Part 8: Policy Evaluation- Marginal Treatment Effects

Chris Conlon

March 31, 2020

Applied Econometrics

Marginal Treatment Effects

One quantity to rule them all: MTE

- Consider a treatment effect $\beta_i = Y_i(1) - Y_i(0)$.
- Think about a single-index such that $T_i = 1(v_i \leq Z_i'\gamma)$.
- Think about the person for whom $v_i = Z_i'\gamma$ (just barely untreated).

$$\Delta^{MTE}(X_i, v_i) = E[\beta_i | X_i, v_i = Z_i'\gamma]$$

- MTE is average impact of receiving a treatment for everyone with the same $Z'\gamma$.
- For any single index model we can rewrite

$$T_i = 1(v_i \leq Z_i'\gamma) = 1(u_{is} \leq F(Z_i'\gamma)) \text{ for } u_s \in [0,1]$$

- F is just the cdf of v_i
- Now we can write $P(Z) = Pr(T = 1|Z) = F(Z'\gamma)$.

MTE: Derivation

Now we can write,

$$\begin{aligned} Y_0 &= \gamma_0' X + U_0 \\ Y_1 &= \gamma_1' X + U_1 \end{aligned}$$

$P(T=1|Z) = P(Z)$ works as our instrument with two assumptions:

1. $(U_0, U_1, u_s) \perp P(Z)|X$. (Exogeneity)
2. Conditional on X there is enough variation in Z for $P(Z)$ to take on all values $\in (0,1)$.
   - This is much stronger than typical relevance condition. Much more like the special regressor method we will discus next time.

MTE: Derivation

Now we can write,

$$\begin{aligned} Y &= \gamma_0' X + T(\gamma_1 - \gamma_0)' X + U_0 + T(U_1 - U_0) \\ E[Y|X, P(Z) = p] &= \gamma_0' X + p(\gamma_1 - \gamma_0)' X + E[T(U_1 - U_0)|X, P(Z) = p] \end{aligned}$$

Observe $T = 1$ over the interval $u_s = [0, p]$ and zero for higher values of u_s. Let $U_1 - U_0 \equiv \eta$.

$$\begin{aligned} E[T(U_1 - U_0)|P(Z) = p, X] &= \int_{-\infty}^{\infty} \int_0^p (U_1 - U_0) f((U_1 - U_0)|U_s = u_s) du_s d(U_1 - U_0) \\ E[T(\eta)|P(Z) = p, X] &= \int_{-\infty}^{\infty} \int_0^p \eta f(\eta|U_s = u_s) d\eta d u_s \end{aligned}$$

$$\begin{aligned} \Delta^{MTE}(p) &= \frac{\partial E[Y|X, P(Z) = p]}{\partial p} = (\gamma_1 - \gamma_0)' X + \int_{-\infty}^{\infty} \eta f(\eta|U_s = p) d\eta \\ &= (\gamma_1 - \gamma_0)' X + E[\eta|u_s = p] \end{aligned}$$

What is $E[\eta|u_s = p]$? The expected unobserved gain from treatment of those people who are on the treatment/no-treatment margin $P(Z) = p$.

How to Estimate an MTE

Easy

1. Estimate $P(Z) = Pr(T = 1|Z)$ nonparametrically (include exogenous part of X in Z).
2. Nonparametric regression of Y on X and $P(Z)$ (polynomials?)
3. Differentiate w.r.t. $P(Z)$
4. plot it for all values of $P(Z) = p$.

So long as $P(Z)$ covers $(0, 1)$ then we can trace out the full distribution of $\Delta^{MTE}(p)$.

Everything is an MTE

Calculate the outcome given (X, Z) (actually X and $P(Z) = p$).

- ATE : This one is obvious. We treat everyone!

$$\int_{-\infty}^{\infty} \Delta^{MTE}(p) = (\gamma_1 - \gamma_0)'X + \underbrace{\int_{-\infty}^{\infty} E(\eta|u_s)d\,u_s}_{0}$$

- LATE: Fix an X and $P(Z)$ varies from $b(X)$ to $a(X)$ and we integrated over the area between (compliers).

$$LATE(X) = \int_{-\infty}^{\infty} \Delta^{MTE}(p) = (\gamma_1 - \gamma_0)'X + \frac{1}{a(X) - b(X)} \int_{b(X)}^{a(X)} E(\eta|u_s)d\,u_s$$

- ATT

$$TT(X) = \int_{-\infty}^{\infty} \Delta^{MTE}(p) \frac{Pr(P(Z|X) > p)}{E[P(Z|X)]} d\,p$$

- Weights for IV and OLS are a bit more complicated. See the Heckman and Vytlacil paper(s).

Carneiro, Heckman and Vytlacil (AER 2010)

- Estimate returns to college (including heterogeneity of returns).
- NLSY 1979
- $Y = \log(wage)$
- Covariates X: Experience (years), Ability (AFQT Score), Mother's Education, Cohort Dummies, State Unemployment, MSA level average wage.
- Instruments Z: College in MSA at age 14, average earnings in MSA at 17 (opportunity cost), avg unemployment rate in state.

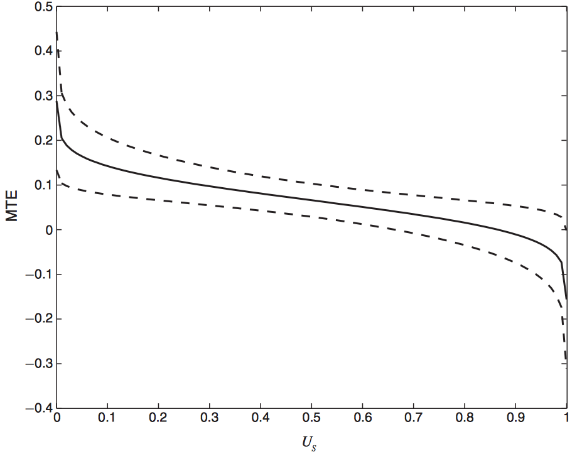


FIGURE 1. MTE ESTIMATED FROM A NORMAL SELECTION MODEL

Notes: To estimate the function plotted here, we estimate a parametric normal selection model by maximum likelihood. The figure is computed using the following formula:

TABLE 4— TEST OF LINEARITY OF $E(Y|\mathbf{X}, P = p)$ USING POLYNOMIALS IN P; AND TEST OF EQUALITY OF LATEs OVER DIFFERENT INTERVALS (H_0: $LATE^j\,(U_S^{Lj}, U_S^{Hj}) - LATE^{j+1}\,(U_S^{Lj+1}, U_S^{Hj+1}) = 0$)

Panel A. Test of linearity of $E(Y|\mathbf{X}, P = p)$ using models with different orders of polynomials in P[a]

Degree of polynomial for model	2	3	4	5
p-value of joint test of nonlinear terms	0.035	0.049	0.086	0.122
Adjusted critical value		0.057		
Outcome of test		Reject		

Panel B. Test of equality of LATEs (H_0: $LATE^j\,(U_S^{Lj}, U_S^{Hj}) - LATE^{j+1}\,(U_S^{Lj+1}, U_S^{Hj+1}) = 0$)[b]

Ranges of U_S for LATE^j	(0, 0.04)	(0.08, 0.12)	(0.16, 0.20)	(0.24, 0.28)	(0.32, 0.36)	(0.40, 0.44)
Ranges of U_S for LATE^{j+1}	(0.08, 0.12)	(0.16, 0.20)	(0.24, 0.28)	(0.32, 0.36)	(0.40, 0.44)	(0.48, 0.52)
Difference in LATEs	0.0689	0.0629	0.0577	0.0531	0.0492	0.0459
p-value	0.0240	0.0280	0.0280	0.0320	0.0320	0.0520
Ranges of U_S for LATE^j	(0.48, 0.52)	(0.56, 0.60)	(0.64, 0.68)	(0.72, 0.76)	(0.80, 0.84)	(0.88, 0.92)
Ranges of U_S for LATE^{j+1}	(0.56, 0.60)	(0.64, 0.68)	(0.72, 0.76)	(0.80, 0.84)	(0.88, 0.92)	(0.96, 1)
Difference in LATEs	0.0431	0.0408	0.0385	0.0364	0.0339	0.0311
p-value	0.0520	0.0760	0.0960	0.1320	0.1800	0.2400
Joint p-value			0.0520			

TABLE 5—RETURNS TO A YEAR OF COLLEGE

Model		Normal	Semiparametric
$ATE = E(\beta)$		0.0670 (0.0378)	Not identified
$TT = E(\beta \mid S = 1)$		0.1433 (0.0346)	Not identified
$TUT = E(\beta \mid S = 0)$		−0.0066 (0.0707)	Not identified
MPRTE			
Policy perturbation	Metric		
$Z_\alpha^k = Z^k + \alpha$	$\lvert \mathbf{Z}\gamma - V \rvert < e$	0.0662 (0.0373)	0.0802 (0.0424)
$P_\alpha = P + \alpha$	$\lvert P - U \rvert < e$	0.0637 (0.0379)	0.0865 (0.0455)
$P_\alpha = (1 + \alpha)P$	$\lvert \frac{P}{U} - 1 \rvert < e$	0.0363 (0.0569)	0.0148 (0.0589)
Linear IV (Using $P(\mathbf{Z})$ as the instrument)		0.0951 (0.0386)	
OLS		0.0836 (0.0068)	

Notes: This table presents estimates of various returns to college, for the semiparametric and the normal selection models: average treatment effect (ATE), treatment on the treated (TT), treatment on the untreated (TUT), and different versions of the marginal policy relevant treat-

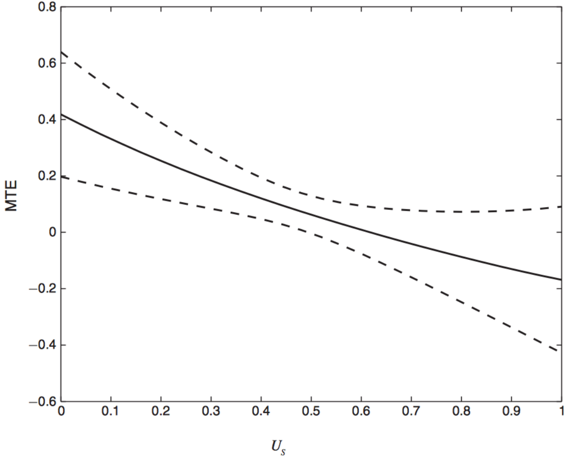


FIGURE 4. $E(Y_1 - Y_0 | \mathbf{X}, U_S)$ WITH 90 PERCENT CONFIDENCE INTERVAL—
LOCALLY QUADRATIC REGRESSION ESTIMATES

Notes: To estimate the function plotted here, we first use a partially linear regression of log wages on polynomials

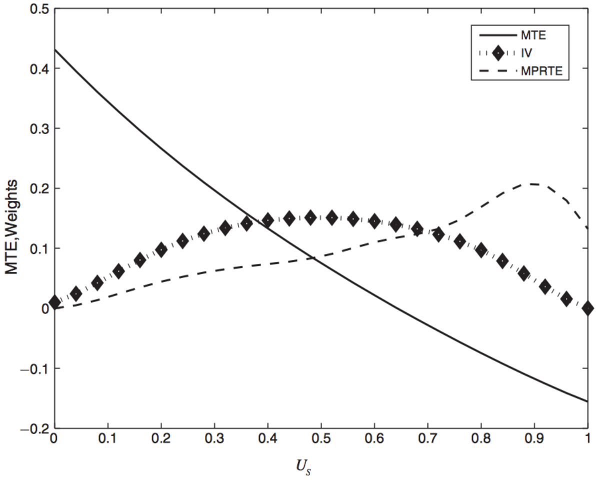


FIGURE 6. WEIGHTS FOR IV AND MPRTE

Note: The scale of the *y*-axis is the scale of the MTE, not the scale of the weights, which are scaled to fit the picture.

Diversion Example

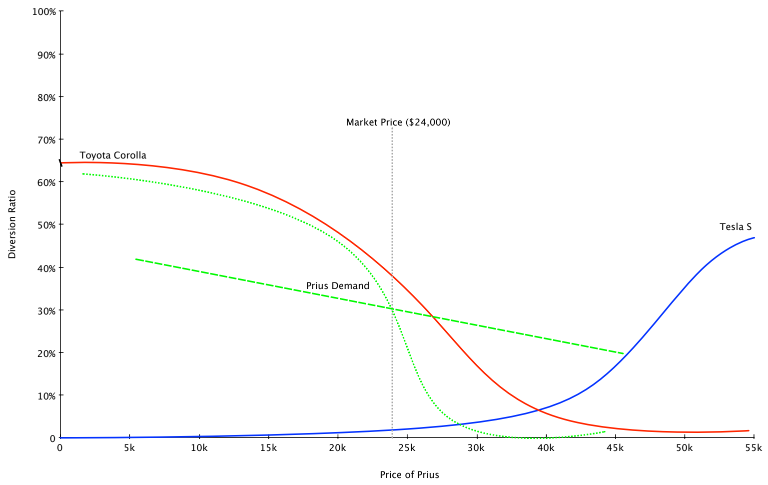
I have done some work trying to bring these methods into merger analysis.

- Key quantity: Diversion Ratio as I raise my price, how much do people switch to a particular competitor's product

$$D_{jk}(p_j, p_{-j}) = \left| \frac{\partial q_k}{\partial p_j}(p_j, p_{-j}) / \frac{\partial q_j}{\partial p_j}(p_j, p_{-j}) \right|$$

- We hold p_{-j} fixed and trace out $D_{jk}(p_j)$.
- The treatment is leaving good j.
- The Y_i is increased sales of good k.
- The Z_i is the price of good j.
- The key is that all changes in sales of k come through people leaving good j (no direct effects).

Diversion for Prius (FAKE!)



Diversion Example

$$\widehat{D_{jk}^{LATE}} = \frac{1}{\Delta q_j} \int_{p_j^0}^{p_j^0+\Delta p_j} \underbrace{\frac{\partial q_k(p_j, p_{-j}^0)}{\partial q_j}}_{\equiv D_{jk}(p_j, p_{-j}^0)} \left| \frac{\partial q_j(p_j, p_{-j}^0)}{\partial p_j} \right| dp_j$$

- $D_{jk}(p_j, p_{-j}^0)$ is the MTE.
- Weights $w(p_j) = \frac{1}{\Delta q_j} \frac{\partial q_j(p_j, p_{-j}^0)}{\partial p_j}$ correspond to the lost sales of j at a particular p_j as a fraction of all lost sales.
- When is $LATE \approx ATE$?
    - Demand for Prius is steep: everyone leaves right away
    - $D_{j,k}(p_j)$ is relatively flat.
    - We might want to think about raising the price to choke price (or eliminating the product from the consumers choice set) same as treating everyone!