Framework for Marginal Treatment Effects

Point Identification

- Linear-in-Parameters Models of the MTR Functions
- Partially Linear Models of the MTR Functions

Partial Identification (Mogstad, Santos, and Torgovitsky 2018)

Summary
Framework for Marginal Treatment Effects

Point Identification

- Linear-in-Parameters Models of the MTR Functions
- Partially Linear Models of the MTR Functions

Partial Identification (Mogstad, Santos, and Torgovitsky 2018)

Summary
Framework for Marginal Treatment Effects

- $Y \in \mathbb{R}$ is a scalar \textbf{outcome} of interest, $D \in \{0, 1\}$ is a \textbf{binary treatment}

- $D$ and $Y$ are linked by \textbf{potential outcomes} $Y(0), Y(1)$

- $X \in \mathcal{X} \subseteq \mathbb{R}^{d_x}$ is a vector of predetermined, \textbf{observable} characteristics

- $U \in \mathbb{R}$ is an \textbf{unobserved} and continuously distributed \textbf{latent variable}

- $Z \in \mathcal{Z} \subseteq \mathbb{R}$ is a scalar \textbf{instrumental variable}
  - $Z$ satisfies the conditional \textbf{exogeneity} assumption $(Y(0), Y(1), U) \perp\!\!\!\!\perp Z|X$
Framework for Marginal Treatment Effects

- \( \nu(\cdot) \) is an **unknown function** of \( X \) and \( Z \) such that \( D = I[U \leq \nu(X, Z)] \)
  - \( U, \nu(X, Z) \) are **additively separable** (no interaction between observables and unobservables)
  - \( \nu(X, Z) - U \) denotes the **net utility** from choosing treatment state \( D = 1 \)

- Without loss, the **selection equation** can be normalized to \( D = I[U \leq p(X, Z)] \)
  - \( p(X, Z) \equiv P(D = 1|X, Z) \) is the **propensity score** (also denoted as \( P \))
  - \( U \) is a latent random variable **uniformly** distributed on \([0, 1]\)

- \( \text{MTE}(u) \equiv E[Y(1) - Y(0)|U = u] \) is the **Marginal Treatment Effect** of \( D \) on \( Y \)

- \( \text{MTR}(u)(d|u) \equiv E[Y(d)|U = u] \) is the **Marginal Treatment Response**
  - The Marginal Treatment Effect of \( D \) on \( Y \) at \( U = u \) is \( \text{MTE}(u) = \text{MTR}(1|u) - \text{MTR}(0|u) \)
Identification

- Several standard parameters are \textit{weighted averages} of marginal treatment responses
  - \textbf{Target parameters}: ATE, ATT, ATU, LATE, PRTE, Average Selection Bias
  - \textbf{Estimands}: IV, TSLS, OLS (with and without covariates)

- Multiple identification approaches have been proposed within the MTE framework
  - \textbf{Point identification}: these approaches can be broadly classified into
    - \textbf{Nonparametric}: Heckman and Vytlacil (1999)’s Local IV Estimand if $Z$ is continuous
    - \textbf{Parametric}: linear-in-parameters and partially linear models of the MTR functions
  - \textbf{Partial identification}: Mogstad, Santos, and Torgovitsky (2018)
1 Framework for Marginal Treatment Effects

2 Point Identification
   - Linear-in-Parameters Models of the MTR Functions
   - Partially Linear Models of the MTR Functions

3 Partial Identification (Mogstad, Santos, and Torgovitsky 2018)

4 Summary
1 Framework for Marginal Treatment Effects

2 Point Identification
   • Linear-in-Parameters Models of the MTR Functions
   • Partially Linear Models of the MTR Functions

3 Partial Identification (Mogstad, Santos, and Torgovitsky 2018)

4 Summary
A general linear-in-parameters model of the MTR functions is

\[
\text{MTR}(d|u, x) \equiv \mathbb{E}[Y(d)|U = u, X = x] = \sum_{k=1}^{\bar{k}} \theta_k b_k (d|u, x) \quad \text{for } d = 0, 1
\]

where \(\{\theta_k\}_{k=1}^{\bar{k}}\) are unknown coefficients and \(\{b_k\}_{k=1}^{\bar{k}}\) are known functions.

When constructing a linear-in-parameters model, a researcher must choose:

- Whether to allow for additive separability between \(U\) and \(X\)
- The order of the polynomials of \(U\) and \(X\) and/or the sieve for \(U\) and \(X\)

If observables and unobservables are assumed not to be additively separable:

\[
\text{MTR}(d|u, x) \equiv \mathbb{E}[Y(d)|X = x, U = u] = \alpha_d + \beta_d u + x'\gamma_d + ux'\delta_d \quad \text{for } d = 0, 1
\]
\[ \mathbb{E}[Y|D = 1, P = u, X = x] = \mathbb{E}[DY(1) + (1 - D)Y(0)|D = 1, P = u, X = x] \\
= \mathbb{E}[Y(1)|D = 1, P = u, X = x] \\
= \mathbb{E}[Y(1)|U \leq P, P = u, X = x] \quad (D = \mathbb{I}[U \leq p(X, Z)]) \\
= \mathbb{E}[Y(1)|U \leq u, X = x] \quad (Z \perp \!\!\!\!\perp U|X) \\
= \frac{1}{u} \int_{0}^{u} \mathbb{E}[Y(1)|W = w, X = x]dw \quad (U \sim \mathcal{U}[0, 1]) \\
= \frac{1}{u} \int_{0}^{u} \left[ \alpha_1 + \beta_1 w + x'\gamma_1 + wx'\delta_1 \right]dw \\
= \frac{1}{u} \left[ \alpha_1 u + \frac{\beta_1}{2} u^2 + ux'\gamma_1 + u^2x'\delta_1 \right] \\
= \alpha_1 + \frac{\beta_1}{2} u + x'\gamma_1 + ux'\delta_1/2 \]
Point Identification: Linear-in-Parameters Models of the MTR Functions

• Thus: \[ E[Y|D = 1, P = u, X = x] = \alpha_1 + \frac{\beta_1}{2} u + x' \gamma_1 + ux' \delta_1 \]

• Analogously: \[ E[Y|D = 0, P = u, X = x] = \left( \alpha_0 + \frac{\beta_0}{2} \right) + \frac{\beta_0}{2} u + x' \left( \gamma_0 + \frac{\delta_0}{2} \right) + ux' \frac{\delta_0}{2} \]

• Goal: **point identify** parameters \( \{\alpha_d, \beta_d, \gamma_d, \delta_d\}_{d\in\{0,1\}} \) of the linear MTR functions

• Implementation: **regress** \( Y \) on 1, \( P \), \( X \), and \( PX \) separately for units with \( D \in \{0,1\} \)

\[ Y = \alpha_0^* + \beta_0^* P + X' \gamma_0^* + PX' \delta_0^* + R_d \quad \text{for } d = 0, 1 \]

• Back out **MTR parameters** using **regression coefficients**:

\[
\begin{align*}
\alpha_1 &= \alpha_1^* \\
\beta_1 &= 2\beta_1^* \\
\gamma_1 &= \gamma_1^* \\
\delta_1 &= 2\delta_1^* \\
\alpha_0 &= \alpha_0^* - \beta_0^* \\
\beta_0 &= 2\beta_0^* \\
\gamma_0 &= \gamma_0^* - \delta_0^* \\
\delta_0 &= 2\delta_0^*
\end{align*}
\]
Gelbach (2002)

- Example: “Public Schooling for Young Children and Maternal Labor Supply” (AER, 2002)

- This paper by Jonah Gelbach provides an interesting setup for the MTE framework

- **Goal**: estimate the effect of public school enrollment on women’s labor supply

- Public school enrollment is not as-good-as randomly assigned
  - Parents may choose to hold their children back a year or enroll them in private school

- **Institutional framework**: parents’ ability to enroll a child in public kindergarten in the academic year during which the child turns five depends on the calendar date of the child’s birth

- **Empirical strategy**: instrument public school enrollment with child’s quarter of birth
Gelbach (2002)

- The author’s **TSLS estimate** is $\approx 2.71$ and statistically significant at conventional levels.

- However, it is hard to provide a clear **economic interpretation** to this estimate:
  - The main specification conditions linearly on **covariates** and uses **four instruments**.
  - The TSLS estimand is a weighted average (likely with **negative weights**) of treatment effects.

- Let us explore treatment effect **heterogeneity** in a MTE framework.

- A **linear-in-parameters** model of the MTR functions:

  $$\text{MTR}(d|u, x) \equiv \mathbb{E}[Y(d)|X = x, U = u] = \alpha_d + \beta_d u + x'\gamma_d + u x'\delta_d \quad \text{for} \ d = 0, 1$$

  where $D \in \{0, 1\}$ denotes public school enrollment and $X$ is a vector of covariates.
This figure plots the estimated MTE function, where the vector $X$ is evaluated at its mean.
Gelbach (2002)

- The child’s *quarter-of-birth* instrument vector is defined as

\[
Z ≡ \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{bmatrix} = \begin{bmatrix} \mathbb{I} [\text{QOB }= \text{Q2-1974}] \\ \mathbb{I} [\text{QOB }= \text{Q3-1974}] \\ \mathbb{I} [\text{QOB }= \text{Q4-1974}] \\ \mathbb{I} [\text{QOB }= \text{Q1-1975}] \end{bmatrix}
\]

- The estimated MTE function can be used to compute **interpretable target parameters**:

\[
\text{LATE}_{z_4 \rightarrow z_3}(x) = \int_0^1 \text{MTE}(u, x) \frac{\mathbb{I} \left[ \tilde{\rho}(x, z_4) < u \leq \tilde{\rho}(x, z_3) \right]}{\tilde{\rho}(x, z_3) - \tilde{\rho}(x, z_4)} du \approx 3.45
\]

\[
\text{LATE}_{z_3 \rightarrow z_2}(x) = \int_0^1 \text{MTE}(u, x) \frac{\mathbb{I} \left[ \tilde{\rho}(x, z_3) < u \leq \tilde{\rho}(x, z_2) \right]}{\tilde{\rho}(x, z_2) - \tilde{\rho}(x, z_3)} du \approx 2.77
\]

\[
\text{LATE}_{z_2 \rightarrow z_1}(x) = \int_0^1 \text{MTE}(u, x) \frac{\mathbb{I} \left[ \tilde{\rho}(x, z_2) < u \leq \tilde{\rho}(x, z_1) \right]}{\tilde{\rho}(x, z_1) - \tilde{\rho}(x, z_2)} du \approx 2.38
\]
Gelbach (2002)

- Enrolling a child in public school in Q1-1975 implies the child is not even five years old.

- Mothers who are willing to do so are likely to be more sensitive to public subsidies than mothers who are shifted into the treated arm when a child was born in Q2-1974.

- This unobserved heterogeneity may explain $\widehat{\text{LATE}}_{z_4 \rightarrow z_3} > \widehat{\text{LATE}}_{z_3 \rightarrow z_2} > \widehat{\text{LATE}}_{z_2 \rightarrow z_1}$.
  - A mother's opportunity cost of not working (i.e., her return from working) is increasing in her willingness to delay the enrollment of a five-year old child in a public kindergarten.

- Modeling the MTR functions allows an empiricist to analyze unobserved heterogeneity.

- Linear-in-parameters models of the MTR functions are not the only option...
1 Framework for Marginal Treatment Effects

2 Point Identification
   • Linear-in-Parameters Models of the MTR Functions
   • Partially Linear Models of the MTR Functions

3 Partial Identification (Mogstad, Santos, and Torgovitsky 2018)

4 Summary
Point Identification: Partially Linear Models of the MTR Functions

• An alternative to linear-in-parameters models is partially linear models

• A common partially linear model of the MTR functions is

\[
\text{MTR}(d|u, x) \equiv \mathbb{E} [Y(d)|X = x, U = u] = g_d(u) + x' \beta_d
\]

where \(g_d\) is an unknown function of the latent variable \(U\)

• This model assumes additive separability between observables and unobservables

• Point identification of the MTR and MTE functions follows Robinson (1988)
\[ \mathbb{E}[Y|D = 1, P = u, X = x] = \mathbb{E}[DY(1) + (1 - D)Y(0)|D = 1, P = u, X = x] 
\]
\[ = \mathbb{E}[Y(1)|D = 1, P = u, X = x] 
\]
\[ = \mathbb{E}[Y(1)|U \leq P, P = u, X = x] \quad (D = 1_{[U \leq p(X, Z)]}) 
\]
\[ = \mathbb{E}[Y(1)|U \leq u, X = x] \quad (Z \perp U|X) 
\]
\[ = \frac{1}{u} \int_0^u \mathbb{E}[Y(1)|W = w, X = x]dw \quad (U \sim U[0, 1]) 
\]
\[ = \frac{1}{u} \int_0^u (g_1(w) + x'\beta_1)dw 
\]
\[ = \frac{1}{u} \left(ux'\beta_1 + \int_0^u g_1(w)dw \right) 
\]
\[ = x'\beta_1 + \frac{1}{u} \int_0^u g_1(w)dw 
\]
Point Identification: Partially Linear Models of the MTR Functions

• Thus: \( \mathbb{E}[Y|D = 1, P = u, X = x] = x'\beta_1 + \frac{1}{u} \int_0^u g_1(w) \, dw \)

• Analogously: \( \mathbb{E}[Y|D = 0, P = u, X = x] = x'\beta_0 + \frac{1}{1-u} \int_u^1 g_0(w) \, dw \)

• The Law of Iterated Expectations implies that

\[
\begin{align*}
\mathbb{E}[Y|P = u, X = x] &= \mathbb{E}[Y|D = 1, P = u, X = x] \times \mathbb{P}(D = 1|P = u, X = x) \\
&\quad + \mathbb{E}[Y|D = 0, P = u, X = x] \times \mathbb{P}(D = 0|P = u, X = x) \\
&= \mathbb{E}[Y|D = 1, P = u, X = x] \times u \\
&\quad + \mathbb{E}[Y|D = 0, P = u, X = x] \times (1 - u) \\
&= ux'\beta_1 + \int_0^u g_1(w) \, dw + (1 - u)x'\beta_0 + \int_u^1 g_0(w) \, dw \\
&= x'\beta_0 + ux'(\beta_1 - \beta_0) + \int_0^u g_1(w) \, dw + \int_u^1 g_0(w) \, dw
\end{align*}
\]
Point Identification: Partially Linear Models of the MTR Functions

• Under this parameterization, the **conditional mean of the observed outcome** is

\[
\mathbb{E}[Y|P = u, X = x] = x'\beta_0 + ux'(\beta_1 - \beta_0) + \bar{g}(u)
\]

where \(\bar{g}(u) \equiv \int_0^u g_1(w) \, dw + \int_u^1 g_0(w) \, dw\) is an **unknown function** of the latent variable

• In a **linear-in-parameters** model, \(\bar{g}(u)\) would be **sieved**

• In a **partially linear** model, \(\bar{g}(u)\) can be estimated with a **kernel-based approach**

• The goal is to **point identify** the **Marginal Treatment Effect function**:

\[
\text{MTE}(u, x) = \text{MTR}(1|u, x) - \text{MTR}(0|u, x)
\]

\[
= (g_1(u) + x'\beta_1) - (g_0(u) + x'\beta_0)
\]

\[
= x'\beta_1 - x'\beta_0 + g_1(u) - g_0(u)
\]
Point Identification: Partially Linear Models of the MTR Functions

- Using the same derivation as Heckman and Vytlacil (1999)’s **Local IV Estimand**:

\[ \text{MTE}(u, x) = \frac{\partial}{\partial p} E[Y|P = p, X = x] \bigg|_{p=u} = x'(\beta_1 - \beta_0) + \bar{g}'(u) \]

- Combining the two previous expressions for \( \text{MTE}(u, x) \):

\[ \text{MTE}(u, x) = x'(\beta_1 - \beta_0) + g_1(u) - g_0(u) = x'(\beta_1 - \beta_0) + \bar{g}'(u) \]

- This is not surprising if one exploits the definition of \( \bar{g}(u) \):

\[ \bar{g}'(u) = \frac{\partial}{\partial u} \left( \int_0^u g_1(w) \, dw + \int_u^1 g_0(w) \, dw \right) = g_1(u) - g_0(u) \]

which follows from an application of **Leibniz’s rule**

- **Implication**: estimating the MTE function entails estimating the **derivative** of \( \bar{g}(U) \)
Point Identification: Partially Linear Models of the MTR Functions

• Identification of the MTEs in this class of partially linear models follows Robinson (1988)

• Recall that the conditional mean of the observed outcome is
  \[ E[Y|P, X] = X'\beta_0 + PX'(\beta_1 - \beta_0) + \bar{g}(P) \]

• The Law of Iterated Expectations implies that
  \[
  = E[X'\beta_0 + PX'(\beta_1 - \beta_0) + \bar{g}(P)|P] \\
  = E[X'|P]\beta_0 + PE[X'|P](\beta_1 - \beta_0) + \bar{g}(P)
  \]

• Define \( \tilde{Y} \equiv Y - E[Y|P] \) and \( \tilde{X} \equiv X - E[X|P] \), then add and subtract \( E[Y|X, P] \):
  \[
  \]
Point Identification: Partially Linear Models of the MTR Functions

- Replace \( \mathbb{E}[Y|P,X] \) and \( \mathbb{E}[Y|P] \) with their expressions above:

\[
\tilde{Y} = \tilde{X}'\beta_0 + P\tilde{X}'(\beta_1 - \beta_0) + R
\]

where \( R \) is a residual defined as \( R \equiv Y - \mathbb{E}[Y|X,P] \)

- By the Law of Iterated Expectations, this residual has two convenient properties:
  - It is **mean independent** of \( X \):
    \[
    \]
  - It is **mean independent** of \( P \):
    \[
    \]

- If \( \mathbb{E}[R|X] = \mathbb{E}[R|P] = 0 \), both \( \beta_0 \) and \( \beta_1 - \beta_0 \) are linear regression coefficients
Point Identification: Partially Linear Models of the MTR Functions

1. **Estimate** $\tilde{Y} \equiv Y - \mathbb{E}[Y|P]$ and $\tilde{X} \equiv X - \mathbb{E}[X|P]$ **nonparametrically** ($P$ is a scalar)
   
   - Local constant regression suffers from boundary bias, so local linear regression is preferable

2. Perform a **linear regression** of $\tilde{Y}$ on $\tilde{X}$ and $P\tilde{X}$ and store the estimated $\beta_0$ and $\beta_1$

3. **Estimate** $\overline{g}(P)$

   - The mean of $Y$ conditional on $P$, derived above, can be rearranged as
     \[
     \overline{g}(P) = \mathbb{E}[Y - X' \beta_0 - PX' (\beta_1 - \beta_0) | P]
     \]

   - $Y$, $X$, $P$, $\beta_0$, and $\beta_1$ are now **known**, so $\overline{g}(P)$ can be estimated **nonparametrically**

   - Recall that $\text{MTE}(u, x) = x' (\beta_1 - \beta_0) + \overline{g}'(u)$, so $\overline{g}'$ is of interest

   - Local linear suffers from boundary bias in the first derivative, **local quadratic** is preferable
This figure plots the estimated MTE function ($X = \bar{x}$) using data from Gelbach (AER, 2002)
1 Framework for Marginal Treatment Effects

2 Point Identification
   - Linear-in-Parameters Models of the MTR Functions
   - Partially Linear Models of the MTR Functions

3 Partial Identification (Mogstad, Santos, and Torgovitsky 2018)

4 Summary

• **Target parameters** and **common estimands** are weighted averages of the MTR pairs

• **Target parameters** (ATE, ATT, ATU, LATE, PRTE, Average Selection Bias):

\[
\beta^* = \Gamma^*(m) \equiv \mathbb{E} \left[ \int_0^1 m_0(u, X) \omega_0^*(u, Z) \, du \right] + \mathbb{E} \left[ \int_0^1 m_1(u, X) \omega_1^*(u, Z) \, du \right]
\]

• **Common estimands** (IV, TSLS, OLS with and without covariates):

\[
\beta_s = \Gamma_s(m) \equiv \mathbb{E} \left[ \int_0^1 m_0(u, X) \omega_{0s}(u, Z) \, du \right] + \mathbb{E} \left[ \int_0^1 m_1(u, X) \omega_{1s}(u, Z) \, du \right]
\]

where \( \omega_{0s}(u, z) \equiv s(0, z) \times \mathbb{I}[u > p(z)] \) and \( \omega_{1s}(u, z) \equiv s(1, z) \times \mathbb{I}[u \leq p(z)] \)

• \( \Gamma^* \) and \( \Gamma_s \) are **identified linear maps** of the MTR functions

- **Estimands** $\beta_s$ are functions of the data and are thus known.

- **Weights** $\omega^*_d(U, Z)$ and $\omega^*_{ds}(U, Z)$, for $d = 0, 1$, are functions of the data and identified.

- The Marginal Treatment Response functions, $m_d(U, X)$ for $d = 0, 1$, are unknown.
  - As a consequence, target parameters are unknown.

- Intuition: bound target parameters such that the implied MTR functions are “consistent” with the data, i.e., they match known estimands via their (identified) weights.

- Formally, these bounds solve two convex optimization problems:

  $\underline{\beta}^* \equiv \inf_{m \in \mathcal{M}_S} \Gamma^*(m)$
  $\overline{\beta}^* \equiv \sup_{m \in \mathcal{M}_S} \Gamma^*(m)$

  where $\mathcal{M}_S \equiv \{m \in \mathcal{M} : \Gamma_s(m) = \beta_s \text{ for all } s \in S\}$.

- **Issue**: the parameter space of MTR functions, $\mathcal{M}$, is possibly **infinite-dimensional**

- **Solution**: replace $\mathcal{M}$ with a **finite-dimensional subset** $\mathcal{M}_{fd} \subseteq \mathcal{M}$

  $\mathcal{M}_{fd}$ could be specified as the **finite linear basis**

  \[
  \mathcal{M}_{fd} \equiv \left\{ (m_0, m_1) \in \mathcal{M} : m_d (u, x) = \sum_{k=1}^{\bar{k}_d} \theta_{dk} b_{dk} (u, x) \text{ for some } \{\theta_{dk}\}_{k=1}^{\bar{k}_d}, d = 0, 1 \right\}
  \]

  where $\{\theta_{dk}\}_{k=1}^{\bar{k}_d}$ are **unknown coefficients** and $\{b_{dk}\}_{k=1}^{\bar{k}_d}$ are **known basis functions**

- This is effectively a **parameterization** of the Marginal Treatment Response functions

• Parameterizing MTR functions as finite linear bases reduces the optimization problems to

\[
\bar{\beta}_{fd}^* \equiv \sup_{\theta_0, \theta_1 \in \Theta} \sum_{k=1}^{k_0} \theta_0k \Gamma_0^* (b_{0k}) + \sum_{k=1}^{k_1} \theta_1k \Gamma_1^* (b_{1k})
\]

s.t.

\[
\sum_{k=1}^{k_0} \theta_0k \Gamma_0s (b_{0k}) + \sum_{k=1}^{k_1} \theta_1k \Gamma_1s (b_{1k}) = \beta_s \quad \text{for all } s \in S
\]

and analogously for \( \beta_{fd}^* \)

• Recall that the (identified) linear maps of the MTR functions are

\[
\Gamma_d^* (m_d) = \mathbb{E} \left[ \int_0^1 m_d (u, X) \omega_d^* (u, Z) \, du \right] \quad \Gamma_{ds} (m_d) = \mathbb{E} \left[ \int_0^1 m_d (u, X) \omega_{ds} (u, Z) \, du \right]
\]
Mogstad, Santos, and Torgovitsky (2018) considers two main sets of finite linear basis:

1. **Bernstein Polynomials**: the $k$th Bernstein basis polynomial of degree $\bar{k}$ is

$$b_k^\bar{k} : [0, 1] \to \mathbb{R} \text{ s.t. } b_k^\bar{k}(u) \equiv \binom{\bar{k}}{k} u^k (1 - u)^{\bar{k}-k} \text{ for } k = 0, 1, \ldots, \bar{k}$$

2. **Constant Splines** for exact computation of nonparametric bounds

   - Suppose $Z$ has discrete support and $\omega_d^*(u, z)$, $d = 0, 1$, are piecewise constant in $u$
   - Define a partition $\{U_j\}_{j=1}^{\bar{j}}$ of $[0, 1]$ such that $\omega_d^*(u, z)$, $\mathbb{I}[u \leq p(z)]$ are constant in each $U_j$
   - Construct the basis functions

$$b_{jl}(u, x) \equiv \mathbb{I}[u \in U_j, x = x_l] \text{ for } 1 \leq j \leq \bar{j} \text{ and } 1 \leq l \leq \bar{l}$$

whose linear combinations form constant splines over $[0, 1]$ for each $x$

• For illustration purposes, the MTR functions are assumed to be known:

\[m_0(u) = 0.6(1-u)^2 + 0.4u(1-u) + 0.3u^2\]
\[m_1(u) = 0.75(1-u)^2 + 0.5u(1-u) + 0.25u^2\]

• Outcome: \(Y \in \{0, 1\}\) is trivially bounded

• Instrument: \(Z \in \{0, 1, 2\}\), with \(P(Z = 0) = 0.5, P(Z = 1) = 0.4, P(Z = 2) = 0.1\)
  
  • Note: some of the paper’s figures incorrectly refer to \(Z \in \{1, 2, 3\}\) rather than \(Z \in \{0, 1, 2\}\)

• Propensity scores: \(p(0) = 0.35, p(1) = 0.6, p(2) = 0.7\)

• Target parameter: \( \text{LATE} (0.35, 0.9) \equiv E[Y(1) - Y(0) | U \in (0.35, 0.9)]\)
  
  • This target parameter requires extrapolation since the complier subpopulation is expanded
This figure plots the **DGP MTE function** in Mogstad, Santos, and Torgovitsky (2018)

Nonparametric bounds: [-0.421, 0.500]

This figure plots maximizing MTRs when using only the IV slope coefficient.
This figure plots maximizing MTRs when using both the **IV and OLS** slope coefficients.
This figure plots maximizing MTRs when breaking the IV slope into two components.

Nonparametric bounds: [-0.138, 0.407]

This figure plots maximizing MTRs when using all IV-like estimands (sharp bounds)
This figure plots maximizing MTRs when restricted to be decreasing.

Order 9 polynomial bounds, MTRs decreasing: [0.000, 0.067]

This figure plots maximizing MTRs when further **restricted** to be a 10th-order **polynomial**
1 Framework for Marginal Treatment Effects

2 Point Identification
   - Linear-in-Parameters Models of the MTR Functions
   - Partially Linear Models of the MTR Functions

3 Partial Identification (Mogstad, Santos, and Torgovitsky 2018)

4 Summary
Summary

- **Target parameters** and **common estimands** are weighted averages of MTRs.

- Within a MTE framework, **point identification** of target parameters usually entails:
  1. Specifying **linear-in-parameters** models of the MTR functions, or
  2. Specifying **partially linear** models of the MTR functions.

- Within a MTE framework, **partial identification** of target parameters entails computing **bounds** such that the implied MTR functions are consistent with known estimands.