Econometrics of Policy Evaluation

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Introduction to the Econometrics of Policy Evaluation

Philipp Eisenhauer

Heckman (2008) defines three policy evaluation tasks:

- Evaluating the impact of historical interventions on outcomes including their impact in terms of well-being of the treated and the society at large.
- Forecasting the impact of historical interventions implemented in one environment in other environments, including their impact in terms of well-being.
- Forecasting the impacts of interventions never historically experienced to various environments, including their impact on well-being.

Policy Evaluation Tasks

The Econometrics of Policy Evaluation:

- is important
- is complicated
- is multifaceted

Numerous Applications:

- Labor Economics
- Development Economics
- Industrial Economics
- Health Economics

Numerous Effects:

- Conventional Average Effects
- Policy-relevant Average Effects
- Marginal Effects
- Distributional Effects
- Effects on Distributions

Numerous Estimation Strategies:

- Instrumental Variables
- (Quasi-)Experimental Methods
- Matching

Generalized Roy Model as Unifying Framework

Potential Outcomes

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

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

Cost

$$C = \mu_D(Z) + U_C$$

Observed Outcomes

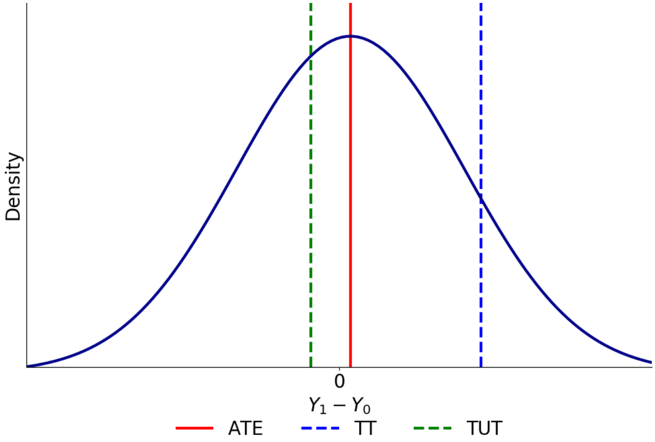
$$Y = DY_1 + (1 - D)Y_0$$

Choice

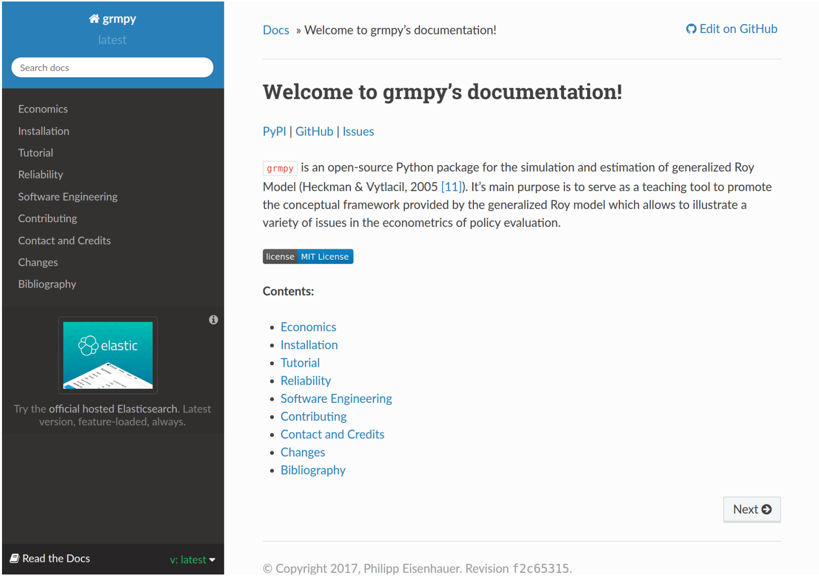
$$S = Y_1 - Y_0 - C$$

$$D = \mathrm{I}[S > 0]$$

Figure: Treatment Effects



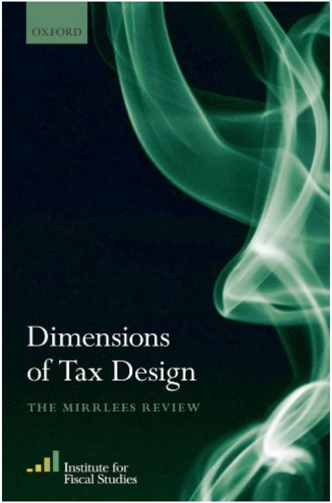
Teaching Tool



The Mirrlees Review

The Mirrlees Review brought together a high-profile group of international experts and early career researchers to identify the characteristics of a good tax system for any open developed economy in the 21st century, assess the extent to which the UK tax system conforms to these ideals, and recommend how it might realistically be reformed in that direction.

Figure: The Mirrlees Review



- **Taxation of Earnings**
  - A single integrated benefit should be introduced to replace all or most of the current multiplicity of benefits, rationalising the way in which total support varies with income and other characteristics.
- **Indirect Taxes**
  - VAT should be extended to nearly all spending. This would reduce complexity and avoid costly distortions to consumption choices.

- **Environmental Taxes**
  - We should work towards a comprehensive system of congestion charging on the roads, replacing most of fuel duty.
- **Taxes on Saving**
  - The risk-free return to saving should not be taxed, so that saving is not discouraged.
- **Business Taxes**
  - The tax treatment of employment, self-employment and corporate source income should be aligned.

Appendix

References

Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., . . . Poterba, J. (2010). *Dimensions of tax design: The mirrlees review* (O. U. Press, Ed.). Institute for Fiscal Studies (ifs).

Cho, S., & Rust, J. (2010). The flat rental puzzle. *Review of Economic Studies*, *77*(2), 560–594.

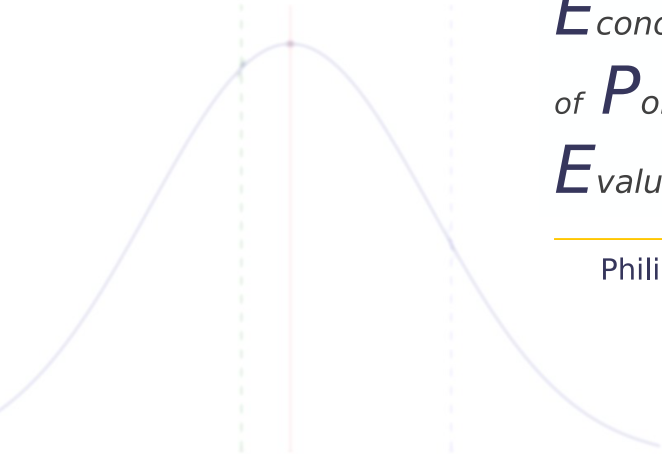
Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J., . . . Group, O. H. S. (2012). The Oregon health insurance experiment: Evidence from the first year. *The Quarterly Journal of Economics*, *127*(3), 1057-1106. Retrieved from http://dx.doi.org/10.1093/qje/qjs020 doi: doi: 10.1093/qje/qjs020

Gordon, R. (2011). Commentary on tax by design: The Mirrlees Review. *Fiscal Studies*, *32*(3), 395-414.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P., & Yavitz, A. (2010). The rate of return of the High-Scope Perry Preschool Program. *Journal of Public Economics*, *94*(1), 114–128.

Mirrlees, J., Adam, S., Besley, T., Bond, S., Chote, R., Gammie, M., . . . Poterba, J. (2011). *Tax by design* (O. U. Press, Ed.). Institute for Fiscal Studies (ifs) and Mirrlees, J.

Todd, P. E., & Wolpin, K. I. (2006). Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review*, *96*(5), 1384–1417.



Econometrics of Policy Evaluation

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Generalized Roy Model

Philipp Eisenhauer

- Does the pursuit of comparative advantage increase or decrease earnings in equality within sectors and in the overall economy?
- Do the people with the highest i skill actually work in sector i?
- As people enter a sector in response to an increase in the demand for its services, does the average skill level employed there rise or fall?

(Roy, 1951) Model

- Individuals are income maximizing, act under perfect information, and possess skills S_1 and S_2.
- The economy offers two employment opportunities associated with skill prices π_1 and π_2 and skill i is only useful in sector i.

An individual chooses sector one if earnings are greater there:

$$w_1 > w_2 \iff \pi_1 S_1 > \pi_2 S_2$$

Econometric Problems

- **Evaluation Problem** We only observe an individual's wage in the sector they are working in.
- **Selection Problem** As individuals pursue their comparative advantage, we only observe selected samples from the latent skill distribution in either sector.

Key Questions

- What economic concepts are accounted for, which are not?
- What does the individual, what does the econometrician know?
- What gives rise to heterogeneity in skills?

- Skills follow a bivariate normal distribution denoted by $F(s_1, s_2)$.

$$\begin{pmatrix} \ln S_1 \\ \ln S_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right)$$

Figure: Joint Distribution of Skills

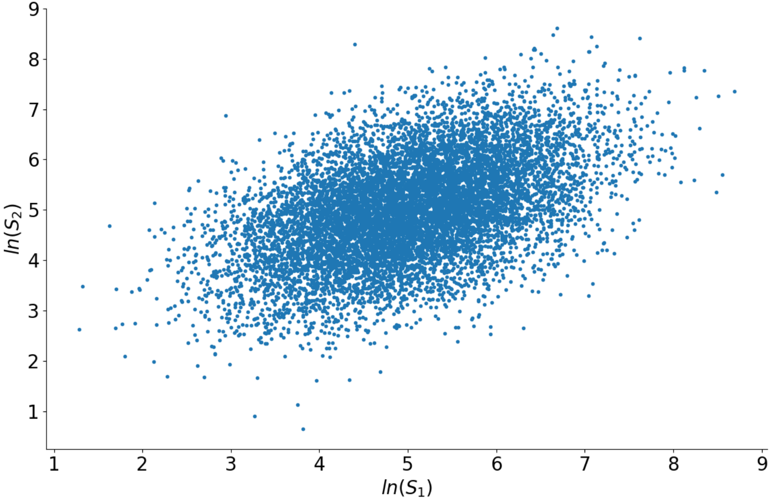
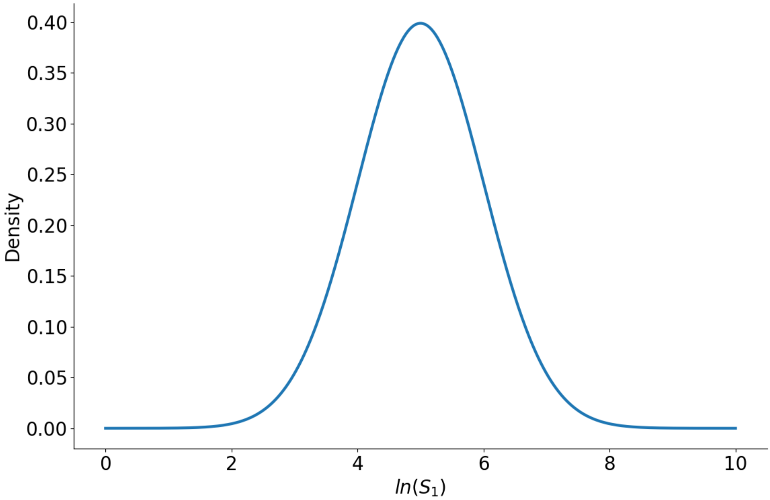


Figure: Marginal Distribution of Skill



The proportion of the population working in sector one P_1

$$P_1 = \int_0^\infty \int_0^{\pi_1 s_1/\pi_2} f(s_1, s_s) ds_1 ds_2$$

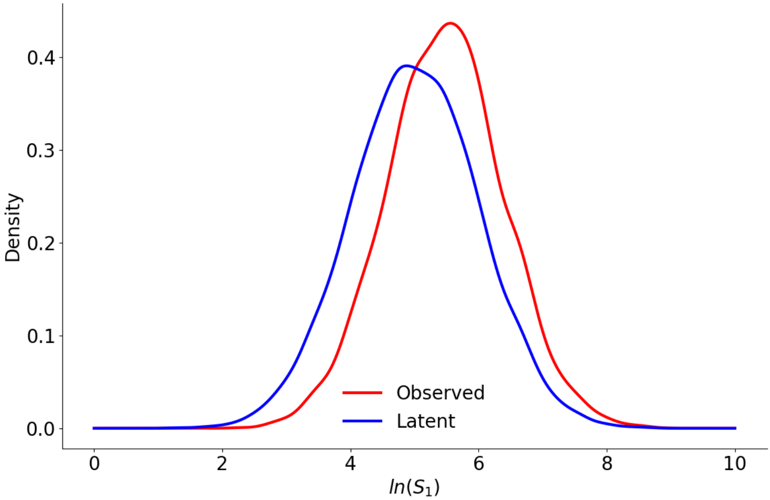
The density of skills employed in sector one differs from the population density of skills.

$$f(s_1) = \int_0^\infty f(s_1, s_2) ds_2$$

$$g_1(s_1 \mid \pi_1 S_1 > \pi_2 S_2) = \frac{1}{P_1} \int_0^{\pi_1 s_1/\pi_2} f(s_1, s_2) ds_2$$

The distribution of skills employed in sector 1 differs from the population distribution of skills due to comparative advantage.

Figure: Latent and Observed Distribution of Skill



Wage Equations

$$\ln W_1 = \ln \pi_1 + \mu_1 + U_1$$
$$\ln W_2 = \ln \pi_2 + \mu_2 + U_2,$$

where $U_i = \ln S_i - \mu_i$.

Sorting

$$E[\ln S_1 \mid \ln W_1 > \ln W_2] = \mu_1 + \frac{\sigma_{11} - \sigma_{12}}{\sigma^*}\lambda(-c_1)$$

$$E[\ln S_2 \mid \ln W_2 > \ln W_1] = \mu_2 + \frac{\sigma_{22} - \sigma_{12}}{\sigma^*}\lambda(-c_2)$$

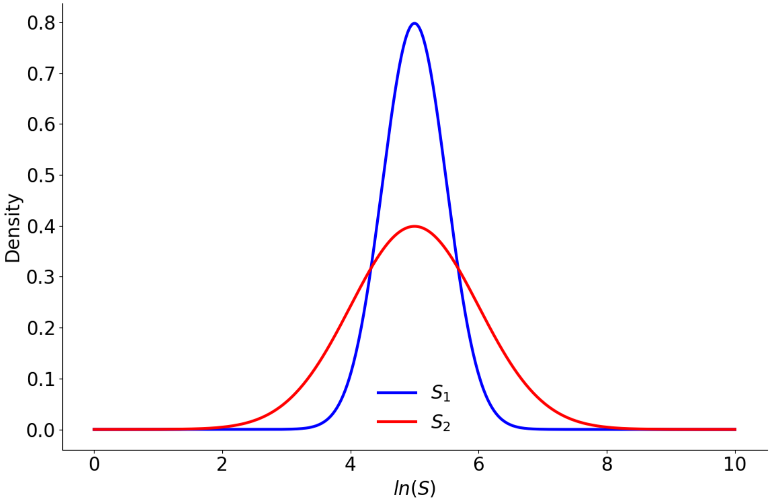
We know the following:

$$\sigma^* = (\sigma_{11} - \sigma_{12}) + (\sigma_{22} - \sigma_{12}) > 0$$

$$\lambda > 0$$

- There must be positive selection into one of the occupations and there can be positive selection into both.

Figure: Marginal Distributions of Skills



What do we know?

- There is positive selection in Sector 2, because there cannot be negative selection in both and $\sigma_{22} > \sigma_{11}$.

How about Sector 1?

- If $\sigma_{12} < 0$, then there is also positive selection in Sector 1.

- If $\rho_{12} = 1$, then there is negative selection into Sector 1 as $\sigma_{12} > \sigma_{11}$

We gain further insights into the effect of self-selection on the distribution of earnings for workers in sector 1 by looking at the distribution of $\ln S_1$ conditional on $\ln S_2$.

$$\ln S_1 \mid \ln S_2 \sim \mathbb{N}(\mu, \sigma),$$

where

$$\mu = \mu_1 + \frac{\sigma_{12}}{\sigma_{22}}\left(\ln S_2 - \mu_2\right) \quad \text{and} \quad \sigma = \sigma_{11}\left(1 - \left(\frac{\sigma_{12}}{\sigma_1 \sigma_2}\right)^2\right)$$

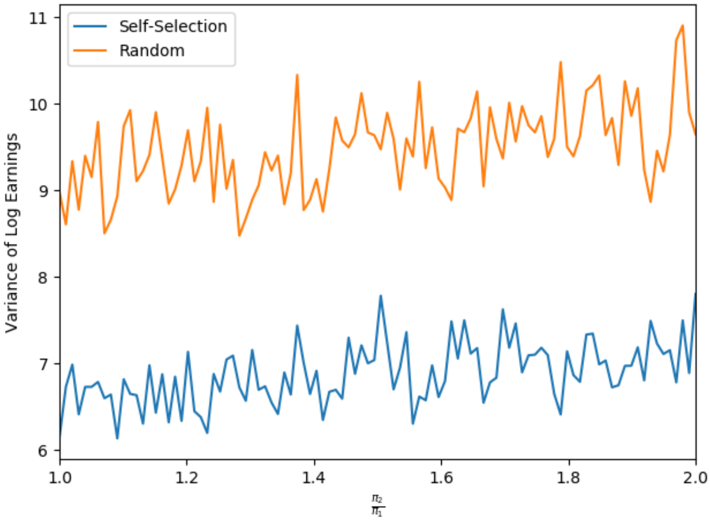
Heckman Productions

Importance of Assignment Mechanism

Heckman and Honore (1990) show that ...

> *For a log normal Roy economy, any random assignment of persons to sectors with the same proportion of persons in each sector as in the Roy economy has higher variance of log earnings provided the proportions lie strictly in the unit interval. This is true whether or not skill prices in the two economies are the same.*

Choices over Time



Incarnations of the Roy Model

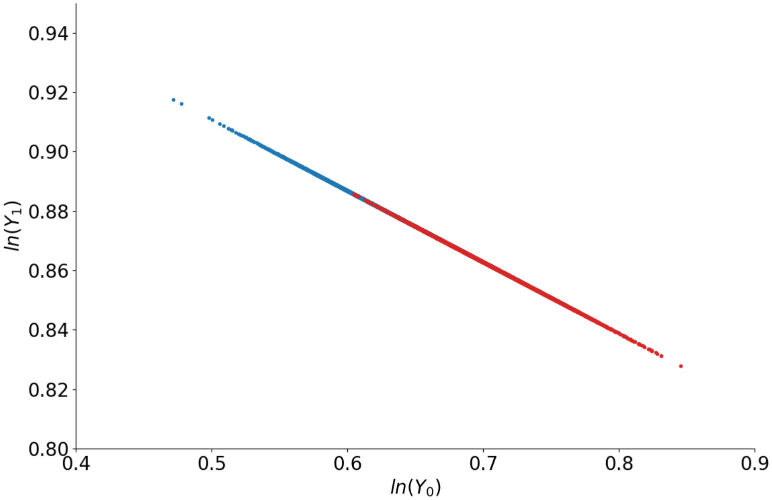
Incarnations of the Roy Model

The Generalized Roy Model

Potential Outcomes	Cost
$Y_1 = \mu_1(X) + U_1$	$C = \mu_D(Z) + U_C$
$Y_0 = \mu_0(X) + U_0$	

Observed Outcomes	Choice
$Y = DY_1 + (1-D)Y_0$	$S = Y_1 - Y_0 - C$
	$D = \mathrm{I}[S > 0]$

Figure: Occupational Sorting in the Generalized Roy Model

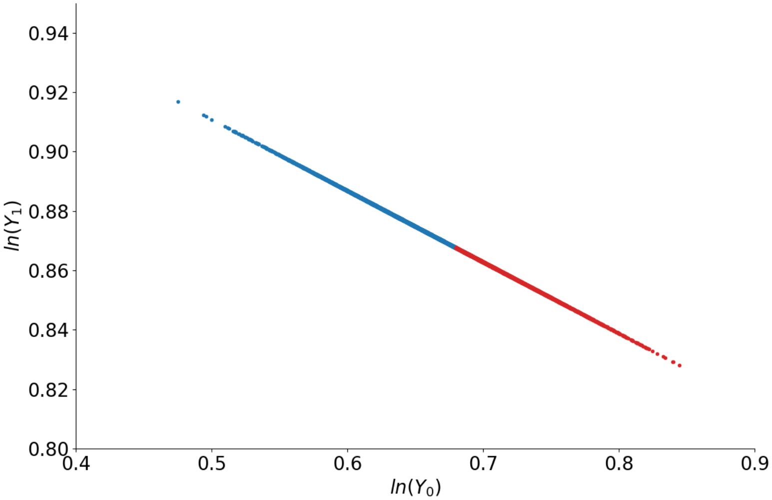


Extended Roy Model

Potential Outcomes	Cost
$Y_1 = \mu_1(X) + U_1$	$C = \mu_D(Z)$
$Y_0 = \mu_0(X) + U_0$	

Observed Outcomes	Choice
$Y = DY_1 + (1 - D)Y_0$	$S = Y_1 - Y_0 - C$
	$D = \mathrm{I}[S > 0]$

Figure: Occupational Sorting in the Extended Roy Model

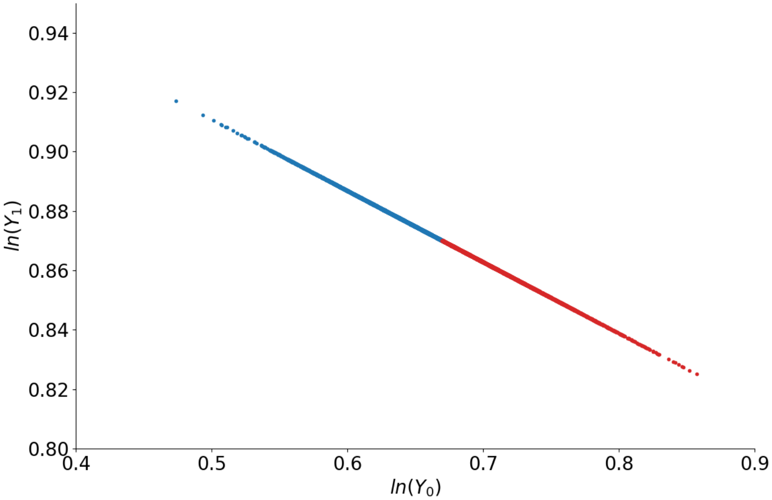


Mapping Notation to original Roy Model

Potential Outcomes	Cost
$W_1 = \pi_1 S_1$	$C = 0$
$W_2 = \pi_2 S_2$	

Observed Outcomes	Choice
$W = DW_1 + (1 - D)W_2$	$S = W_1 - W_2$
	$D = \mathrm{I}[S > 0]$

Figure: Occupational Sorting in the Original Roy Model



Appendix

References

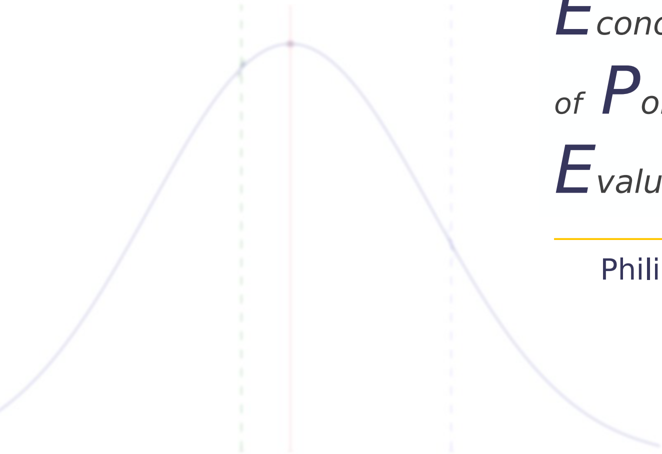
Carneiro, P., Heckman, J. J., & Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, *101*(6), 2754–2781.

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., & Honore, B. E. (1990). The empirical content of the Roy model. *Econometrica*, *58*(5), 1121–1149.

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

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, *3*(2), 135–146.



Econometrics of Policy Evaluation

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Parameters of Interest

Philipp Eisenhauer

Heckman (2008) sets out three tasks for us:

- Defining the Set of Hypotheticals or Counterfactuals
  $\Rightarrow$ A Scientific Theory

- Identifying Causal Parameters from Real Data
  $\Rightarrow$ Mathematical Analysis of Data Point or Set Identification

- Identifying Parameters from Real Data
  $\Rightarrow$ Estimation and Testing Theory

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

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

Useful Notation

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

Specification We follow the parameterization in Heckman and Vytlacil (2005):

$$
\begin{array}{llll}
Y_1 = \gamma + \alpha + U_1 & U_1 = \sigma_1 \epsilon & \gamma = 0.670 & \sigma_1 = 0.012 \\
Y_0 = \gamma + U_0 & U_0 = \sigma_0 \epsilon & \alpha = 0.200 & \sigma_0 = -0.050 \\
D = \mathrm{I}[Z - V > 0] & V \ \ = \sigma_V \epsilon & \epsilon \sim \mathbb{N}(0, 1) & \sigma_V = -1.000
\end{array}
$$

$$
Z \sim \mathbb{N}(-0.0026, 0.2700) \qquad U_D = \Phi\left(\frac{V}{\sigma_V \sigma_\epsilon}\right)
$$

Individual Heterogeneity

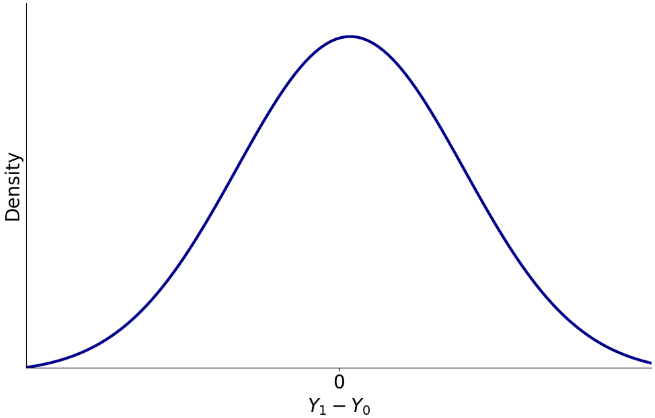
Individual-specific Benefit of Treatment

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

Sources of Heterogeneity

- Difference in Observable Characteristics
- Difference in Unobservable Characteristics
    - Uncertainty
    - Private Information

Figure: Distribution of Benefits

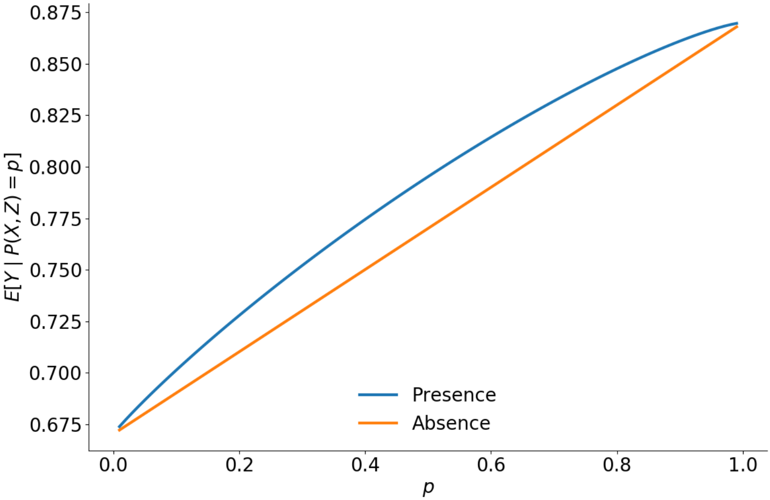


Essential Heterogeneity Definition: Individuals select their treatment status based on gains unobservable by the econometrician. More formally,

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

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

Figure: Conditional Expectation and Essential Heterogeneity



Conventional Average Treatment Effects

Conventional Average Treatment Effects

$$\begin{aligned} B^{ATE} &= E[Y_1 - Y_0] \\ B^{TT} &= E[Y_1 - Y_0 \mid D = 1] \\ B^{TUT} &= E[Y_1 - Y_0 \mid D = 0] \end{aligned}$$

$\Rightarrow$ correspond to *extreme* policy alternatives

Selection Problem

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

$$E[Y \mid D=1] - E[Y \mid D=0] = \underbrace{E[Y_1 - Y_0 \mid D=1]}_{B^{TT}} + \underbrace{E[Y_0 \mid D=1] - E[Y_0 \mid D=0]}_{\text{Selection Bias}}$$

$\Rightarrow$ the bias depends on the parameter of interest

Figure: Distribution of Effects with Essential Heterogeneity

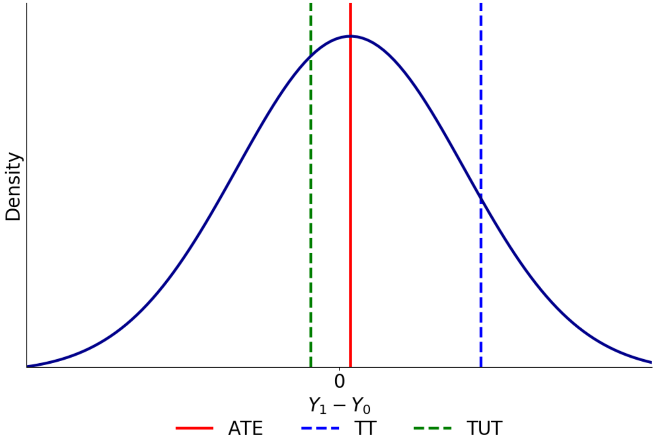


Figure: Distribution of Effects without Essential Heterogeneity

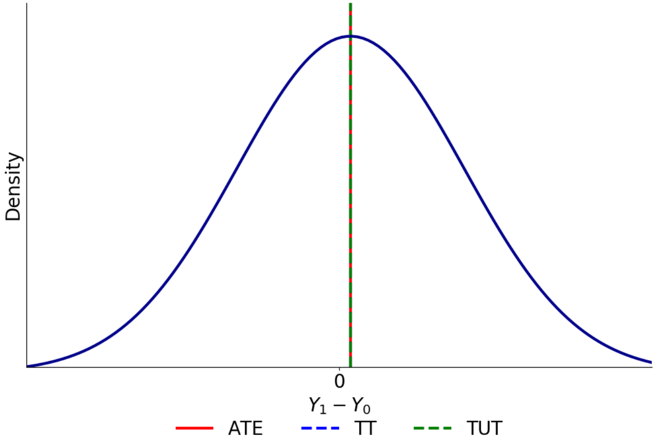
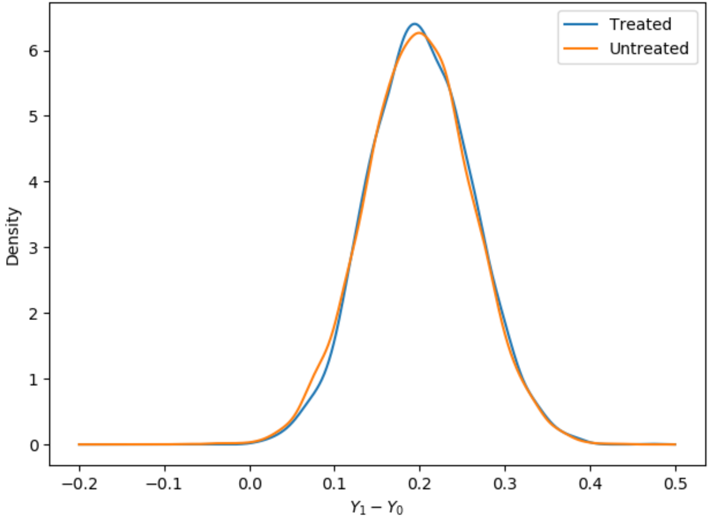


Figure: Distribution of Benefits by Treatment Status



Policy-Relevant Average Treatment Effects

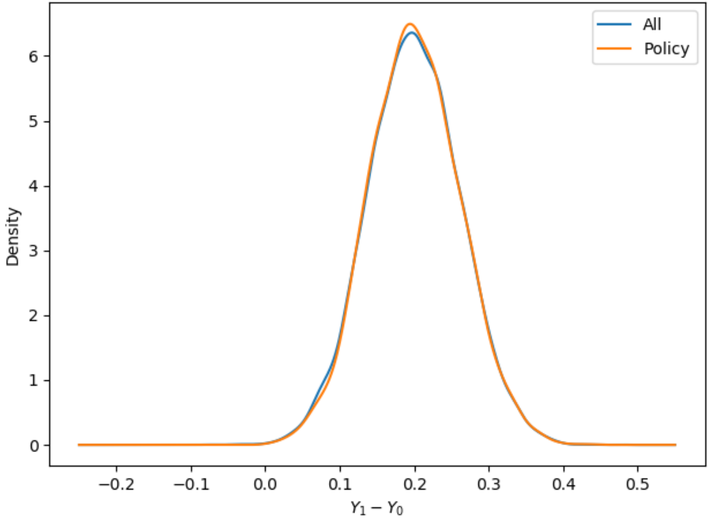
Observed Outcomes

$$Y_B = D_B Y_1 + (1 - D_B) Y_0$$
$$Y_A = D_A Y_1 + (1 - D_A) Y_0$$

Effect of Policy

$$B^{PRTE} = \frac{1}{E[D_A] - E[D_B]} (E[Y_A] - E[Y_B])$$

Figure: Distribution of Benefits for Policy



Marginal Effect of Treatment

Marginal Benefit of Treatment

$$B^{MTE}(x, u_D) = E[Y_1 - Y_0 \mid X = x, U_D = u_D]$$

Intuition: Mean gross return to treatment for persons at quantile u_D of the first-stage unobservable V or a willingness to pay for individuals at the margin of indifference.

Figure: Margin of Indifference

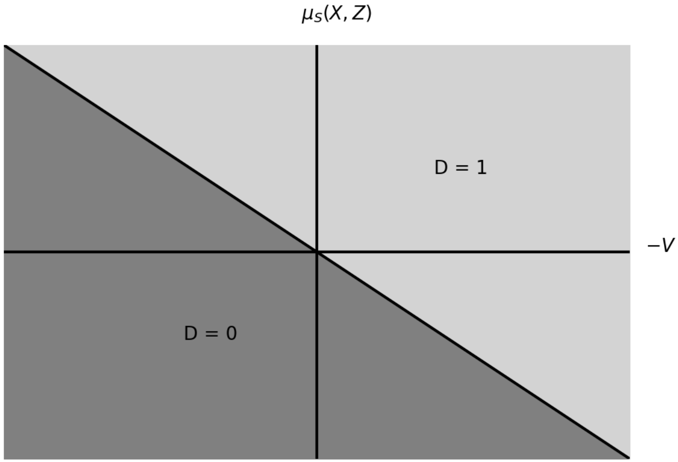
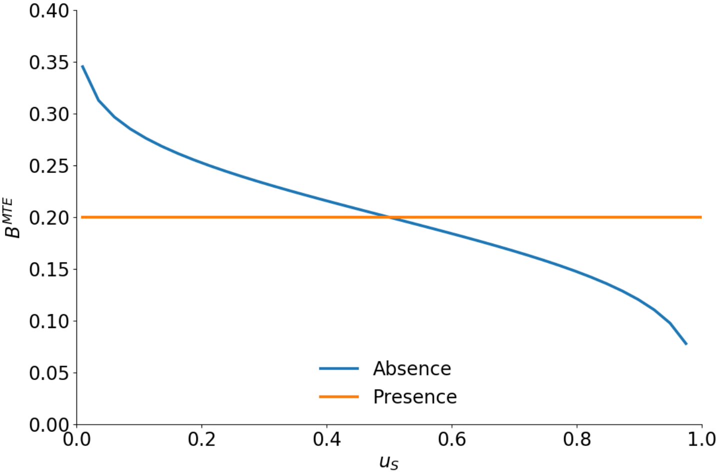


Figure: Marginal Benefit of Treatment



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

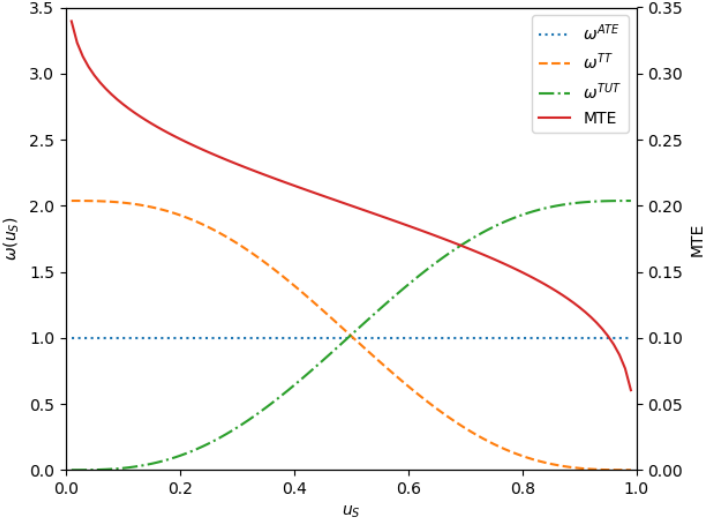
$$\Delta_j(x) = \int_0^1 B^{MTE}(x, u_D)\omega^j(x, u_D)du_D,$$

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

Weights

$$
\begin{aligned}
\omega^{ATE}(x, u_D) &= 1 \\
\omega^{TT}(x, u_D) &= \frac{1 - F_{P|X=x}(u_D)}{E[P \mid X = x]} \\
\omega^{TUT}(x, u_D) &= \frac{F_{P|X=x}(u_D)}{E[1 - P \mid X = x]}
\end{aligned}
$$

Figure: Effects of Treatment as Weighted Averages



Local Average Treatment Effect

Local Average Treatment Effect

- **Local Average Treatment Effect:** Average effect for those induced to change treatment because of a change in the instrument.$\Rightarrow$ instrument-dependent

  parameter

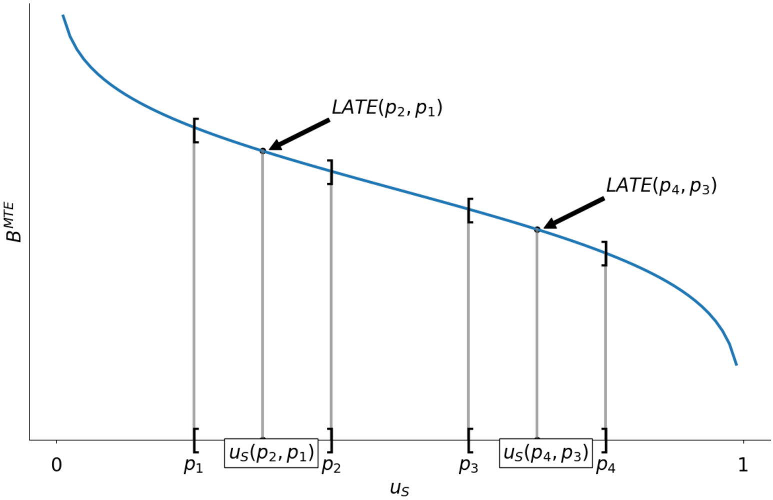
- **Marginal Treatment Effect:** Average effect for those individuals with a given unobserved desire to receive treatment.

  $\Rightarrow$ deep economic parameter

$$B^{LATE} = \frac{E(Y \mid Z = z) - E[Y \mid Z = z']}{P(z) - P(z')}$$

$$B^{LATE}(x, u_D, u_{S'}) = \frac{1}{u_D - u_{D'}} \int_{u_D}^{u_{S'}} B^{MTE}(x, u)du,$$

Figure: Local Average Treatment Effect



Distributions of Effects

Figure: Distribution of Potential Outcomes

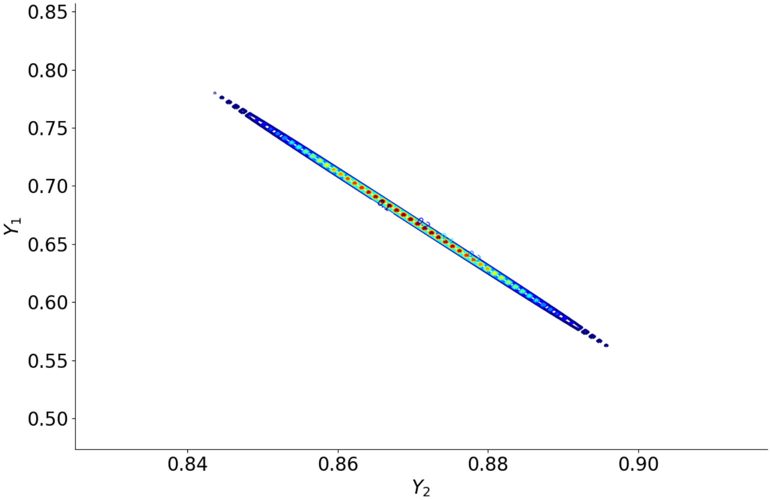
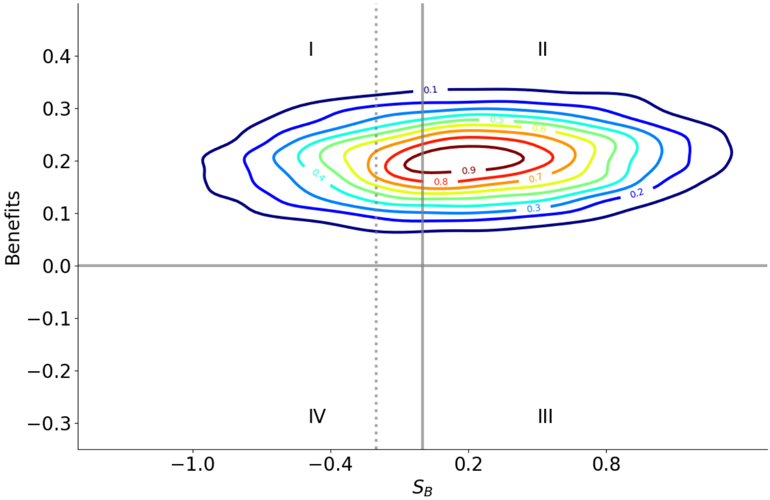


Figure: Distribution of Benefits and Surplus



Appendix

References

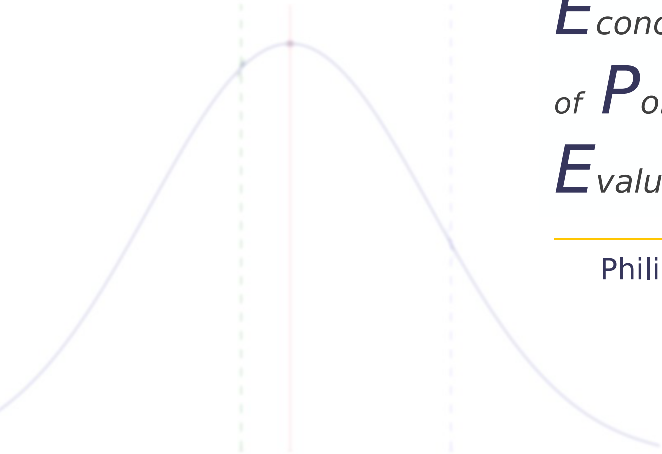
Carneiro, P., Heckman, J. J., & Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, *101*(6), 2754–2781.

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. (2008). Econometric causality. *International Statistical Review*, *76*(1), 1–27.

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

Heckman, J. J., & Vytlacil, E. J. (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica*, *73*(3), 669–739.



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
- $0 < \Pr(D = 1 \mid X) < 1$
- $\ldots$

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

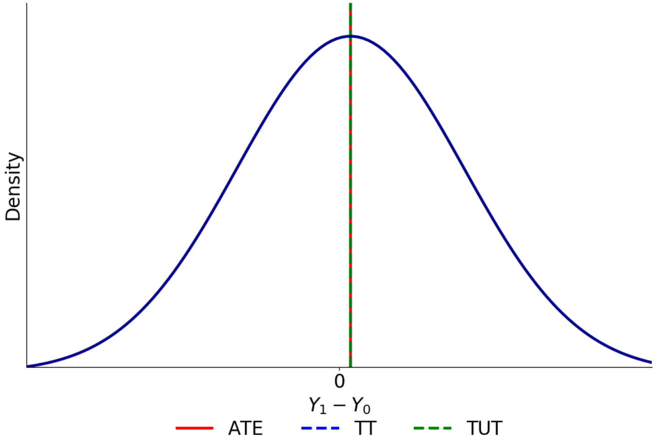
Key Identifying Assumptions

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

When do we have to worry about compliance?

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

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

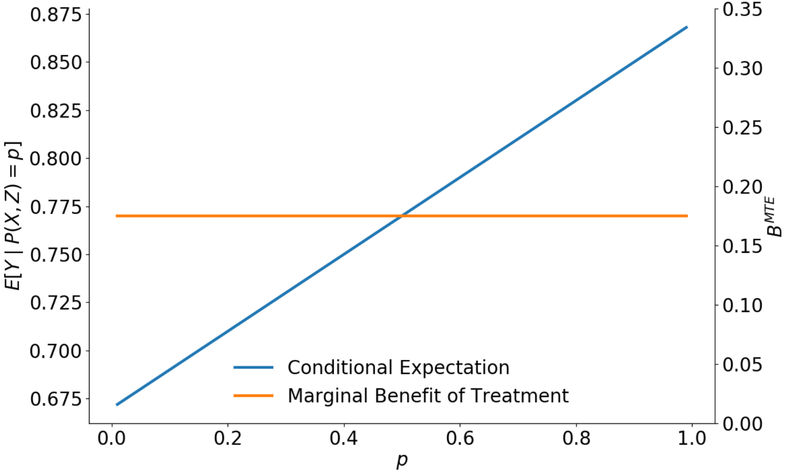
Key Identifiying Assumption

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

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

$$\begin{aligned}
\alpha &= \mu_0(X) \\
\beta &= (Y_1 - Y_0) = \mu_1(X) - \mu_0(X) + (U_1 - U_0) \\
\epsilon &= U_0
\end{aligned}$$

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 ...

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

then the following holds

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

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

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

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

Figure: Observed Outcome and Essential Heterogeneity

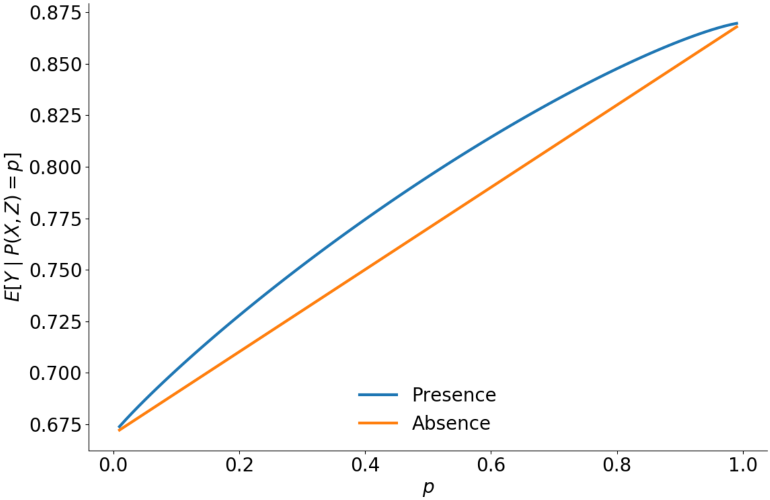
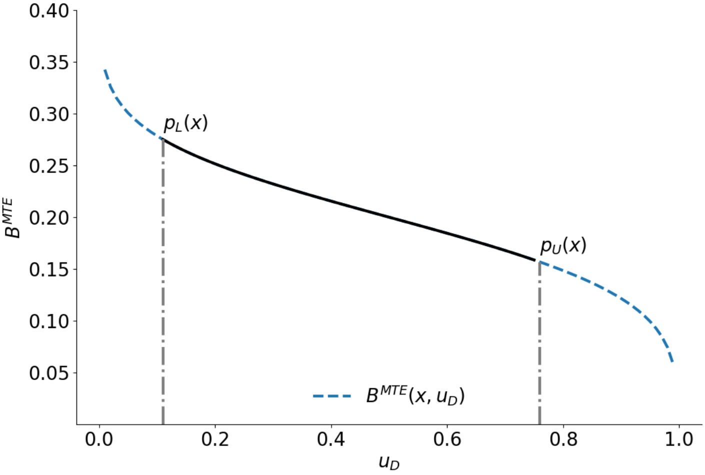


Figure: Identification I



Making $X = x$ explicit

$$E(Y_1 - Y_0|X = x, U_D = u_D)$$
$$= (\mu_1(x) - \mu_0(x)) + E(U_1 - U_0|X = x, U_D = u_D)$$

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

$$E(Y_1 - Y_0|X = x, U_D = u_D)$$
$$= (\mu_1(x) - \mu_0(x)) + E(U_1 - U_0|U_D = u_D)$$

Figure: Identification II

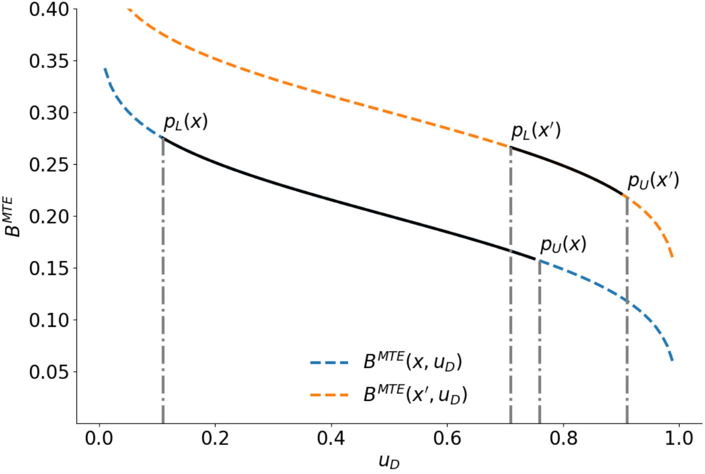
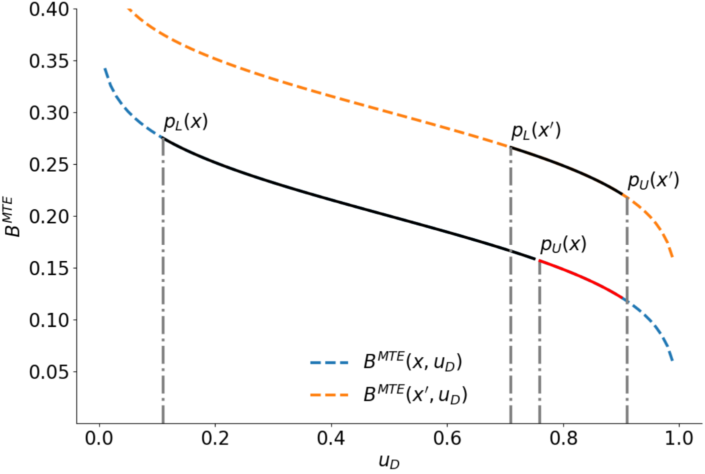


Figure: Identification III



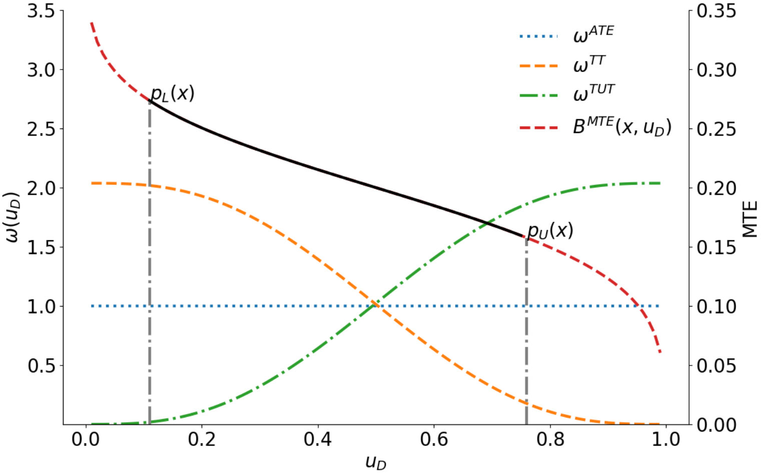
Effects of Treatment as Weighted Averages

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

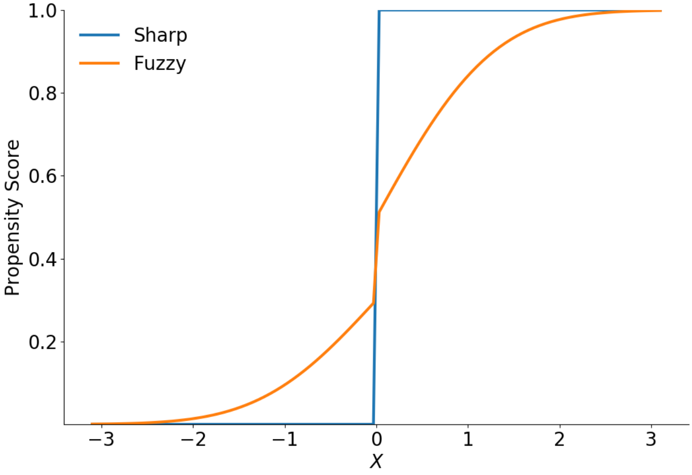


Figure: Observed Outcome in a Sharp Design

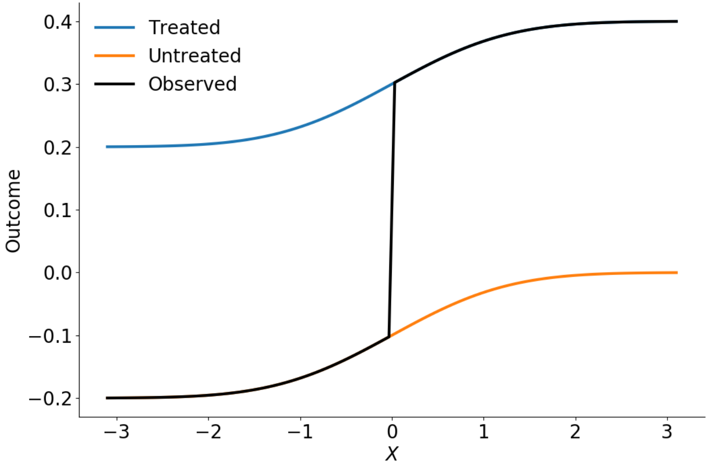
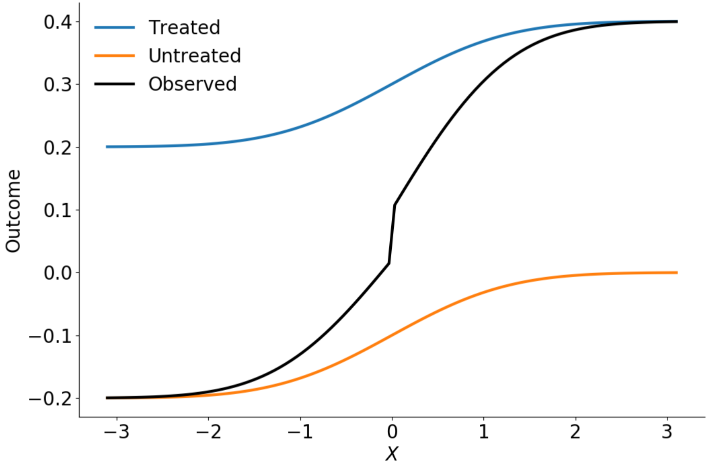


Figure: Observed Outcome in a Fuzzy Design



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.

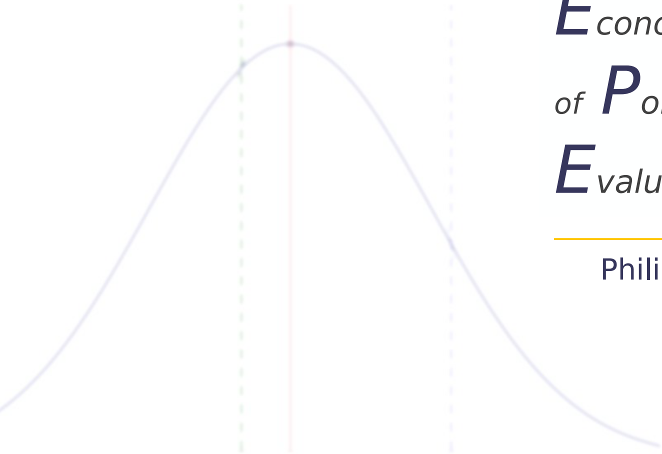
Eliot, T. (1943). Little gidding. In *Four quartets.* Harcourt.

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.



Econometrics of Policy Evaluation

Philipp Eisenhauer

Material available on

Visit us!

grmpy Tutorial

Sebastian Becker

Introduction

grmpy

`grmpy` is an open-source Python package for the simulation and estimation of generalized Roy model. Its main purpose is to serve as a teaching tool to promote the conceptual framework provided by the generalized Roy model to illustrate a variety of issues in the econometrics of policy evaluation.

grmpy

- ***grmpy*** is ...
  - ...an open-source Python Package for the simulation and estimation of the generalized Roy model.
  - ...intended as an useful device to support and improve the understanding of the framework by the opportunity to experience the effect of particular specifications directly.

Setup

Setup

- Normal linear-in-parameters version of the generalized Roy model.

Potential Outcomes	Cost
$Y_1 = \beta_1 X + U_1$	$C = \gamma Z + U_C$
$Y_0 = \beta_0 X + U_0$	

Observed Outcomes	Choice
$Y = DY_1 + (1 - D)Y_0$	$S = Y_1 - Y_0 - C$
	$D = \mathrm{I}[S > 0]$

Features

Features

- *grmpy* is currently capable of the following features:
  - Simulating a dataset based on your own specifications.
  - Providing some useful information about the simulated dataset for instance:
    - Distributional outcome characteristics
    - ATE, TT, TUT
    - MTE by decile
  - Estimating the coefficients of interest given a dataset (of a specific form).

Install the package

- OS, Linux : Use the pip install manager (*pip install grmpy*) or download the package via GitHub and install it manually.
- Windows: The same procedure as for Linux, OS but you have to verify that the numpy package is already installed on your machine.

Initialization file

- The initialization file provides the user the opportunity to specify all parameters of his/her model, for instance:
  - Simulation parameters (number of observations, name of the output files)
  - Estimation parameters (optimization algorithm, start values)
  - Optimization parameters
  - Coefficients and covariance parameters, dummy variables...
- Example
- for a detailed explanation see: *grmpy*-documentation

Simulation

- *grmpy.simulate():* :
  - Input: path of the initialization file.
  - The function returns a data frame based on your specifications and different output files.
    - The data set as a pickle and a txt file.
    - An Info file that provides the distributional characteristics of the data as well as information about the different treatment effects.

Estimation

- *grmpy.estimate():*
  - Input: path of the initialization file.
  - At the moment the estimation process is only capable of two different optimization algorithms:
    - Broyden Fletcher Goldfarb Shanno (BFGS) algorithm
    - Powell’s conjugate direction method

- There are two different options for the start values that could be set in the initialization file:
  - *init:* The estimation process uses the coefficient values specified in the initialization file as the start values for the estimation process.
  - *auto:* The start values are determined via a simple OLS followed by a Probit regression for the choice indicator.
- The estimation results are printed to an output file

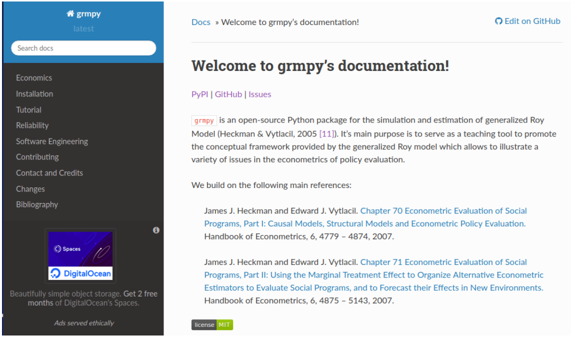
Test battery

- We also provide a test battery that includes several tests to ensure that the processes perform as intended.
  - Property-based testing
  - Reliability testing
  - Regression testing

Application Example

Additional Information

Online documentation



Additional Information

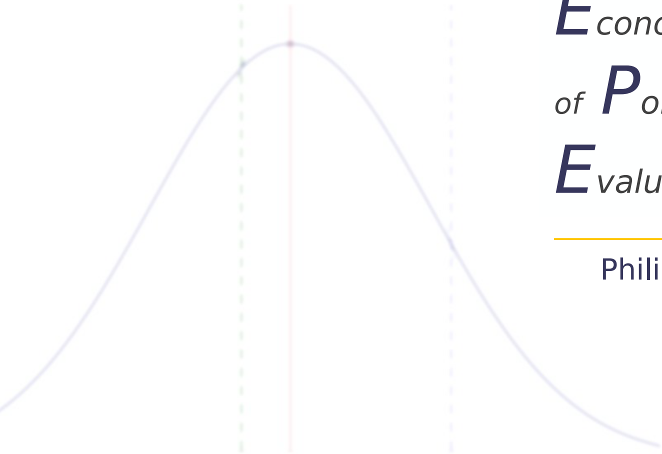
- ***grmpy***-documentation
- Course material regarding the generalized Roy model
- GitHub Repository

Appendix

References

Heckman, J. J., & Vytlacil, E. J. (2007a). Econometric evaluation of social programs, part I: Causal effects, structural models and econometric policy evaluation. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 6B, pp. 4779–4874). Amsterdam, Netherlands: Elsevier Science.

Heckman, J. J., & Vytlacil, E. J. (2007b). Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative economic estimators to evaluate social programs and to forecast their effects in new environments. In J. J. Heckman & E. E. Leamer (Eds.), *Handbook of econometrics* (Vol. 6B, pp. 4875–5144). Amsterdam, Netherlands: Elsevier Science.



Econometrics of Policy Evaluation

Philipp Eisenhauer

Material available on

Visit us!

UNIVERSITÄT BONN

Monte Carlo Exploration

Philipp Eisenhauer

Introduction

The Econometrics of Policy Evaluation

- is important
- is complicated
- is multifaceted

Fundamental Problems

- Evaluation Problem
- Selection problem
  - Essential Heterogeneity

Objects of Interest

- Conventional Average Treatment Effects
- Policy-Relevant Average Treatment Effects
- Local Average Treatment Effect
- Marginal Effect of Treatment
- Distribution of Effects
- Effects on Distribution

Identification Strategies

- Random Assignment
- Matching
- Control Functions and Extensions
- Instrumental Variables

Generalized Roy Model

Potential Outcomes	*Cost*
$Y_1 = \beta_1 X + U_1$	$C = \gamma Z + U_C$
$Y_0 = \beta_0 X + U_0$	

Observed Outcomes	*Choice*
$Y = DY_1 + (1 - D)Y_0$	$S = Y_1 - Y_0 - C$
	$D = \mathrm{I}[S > 0]$

Monte Carlo Exploration

We will touch on all these issues in a Monte Carlo exercise using the **grmpy** package. The notebook is available on the course website.

Appendix

References

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

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