Marginal Treatment Effects: Theory

ECON 31720 Applied Microeconometrics

Francesco Ruggieri

The University of Chicago

November 4, 2020

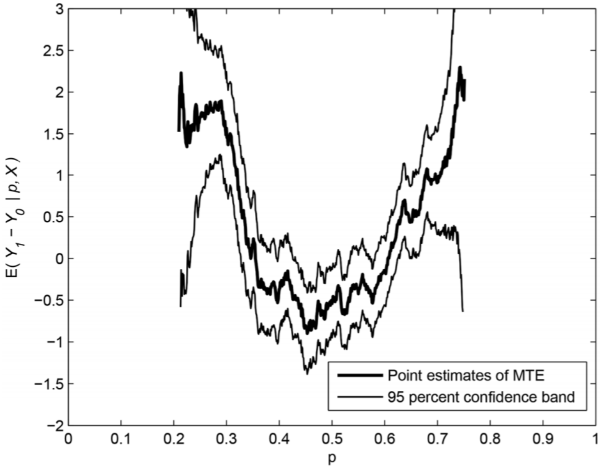
Framework for Marginal Treatment Effects

- $Y \in \mathbb{R}$ is a scalar **outcome** of interest, $D \in \{0,1\}$ is a **binary treatment**
- D and Y are linked by **potential outcomes** $Y(0), Y(1)$
- $X \in \mathbb{R}^{d_x}$ is a vector of predetermined, **observable** characteristics with support $\mathcal{X}$
  - Hereafter, all arguments will be made implicitly conditioning on X
- $U \in \mathbb{R}$ is an **unobserved** and continuously distributed **latent variable**
- $Z \in \mathbb{R}$ is a scalar **instrumental variable** with support $\mathcal{Z}$
  - Z satisfies the **exogeneity** assumption $(Y(0), Y(1), U) \perp\!\!\!\perp Z$

Framework for Marginal Treatment Effects

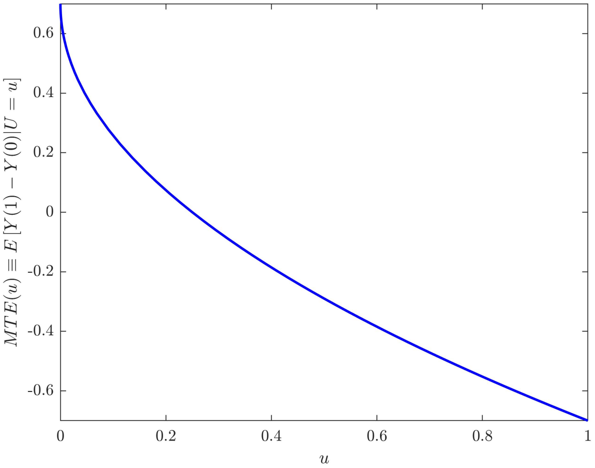
- $\nu(\cdot)$ is an **unknown function** of Z such that $D = \mathbb{I}[U \leq \nu(Z)]$
  - U, $\nu(Z)$ are **additively separable** (no interaction between policy shifters and unobservables)
  - $\nu(Z) - U$ denotes the **net utility** from choosing treatment state $D = 1$
- Without loss, the **selection equation** can be normalized to $D = \mathbb{I}[U \leq p(Z)]$
  - $p(Z) \equiv \mathbb{P}(D = 1|Z)$ is the **propensity score**
  - U is a latent random variable **uniformly** distributed on $[0, 1]$
- $\mathrm{MTE}(u) \equiv \mathbb{E}[Y(1) - Y(0)|U = u]$ is the **Marginal Treatment Effect** of D on Y
  - $\mathrm{MTE}(u)$ is the Average Treatment Effect of D on Y for agents with unobservables $U = u$
- Plotting the Marginal Treatment Effect function is informative about choice heterogeneity

Unobserved Choice Heterogeneity and the MTE Function



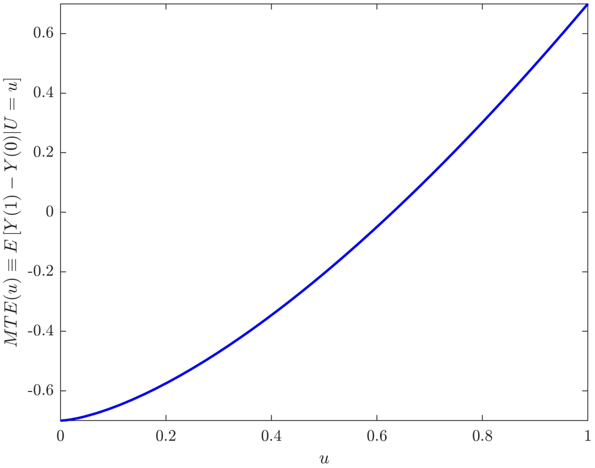
This figure displays the **estimated MTE function** from Brinch, Mogstad, and Wiswall (2017)

Selection on the Gain



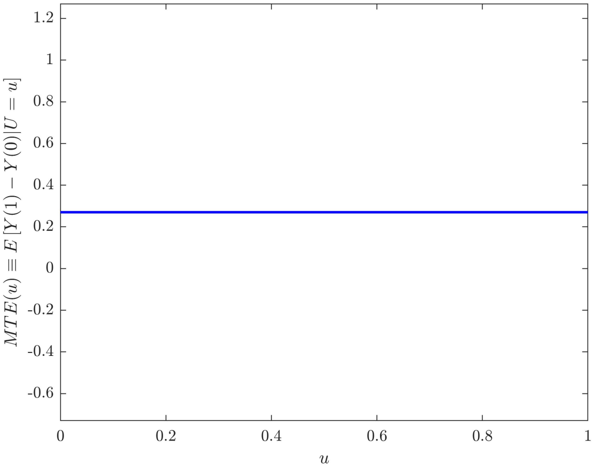
Selection on the gain: **positive correlation** between D and the return from choosing $D = 1$

Selection on the Loss



Selection on the loss: **negative correlation** between D and the return from choosing $D = 1$

Unobserved Homogeneity



Unobserved homogeneity: **zero correlation** between D and the return from choosing $D = 1$

Average Treatment Effect

- Target parameters can be expressed as **weighted averages** of marginal treatment effects

- Consider the **Average Treatment Effect**:

$$\begin{aligned}
\mathrm{ATE} \equiv \mathbb{E}\left[Y(1) - Y(0)\right] &= \mathbb{E}\left[\mathbb{E}\left[Y(1) - Y(0)|U\right]\right] \qquad \text{(LIE)} \\
&= \int_0^1 \mathbb{E}\left[Y(1) - Y(0)|U = u\right] du \qquad (U \sim \mathcal{U}\left[0,1\right]) \\
&= \int_0^1 \mathrm{MTE}(u) du \\
&= \int_0^1 \mathrm{MTE}(u) \times \omega_{\text{ATE}}\, du
\end{aligned}$$

where $\omega_{\text{ATE}} = 1$, i.e., the ATE is a **simple average** of marginal treatment effects

Average Treatment Effect on the Treated

Consider the **Average Treatment Effect on the Treated**:

$$
\begin{aligned}
\mathrm{ATT} \equiv \mathbb{E}\left[Y(1) - Y(0)|D = 1\right] &= \mathbb{E}\left[\mathbb{E}\left[Y(1) - Y(0)|D = 1, p(Z)\right]|D = 1\right] && \text{(LIE)} \\
&= \int_0^1 \mathbb{E}\left[Y(1) - Y(0)|D = 1, p(Z) = p\right] dF_{p(Z)|D=1}(p) \\
&= \int_0^1 \mathbb{E}\left[Y(1) - Y(0)|U \leq p(Z), p(Z) = p\right] dF_{p(Z)|D=1}(p) && (D = \mathbb{I}\left[U \leq p(Z)\right]) \\
&= \int_0^1 \mathbb{E}\left[Y(1) - Y(0)|U \leq p\right] dF_{p(Z)|D=1}(p) && (U \perp\!\!\!\perp Z) \\
&= \int_0^1 \left[\frac{1}{p}\int_0^p \mathbb{E}\left[Y(1) - Y(0)|U = u\right] du\right] dF_{p(Z)|D=1}(p) && (U \sim \mathcal{U}\left[0, 1\right])
\end{aligned}
$$

Average Treatment Effect on the Treated

- In addition, Bayes' rule implies that

$$dF_{p(Z)|D=1} = \frac{\mathbb{P}\left(D = 1|p(Z)\right)}{\mathbb{P}\left(D = 1\right)} dF_{p(Z)} = \frac{p(Z)}{\mathbb{P}\left(D = 1\right)} dF_{p(Z)}$$

- Thus, the Average Treatment Effect on the Treated can be expressed as

$$\begin{aligned}
\mathrm{ATT} &= \int_0^1 \left[\frac{1}{p}\int_0^p \mathbb{E}\left[Y(1) - Y(0)|U = u\right] du\right] \frac{p}{\mathbb{P}\left(D = 1\right)} dF_{p(Z)}(p) \\
&= \frac{1}{\mathbb{P}\left(D = 1\right)} \int_0^1 \left[\int_0^p \mathbb{E}[Y(1) - Y(0)|U = u] du\right] dF_{p(Z)}(p) \\
&= \frac{1}{\mathbb{P}\left(D = 1\right)} \int_0^1 \mathbb{E}[Y(1) - Y(0)|U = u] \left[\int_0^1 \mathbb{I}\left[u \leq p\right] dF_{p(Z)}(p)\right] du \quad \text{(Fubini's)} \\
&= \frac{1}{\mathbb{P}\left(D = 1\right)} \int_0^1 \mathbb{E}[Y(1) - Y(0)|U = u]\, \mathbb{P}\left(u \leq p(Z)\right) du \quad \left(\mathbb{E}\left[\mathbb{I}\left[W\right]\right] = \mathbb{P}\left(W = 1\right)\right)
\end{aligned}$$

Average Treatment Effect on the Untreated

- Rearranging terms:

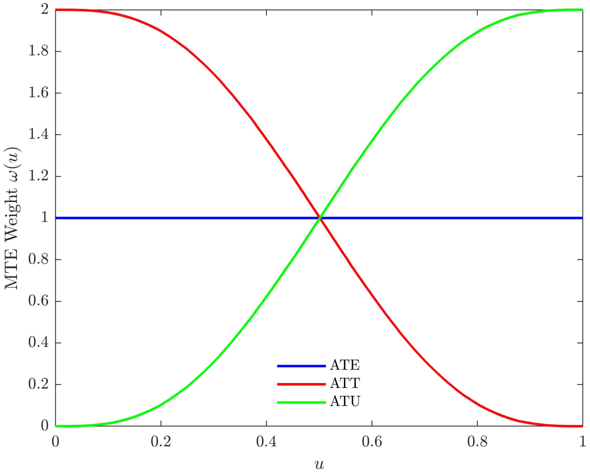
$$\mathrm{ATT} = \int_0^1 \mathrm{MTE}(u) \times \frac{\mathbb{P}\left(u \leq p(Z)\right)}{\mathbb{P}\left(D = 1\right)} du = \int_0^1 \mathrm{MTE}(u) \times \omega_{\mathsf{ATT}}\, du$$

- Analogously, the **Average Treatment Effect on the Untreated** can be expressed as

$$\mathrm{ATU} = \int_0^1 \mathrm{MTE}(u) \times \frac{\mathbb{P}\left(u > p(Z)\right)}{\mathbb{P}\left(D = 0\right)} du = \int_0^1 \mathrm{MTE}(u) \times \omega_{\mathsf{ATU}}\, du$$

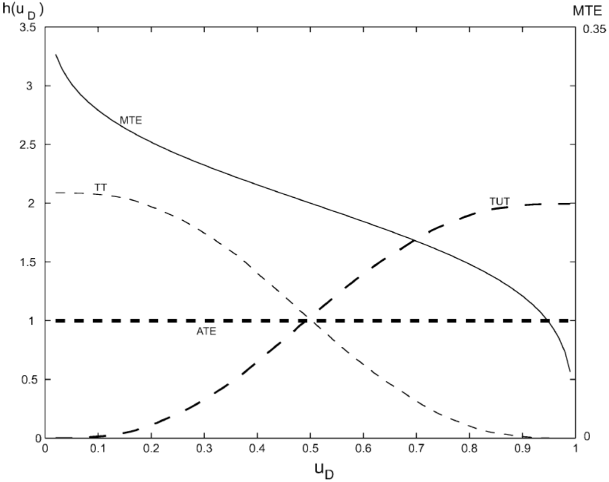
- Intuition: the **ATT (ATU) oversamples** marginal treatment effects for agents who are **more (less)** likely to self-select into treatment state $D = 1$

MTE Weights in a Parametric Normal Roy Model



This figure plots **ATE, ATT, ATU weights** from a parametric normal generalized Roy model

MTE Weights in Heckman and Vytlacil (2005)



Source: Heckman and Vytlacil (2005)

Target Parameters as Weighted Averages of Marginal Treatment Effects

Let us combine information on the **MTE function** and **MTE weights** for target parameters:

- The MTE function is **monotonically decreasing**
  - Agents who **self-select** into treatment state $D = 1$ are more likely to **gain** from it
- The ATT (ATU) weighting function is **monotonically decreasing (increasing)**
  - The ATT **oversamples** MTEs for agents who are more likely to gain from $D = 1$
  - The ATU **undersamples** MTEs for agents who are more likely to gain from $D = 1$
- As a consequence, **selection on the gain** implies $\mathbf{ATT} > \mathbf{ATU}$

Nonparametric Roy vs. Imbens & Angrist with Multivalued Treatment

- The analysis so far has focused on the case in which the **treatment** is **binary**
- Vytlacil (2002) shows that, when $D \in \{0, 1\}$,
  - The nonparametric Roy model **implies** the Imbens and Angrist model
  - The Imbens and Angrist model **implies** the nonparametric Roy model
- Consider the case in which **treatment** is **multivalued**
- With $D \in \mathbb{R}$, the two models are **not nested**:
  - The nonparametric Roy model **does not imply** the Imbens and Angrist model
  - The Imbens and Angrist model **does not imply** the nonparametric Roy model

Nonparametric Roy vs. Imbens & Angrist with Multivalued Treatment

- For simplicity, consider the case in which $D \in \{0, 1, 2\}$ and $Z \in \{0, 1\}$
- The **Imbens and Angrist model** assumes that
  - Either $D(1) \geq D(0)$ or $D(0) \geq D(1)$ with probability one
- The **nonparametric Roy model** assumes that
  - There exists a continuously distributed U and unknown functions ν_1, ν_2 of Z such that

$$D = 1 \times \mathbb{I}\left[\nu_1(Z) < U \leq \nu_2(Z)\right] + 2 \times \mathbb{I}\left[U > \nu_2(Z)\right]$$

    where $\nu_1(z) < \nu_2(z)$ for $z = 0, 1$

Nonparametric Roy vs. Imbens & Angrist with Multivalued Treatment

- Let us assume that the Imbens and Angrist selection model holds

- Without loss, the **monotonicity assumption** is $D(1) \geq D(0)$ with probability one

- The following **inequalities** are consistent with the Imbens and Angrist selection model:
  1. $\mathbb{P}\left(D(0) = 0, D(1) = 0\right) > 0$
  2. $\mathbb{P}\left(D(0) = 1, D(1) = 1\right) > 0$
  3. $\mathbb{P}\left(D(0) = 2, D(1) = 2\right) > 0$
  4. $\mathbb{P}\left(D(0) = 0, D(1) = 1\right) > 0$
  5. $\mathbb{P}\left(D(0) = 1, D(1) = 2\right) > 0$
  6. $\mathbb{P}\left(D(0) = 0, D(1) = 2\right) > 0$

Nonparametric Roy vs. Imbens & Angrist with Multivalued Treatment

- **Potential treatments** can be expressed in terms of U and $\nu_1(Z), \nu_2(Z)$:

$$
\begin{aligned}
D(0) &= 0 \times \mathbb{I}\left[U \leq \nu_1(0)\right] + 1 \times \mathbb{I}\left[\nu_1(0) < U \leq \nu_2(0)\right] + 2 \times \mathbb{I}\left[\nu_2(0) < U\right] \\
D(1) &= 0 \times \mathbb{I}\left[U \leq \nu_1(1)\right] + 1 \times \mathbb{I}\left[\nu_1(1) < U \leq \nu_2(1)\right] + 2 \times \mathbb{I}\left[\nu_2(1) < U\right]
\end{aligned}
$$

- The following **if-and-only-if statements** are true:

$$
\begin{aligned}
D(0) = 0 &\iff U \leq \nu_1(0) & \qquad D(1) = 0 &\iff U \leq \nu_1(1) \\
D(0) = 1 &\iff \nu_1(0) < U \leq \nu_2(0) & \qquad D(1) = 1 &\iff \nu_1(1) < U \leq \nu_2(1) \\
D(0) = 2 &\iff U > \nu_2(0) & \qquad D(1) = 2 &\iff U > \nu_2(1)
\end{aligned}
$$

Nonparametric Roy vs. Imbens & Angrist with Multivalued Treatment

- The **six positive probabilities** consistent with the Imbens and Angrist model are:
  1. $\mathbb{P}(D(0) = 0, D(1) = 0) = \mathbb{P}(U \leq \min\{\nu_1(0), \nu_1(1)\})$
  2. $\mathbb{P}(D(0) = 1, D(1) = 1) = \mathbb{P}(\max\{\nu_1(0), \nu_1(1)\} < U \leq \min\{\nu_2(0), \nu_2(1)\})$
  3. $\mathbb{P}(D(0) = 2, D(1) = 2) = \mathbb{P}(U > \max\{\nu_2(0), \nu_2(1)\})$
  4. $\mathbb{P}(D(0) = 0, D(1) = 1) = \mathbb{P}(\nu_1(1) < U \leq \min\{\nu_1(0), \nu_2(1)\})$
  5. $\mathbb{P}(D(0) = 1, D(1) = 2) = \mathbb{P}(\max\{\nu_1(0), \nu_2(1)\} < U \leq \nu_2(0))$
  6. $\mathbb{P}(D(0) = 0, D(1) = 2) = \mathbb{P}(\nu_2(1) < U \leq \nu_1(0))$
- If $\mathbb{P}(D(0) = 0, D(1) = 2) > 0$, then $\nu_1(0) > \nu_2(1)$. But then $\mathbb{P}(D(0) = 1, D(1) = 1) = 0$
  - This **contradicts** the **strict positivity** of all six probabilities
- Thus, the Imbens and Angrist model **does not imply** the nonparametric Roy model

Nonparametric Roy vs. Imbens & Angrist with Multivalued Treatment

- Let us assume that the nonparametric Roy selection model holds

- Suppose the **unknown functions** $\nu_1(Z)$ and $\nu_2(Z)$ take the following values:

$$\begin{aligned} \nu_1(0) &= 0.4 \qquad & \nu_2(0) &= 0.6 \\ \nu_1(1) &= 0.3 \qquad & \nu_2(1) &= 0.7 \end{aligned}$$

  which meet the condition that $\nu_1(z) < \nu_2(z)$ for $z = 0, 1$

- **Potential treatments** associated with this selection model are

$$\begin{aligned} D(0) &= 0 \times \mathbb{I}\left[U \leq 0.4\right] + 1 \times \mathbb{I}\left[0.4 < U \leq 0.6\right] + 2 \times \mathbb{I}\left[0.6 < U\right] \\ D(1) &= 0 \times \mathbb{I}\left[U \leq 0.3\right] + 1 \times \mathbb{I}\left[0.3 < U \leq 0.7\right] + 2 \times \mathbb{I}\left[0.7 < U\right] \end{aligned}$$

Nonparametric Roy vs. Imbens & Angrist with Multivalued Treatment

- Suppose that the unobservable latent variable is $\boldsymbol{U = 0.35}$. Potential treatments are

$$\boldsymbol{D}(\mathbf{0}) = 0 \times \mathbb{I}\,[0.35 \leq 0.4] + 1 \times \mathbb{I}\,[0.4 < 0.35 \leq 0.6] + 2 \times \mathbb{I}\,[0.6 < 0.35] = \mathbf{0}$$
$$\boldsymbol{D}(\mathbf{1}) = 0 \times \mathbb{I}\,[0.35 \leq 0.3] + 1 \times \mathbb{I}\,[0.3 < 0.35 \leq 0.7] + 2 \times \mathbb{I}\,[0.7 < 0.35] = \mathbf{1}$$

- Suppose that the unobservable latent variable is $\boldsymbol{U = 0.65}$. Potential treatments are

$$\boldsymbol{D}(\mathbf{0}) = 0 \times \mathbb{I}\,[0.65 \leq 0.4] + 1 \times \mathbb{I}\,[0.4 < 0.65 \leq 0.6] + 2 \times \mathbb{I}\,[0.6 < 0.65] = \mathbf{2}$$
$$\boldsymbol{D}(\mathbf{1}) = 0 \times \mathbb{I}\,[0.65 \leq 0.3] + 1 \times \mathbb{I}\,[0.3 < 0.65 \leq 0.7] + 2 \times \mathbb{I}\,[0.7 < 0.65] = \mathbf{1}$$

- $\boldsymbol{D(1) = 1 > 0 = D(0)}$ if $U = 0.35$, but $\boldsymbol{D(1) = 1 < 2 = D(0)}$ if $U = 0.65$
  - This **contradicts** the Imbens and Angrist **monotonicity** assumption

- Thus, the nonparametric Roy model **does not imply** the Imbens and Angrist model

Summary

- The **Marginal Treatment Effect function** is informative about the nature and extent of **unobserved choice heterogeneity** (selection on the gain/loss, unobserved homogeneity)

- **Target parameters** are **weighted averages** of marginal treatment effects
  - The ATT (ATU) oversamples (undersamples) MTEs for agents who more likely choose $D = 1$

- Unlike the case in which the treatment is binary, the **nonparametric Roy** and **Imbens & Angrist** models are **not nested** when the **treatment** is **multivalued**