\!\!\!\perp D \quad | X = x.$$

$\Rightarrow$ consequences for the choice of the estimation strategy

Useful Notation

$$\begin{aligned} P(X,Z) &= \Pr(D = 1 \mid X, Z) = F_V(\mu_D(X,Z)) \\ U_D &= F_V(V) \end{aligned}$$

Key Assumptions

- (U_1, U_0, V) are independent of Z conditional on X
- $\mu_D(X, Z)$ is a nondegenerate random variable conditional on X
- $0 < \Pr(D = 1 \mid X) < 1$
- ...

Evaluation Problem

$$Y = DY_1 + (1 - D)Y_0 = \begin{cases} Y_1 & \text{if} \quad D = 1 \\ Y_0 & \text{if} \quad D = 0 \end{cases}$$

Selection Problem

$$
\begin{aligned}
E[Y \mid D=1]-E[Y \mid D=0] = & \underbrace{E[Y_1-Y_0]}_{B^{ATE}} \\
& + \underbrace{E[Y_1-Y_0 \mid D=1]-E[Y_1-Y_0]}_{\text{Sorting Gain}} \\
& + \underbrace{E[Y_0 \mid D=1]-E[Y_0 \mid D=0]}_{\text{Selection Bias}}
\end{aligned}
$$

Estimation Strategies

- Randomization
- Matching
- Instrumental Variables
  - conventional and local
- Regression Discontinuity
  - fuzzy and sharp design

Randomization

Treatment Status

D	self-selected
ξ	assigned
A	actual

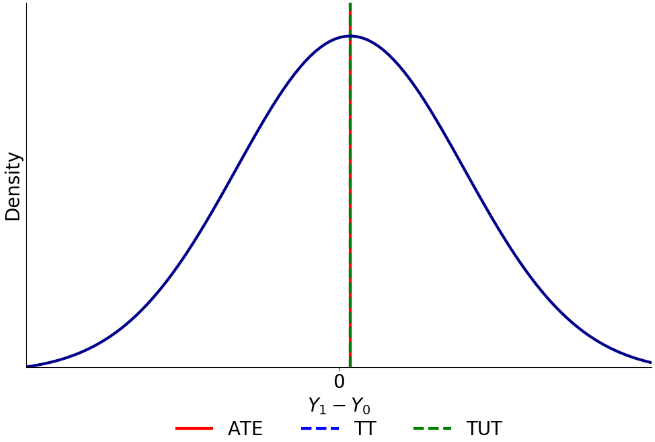
Key Identifying Assumptions

$$(Y_1, Y_0) \perp\!\!\!\perp D$$
$$(Y_1, Y_0) \perp\!\!\!\perp \xi$$
$$(Y_1, Y_0) \perp\!\!\!\perp A$$

When do we have to worry about compliance?

$$
\begin{aligned}
E(Y \mid A=1) - E(Y \mid A=0) \\
&= E(Y_1 \mid A=1) - E(Y_0 \mid A=0) \\
&\qquad\qquad\qquad \text{(by full compliance)} \\
&= E(Y_1) - E(Y_0) \qquad \text{(by randomization)} \\
&= ATE = TT = TUT
\end{aligned}
$$

Figure: Distribution of Effects



What if we can only deny program participation to individuals who are willing to participate?

$$
\begin{aligned}
E(Y \mid D=1, A=1) - E(Y \mid D=1, A=0) \\
&= E(Y_1 \mid D=1, A=1) - E(Y_0 \mid D=1, A=0) \\
&= E(Y_1 \mid D=1) - E(Y_0 \mid D=1) \\
&= TT \neq ATE \neq TUT
\end{aligned}
$$

Issues

- Compliance
- Imperfect Randomization
- Ethical Concerns
- Feasibility
- Expenses
- External Validity

Challenges to Scaling Experiments

- market equilibrium effects
- spillovers
- political reactions
- context dependence
- randomization or site-selection bias
- piloting bias

See Banerjee et al. (2017) for a discussion of these challenges and their attempts to address them in their work.

Matching

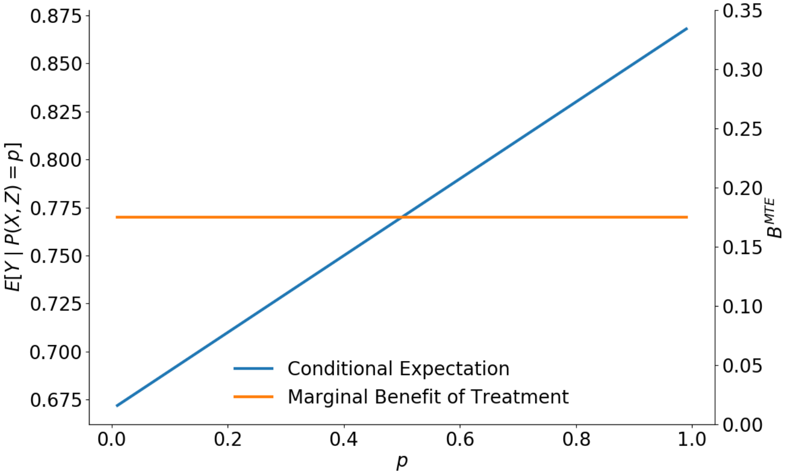
Key Identifiying Assumption

$$(Y_1, Y_0) \perp\!\!\!\perp D \mid X$$

What is in the agent's and econometrician's information set?

- J. J. Heckman and Navarro-Lozano (2004) highlights the sensitivity of results to different conditioning variables.

Figure: Matching and Essential Heterogeneity



Instrumental Variables

Key Identifiying Assumption

$$(Y_1, Y_0) \perp\!\!\!\perp Z \mid X$$

Even in the best cases, this is sometimes not as obvious as you think. See J. Heckman (1997) for a study of implicit behavioral assumptions used in making program evaluations.

Conventional Notation

$$Y = \alpha + \beta D + \epsilon,$$

where

$$\alpha = \mu_0(X)$$
$$\beta = (Y_1 - Y_0) = \mu_1(X) - \mu_0(X) + (U_1 - U_0)$$
$$\epsilon = U_0$$

Assume for now that there is no treatment effect heterogeneity, i.e. $Y_1 - Y_0$ is the same for everybody. If we have access to a variable Z with the following properties ...

$$\text{cov}(Z, D) \neq 0$$
$$\text{cov}(Z, \epsilon) = 0$$

then the following holds

$$\text{plim}\,\hat{\beta}_{IV} = \frac{\text{cov}(Z, Y)}{\text{cov}(Z, D)} = \beta$$

What happens if β varies in the population?

- Do individuals select their treatment status based on gains?

  $\Rightarrow$ essential heterogeneity

Let $\beta = E[\beta] + \eta$, where $U_1 - U_0 = \eta$, then

$$Y = \alpha + \bar{\beta} D + [\epsilon + \eta D].$$

and

$$\operatorname{plim} \hat{\beta}_{IV} = \bar{\beta} + \frac{\operatorname{cov}(Z, \epsilon + \eta D)}{\operatorname{cov}(D, Z)}$$

So we cannot even learn about the mean effect of treatment unless we rule out essential heterogeneity, i.e. individuals selecting their treatment status based on gains.

Local Average Treatment Effect

- Average effect for those induced to change treatment because of a change in the instrument.

  $\Rightarrow$ instrument-dependent parameter

$$\frac{E(Y \mid Z = z) - E[Y \mid Z = z']}{P(z) - P(z')} =$$
$$E(Y_1 - Y_0 \mid D(z) = 1, D(z') = 0)$$

Local Instrumental Variables

Local Instrumental Variable

$$\left.\frac{\partial E(Y \mid P(Z) = p)}{\partial p}\right|_{p=u_D} = E(Y_1 - Y_0 | U_D = u_D)$$
$$= B^{MTE}(u_D)$$

$\Rightarrow$ we can only identify the $B^{MTE}(u_D)$ over the support of p in our data

Figure: Observed Outcome and Essential Heterogeneity

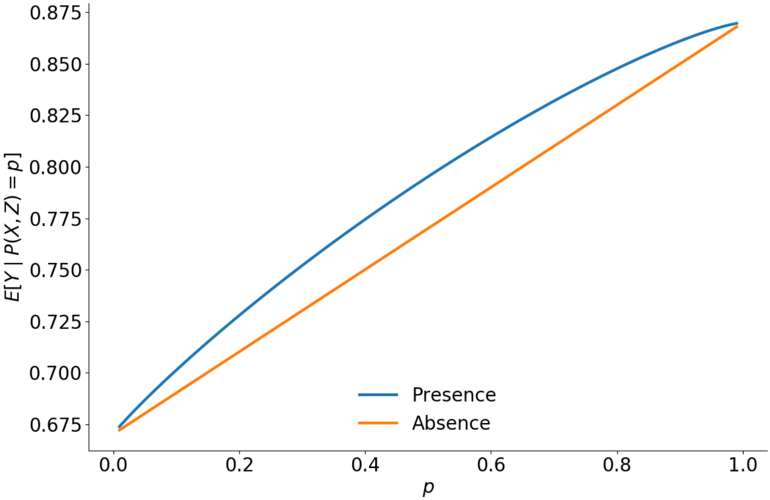
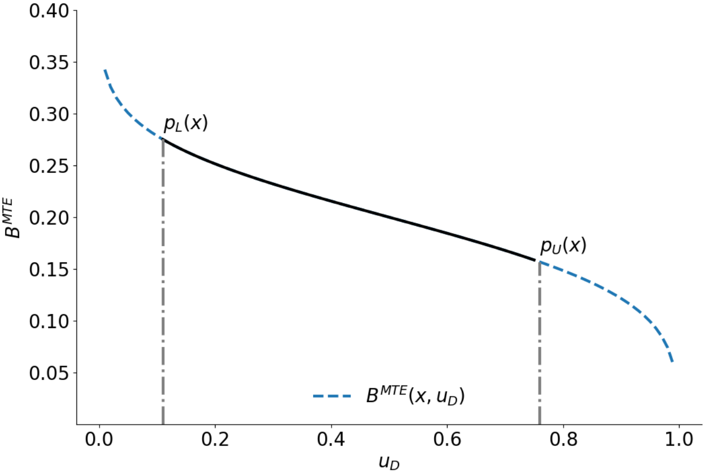


Figure: Identification I



Making $X = x$ explicit

$$\begin{aligned} &E(Y_1 - Y_0 | X = x, U_D = u_D) \\ &\qquad = (\mu_1(x) - \mu_0(x)) + E(U_1 - U_0 | X = x, U_D = u_D) \end{aligned}$$

but if we are willing to assume $(U_1 - U_0) \perp\!\!\!\perp X$ then

$$\begin{aligned} &E(Y_1 - Y_0 | X = x, U_D = u_D) \\ &\qquad = (\mu_1(x) - \mu_0(x)) + E(U_1 - U_0 | U_D = u_D) \end{aligned}$$

Figure: Identification II

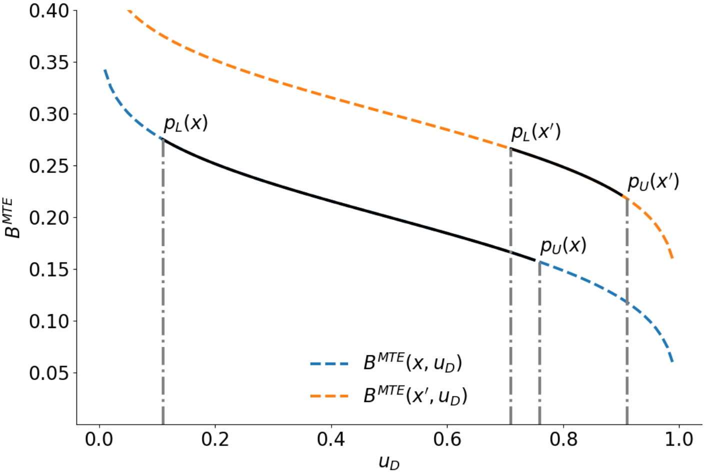
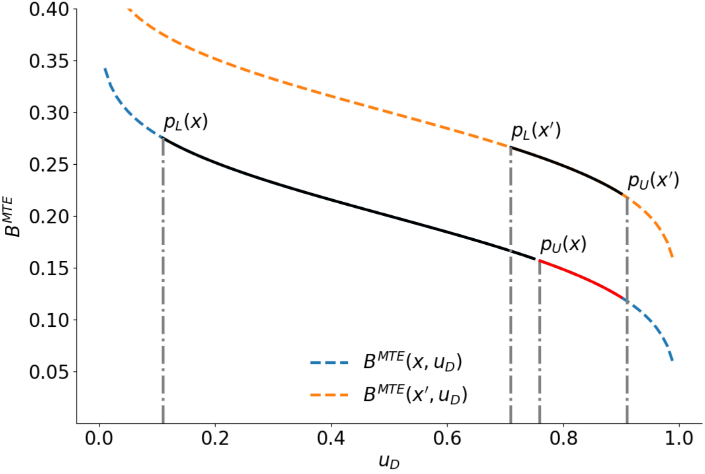


Figure: Identification III



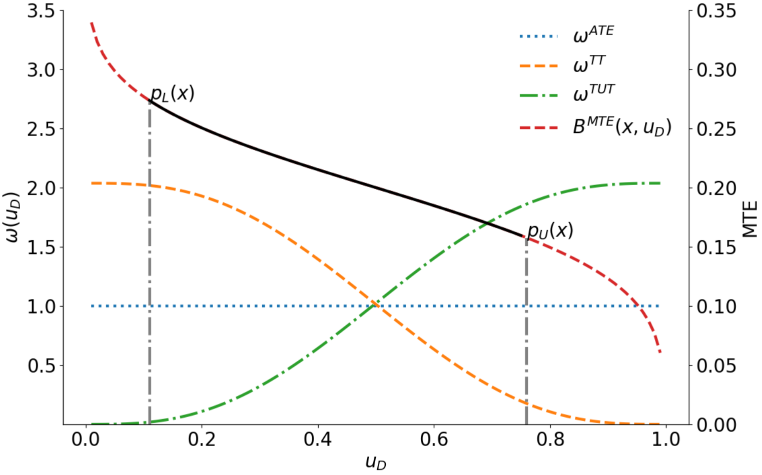
Effects of Treatment as Weighted Averages

Parameter Δ_j, can be written as a weighted average of the $B^{MTE}(x, u_D)$.

$$\Delta_j(x) = \int_0^1 B^{MTE}(x, u_D)\omega^j(x, u_D)du_D,$$

where the weights $\omega^j(x, u_D)$ are specific to parameter j and integrate to one.

Figure: Identification IV



Regression Discontinuity Design

Suppose $D = 1$ if $X \geq x_0$, and $D = 0$ otherwise

$$\Rightarrow \begin{cases} E(Y \mid X = x) = E(Y_0 \mid X = x) & \text{for} \quad x < x_0 \\ E(Y \mid X = x) = E(Y_1 \mid X = x) & \text{for} \quad x \geq x_0 \end{cases}$$

Suppose $E(Y_1 \mid X = x), E(Y_0 \mid X = x)$ are continuous in x.

$$\Rightarrow \begin{cases} \lim_{\epsilon \searrow 0} E(Y_0 \mid X = x_0 - \epsilon) = E(Y_0 \mid X = x_0) \\ \lim_{\epsilon \searrow 0} E(Y_1 \mid X = x_0 + \epsilon) = E(Y_1 \mid X = x_0) \end{cases}$$

$$
\begin{aligned}
\lim_{\epsilon \searrow 0} E(Y \mid X = x_0 + \epsilon) - \lim_{\epsilon \searrow 0} E(Y \mid X = x_0 - \epsilon) \\
&= \lim_{\epsilon \searrow 0} E(Y_1 \mid X = x_0 + \epsilon) - \lim_{\epsilon \searrow 0} E(Y_0 \mid X = x_0 - \epsilon) \\
&= E(Y_1 \mid X = x_0) - E(Y_0 \mid X = x_0) \\
&= E(Y_1 - Y_0 \mid X = x_0)
\end{aligned}
$$

Figure: Probability

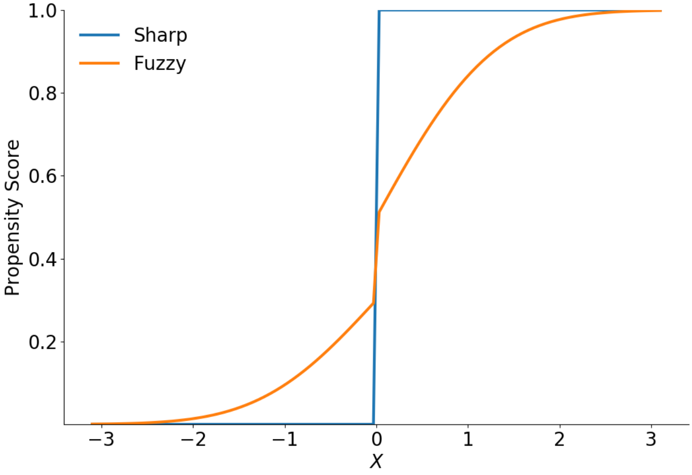


Figure: Observed Outcome in a Sharp Design

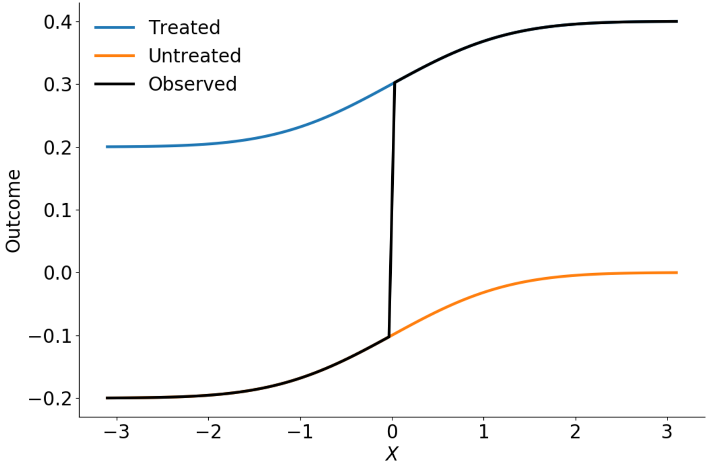
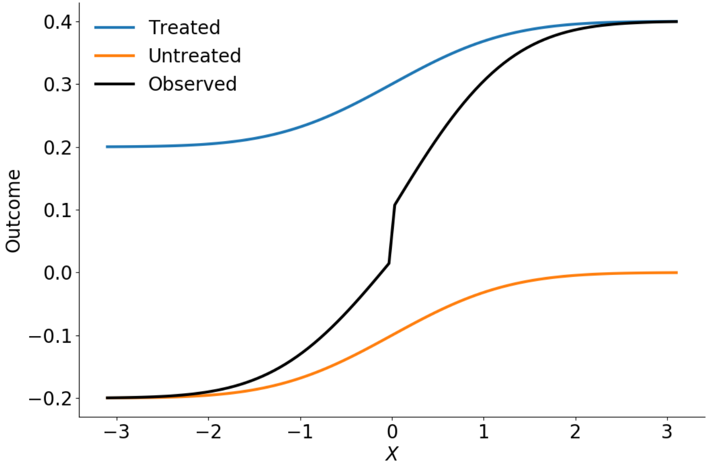


Figure: Observed Outcome in a Fuzzy Design



Conclusion

We must not cease from exploration and the end of all our exploring will be to arrive where we began and to know the place for the first time.

- T. S. Eliot (1943)

Appendix

References

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., . . . Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives*, *31*(4), 73–102.

Carneiro, P., Heckman, J. J., & Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, *101*(6), 2754–2781.

Eliot, T. (1943). Little gidding. In *Four quartets.* Harcourt.

Heckman, J. (1997). Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. *The Journal of Human Resources*, *32*(3), 441. Retrieved from http://www.jstor.org/stable/146178?origin=crossref doi: doi: 10.2307/146178

Heckman, J. J. (1990). Selection bias and self-selection. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Econometrics* (pp. 201–224). London: Palgrave Macmillan.

Heckman, J. J., & Navarro-Lozano, S. (2004). Using matching, instrumental variables, and control functions to estimate economic choice models. *Review of Economics and Statistics*, *86*(1), 30–57.

Heckman, J. J., & Taber, C. (2010). Roy model. In L. E. Blume & S. N. Durlauf (Eds.), *Microeconometrics* (pp. 221–228). London: Palgrave Macmillan.