Applied Microeconometrics Problem Set 3

Yixin Sun

November 30, 2020

Problem 1

(a)

$$\begin{aligned}
E\left[Y_1 - Y_0 \mid D=0\right] &= E\left[Y_1 - Y_0 \mid p(Z) < U\right] \\
&= E\left[E\left[Y_1 - Y_0 \mid p(Z) < U, U=u\right] \mid p(z) < U\right] \\
&= E\left[E\left[Y_1 - Y_0 \mid U=u\right] \mid p(z) < U\right] \\
&= \frac{E[m(u)\mathbb{1}\{p(z) < U\}]}{P(p(Z) < U)} \\
&= E\left[E\left[m(u)\frac{\mathbb{1}\{p(z) < U\}}{P(D=0)} \mid U=u\right]\right] \\
&= E\left[m(u)\frac{E[\mathbb{1}\{p(z) < U\} \mid U=u]}{P(D=0)}\right] \\
&= \int_0^1 m(u)\frac{P(p(Z) < u)}{P(D=0)}du
\end{aligned}$$

Where the first equality uses the fact that $D = 1[\nu(Z) \geq U]$, the second equality uses LIE, the third equality uses the fact that $Z \perp (Y_1, Y_0, U)$, the fifth equality uses the fact that $P(p(Z) < u) = P(D=0)$ and LIE, and the sixth equality uses the definition of expectation. We see that those with higher values of u have a higher weight, since they are the least likely to take treatment.

(b) For any two values of the instrument $Z = z_1$ and $Z = z_0$, we can write

(1) Using instrument exogeneity to get

$$\begin{aligned}
E[Y|Z=z_1] &= E\left[DY_1 + (1-D)Y_0 | Z=z_1\right] \qquad (\because E[U_1 - U_0|Z=z_1] = E[U_1 - U_0]) \\
&= E\left[D(z_1)Y_1 + \left(\mathbb{1} - D(z_1)\right)Y_0 | Z=z_1\right] \\
&= E\left[D(z_1)Y_1 + \left(\mathbb{1} - D(z_1)\right)Y_0\right] \\
E[Y|Z=z_0] &= E\left[D(z_0)Y_1 + \left(\mathbb{1} - D(z_0)\right)Y_0\right]
\end{aligned}$$

(2) Combining these in the numerator, we have

$$\begin{aligned}
E[Y|Z=z_1] - E[Y|Z=z_0] &= E\left[D(z_1)Y_1 + \left(\mathbb{1} - D(z_1)\right)Y_0\right] - E\left[D(z_0)Y_1 + \left(\mathbb{1} - D(z_0)\right)Y_0\right] \\
&= E\left[\left(Y_1 - Y_0\right)\left(D(z_1) - D(z_0)\right)\right]
\end{aligned}$$

(3) Then using the law of total probability, followed by monotonicity, we can write

$$E[Y|Z=z_1] - E[Y|Z=z_0] = E\left[Y_1 - Y_0 | D(z_1) > D(z_0)\right] P\left(D(z_1) > D(z_0)\right)$$

(4) similarly for the denominator we have

$$E[D|Z=z_1]-E[D|Z=z_0]=E\left[D(z_1)-D(z_0)\right]=P\left(D(z_1)>D(z_0)\right)$$

(5) Combining, we get

$$\begin{aligned}\frac{E[Y|Z=z_1]-E[Y|Z=z_0]}{E[D|Z=z_1]-E[D|Z=z_0]} &= \frac{E\left[Y_1-Y_0|D(z_1)>D(z_0)\right]P\left(D(z_1)>D(z_0)\right)}{P\left(D(z_1)>D(z_0)\right)} \\ &= E\left[Y_1-Y_0|D(z_1)>D(z_0)\right] \\ &= E\left[Y_1-Y_0 \mid p\left(z_0\right)\leq u\leq p\left(z_1\right)\right] \\ &= \int_0^1 m(u)\left(\frac{\mathbb{1}\left[p\left(z_0\right)<u\leq p\left(z_1\right)\right]}{p\left(z_1\right)-p\left(z_0\right)}\right)du\end{aligned}$$

(c) Let's start with the numerator:

$$\begin{aligned}E\left[Y^*\right]-E[Y] &= E\left[E\left[\left(D^*-D\right)\left(Y(1)-Y(0)\right)\mid U=u\right]\right] \\ &= \int_0^1 E\left[\left(D^*-D\right)\left(Y(1)-Y(0)\right)\mid U=u\right]\mathrm{d}u \\ &= \int_0^1 E\left[\left(\mathbb{1}\left[u\leq p^*\left(Z^*\right)\right]-\mathbb{1}[u\leq p(Z)]\right)\left(Y(1)-Y(0)\right)\mid U=u\right]\mathrm{d}u \\ &= \int_0^1 \left(E\left[\mathbb{1}\left[u\leq p^*\left(Z^*\right)\right]\mid U=u\right]-E[\mathbb{1}[u\leq p(Z)]\mid U=u]\right)m(u)\mathrm{d}u \\ &= \int_0^1 \left(\left(\mathbb{1}-F_{P^*}^-(u)\right)-\left(\mathbb{1}-F_P^-(u)\right)\right)m(u)\mathrm{d}u \\ &= \int_0^1 \left(F_P^-(u)-F_{P^*}^-(u)\right)m(u)\mathrm{d}u\end{aligned}$$

The denominator can be written as:

$$\begin{aligned}E\left[D^*\right]-E[D] &= E\left[\mathbb{1}\left[U\leq p^*\left(Z^*\right)\right]\right]-E[\mathbb{1}[U\leq p(Z)]] \\ &= E\left[p^*\left(Z^*\right)\right]-E[p(Z)]\end{aligned}$$

Putting this together, we get our desired result.

(d) First break up the object of interest

$$E[s(D,Z)Y]=E\left[s(1,Z)DY(1)\right]+E\left[s(0,Z)(1-D)Y(0)\right]$$

And look at the first component

$$\begin{aligned}E\left[s(1,Z)DY(1)\right] &= E\left[E\left[s(1,Z)\mathbb{1}[U\leq p(Z)]Y(1)\mid U=u\right]\right] \\ &= E\left[E\left[s(1,Z)\mathbb{1}\left\{u\leq P\right\}\mid U=u\right]m(1|u)\right] \\ &= \int_0^1 m(1\mid u)\times E[s(1,Z)\mid P\geq u]\left(1-F_P^-(u)\right)du\end{aligned}$$

A similar argument follows for the second component, and we have our desired result.

(e) Following Mogstad, Torgovitsky, and Walters (2018), if we have two instruments, Z_1 and Z_2, and $\pi_{1c}=P(Z_1 \text{ compliers})>P(Z_2 \text{ compliers})=\pi_{2c}$, then the sign on the weight for Z_2 compliers is determined by

$$\operatorname{sgn}\left(\omega_{2c}\right)=\mathbb{1}\left[\pi_{2c}>0\right]\times\operatorname{sgn}\left(P\left[D_i=1\mid Z_{i,2}=1\right]-P\left[D_i=1\mid Z_{i,2}=0\right]\right)$$

and we can have negative weights when $P[D_i=1|Z_{i,2}=1]>P[D_i=1|Z_{i,2}=0]$.

Problem 2

(a) The MTE function is

$$MTE(u) = (\alpha_1 - \alpha_0) + (\beta_1 - \beta_0)\, u$$

which is constant if and only if $\beta_1 = \beta_0$.

Turning to $E[Y \mid D = 1, Z = 1], E[Y \mid D = 1, Z = 0], E[Y \mid D = 0, Z = 1]$, and $E[Y \mid D = 0, Z = 0]$, we first plug in MTR functions:

$$\begin{aligned}
E[Y|D = 1, Z = 1 &= E[Y(1)|p(1) > u] \\
&= \frac{1}{p(1)} \int_0^{p(1)} (\alpha_1 + \beta_1 u) du \\
&= \alpha_1 + \frac{1}{2}\beta_1 p(1) \\
E[Y \mid D = 1, Z = 0] &= \alpha_1 + \frac{1}{2}\beta_1 p(0) \\
E[Y \mid D = 0, Z = 1] &= E[Y(0)|p(1) < u] \\
&= \frac{1}{1 - p(1)} \int_{p(1)}^{1} (\alpha_0 + \beta_0 u) du \\
&= \frac{1}{1 - p(1)} \left[\alpha_0(1 - p(1)) + \frac{1}{2}\beta_0(1 + p(1))(1 - p(1))\right] \\
&= \alpha_0 + \frac{1}{2}\beta_0(1 + p(1)) \\
E[Y \mid D = 0, Z = 0] &= \alpha_0 + \frac{1}{2}\beta_0(1 + p(0))
\end{aligned}$$

Using these, we can compute

$$\begin{aligned}
E[Y \mid D = 1, Z = 1] - E[Y \mid D = 1, Z = 0] &= -\frac{1}{2}\beta_1(p(1) - p(0)) \\
EY \mid D = 0, Z = 1] - E[Y \mid D = 0, Z = 0] &= -\frac{1}{2}\beta_0(p(1) - p(0))
\end{aligned}$$

and we see that these two are equal if and only if $\beta_1 = \beta_0$.

(b) The estimate for LATE under this specification is

$$\begin{aligned}
E\left[Y(1) - Y(0) \mid D\left(z_1\right) > D\left(z_0\right)\right] &= \int_0^1 m(u) \frac{\mathbb{1}[p(0) \leq u \leq p(1)]}{p(1) - p(0)} \mathrm{d}u \\
&= \int_0^1 \left((\alpha_1 - \alpha_0) + (\beta_1 - \beta_0)\, u\right) \frac{\mathbb{1}[p(0) \leq u \leq p(1)]}{p(1) - p(0)} \mathrm{d}u \\
&= \alpha_1 - \alpha_0 + (\beta_1 - \beta_0)\left(\frac{p(0) + p(1)}{2}\right)
\end{aligned}$$

Now turning to the Wald, let's first write out the numerator:

$$\begin{aligned}
E[Y \mid Z=1]-E[Y \mid Z=0] &= p(1)E[Y \mid Z=1, D=1]+(1-p(1))E[Y \mid Z=1, D=0] \\
&\quad -(p(1)E[Y \mid Z=0, D=1]+(1-p(1))E[Y \mid Z=0, D=0]) \\
&= \left((p(1)-p(0))(\alpha_1-\alpha_0)+\frac{1}{2}\left(p(1)^2-p(0)^2\right)\right)(\beta_1-\beta_0) \\
&= ((p(1)-p(0))\left[(\alpha_1-\alpha_0)+\frac{1}{2}(p(0)+p(1))\right](\beta_1-\beta_0)
\end{aligned}$$

where we plugged in our results from part a. Thus we get that the Wald estimand is

$$\frac{E[Y \mid Z=1]-E[Y \mid Z=0]}{E[D \mid Z=1]-E[D \mid Z=0]}=\alpha_1-\alpha_0+(\beta_1-\beta_0)\left(\frac{p(0)+p(1)}{2}\right)$$

which is the same as our LATE. Tihs proof generalizes to other MTR functions, and does not depend on the linearity of the given functional form.

(c) With a binary instrument and no covariates, the LATE matches the Wald estimand because it is simply a weighted average of MTEs, where the weights are constructed from the propensity scores (see 1b).

Problem 3

(a) To the left of the cutoff, we only observe individuals with $M=0$. Since we are given that $f_{R|M}(r \mid 0)$ is continuous at $r=c$, we know that

$$\lim_{r \uparrow c} f_R(r)=(1-\pi) f_R(c)$$
$$\Rightarrow \pi=1-\frac{\lim_{r \uparrow c} f_{R|M}(r \mid 0)}{f_R(c)}$$

(b) Our target of interest is bounds on

$$\begin{aligned}
\delta &\equiv E[Y(1)-Y(0) \mid R=c, M=0] \\
&= E[Y(1) \mid R=c, M=0]-E[Y(0) \mid R=c, M=0]
\end{aligned}$$

We know $E[Y(0) \mid R=r, M=0]$ is continuous at $r=c$, so can identify the second term from above:

$$E[Y(0) \mid R=c, M=0]=\lim_{r \uparrow c} E[Y \mid R=r]$$

To get bounds on the first term, we can adapt the results from problem set 1, question 2 to this setting. G is the empirical cdf of $Y|R=c$. π is defined in the same way as in 2c. Then we can directly apply 2c to get.

$$E\left[Y \mid R=c, Y \leq G^{-1}(1-\pi)\right] \leq E[Y \mid R=c, M=0] \leq E\left[Y \mid R=c, Y \geq G^{-1}(\pi)\right]$$

Putting this together, our bounds are

$$E\left[Y \mid R=c, Y \leq G^{-1}(1-\pi)\right]-\lim_{r \uparrow c} E[Y \mid R=r] \leq \delta \leq E\left[Y \mid R=c, Y \geq G^{-1}(\pi)\right]-\lim_{r \uparrow c} E[Y \mid R=r]$$

(c) Because there is selection on unobservables, we no longer have continuity from the left in our running variable, differing from the usual sharp regression discontinuity framework.

Problem 4

First note that

$$\begin{aligned}
E[Y(1) - Y(0) \mid R = r] &= P(C = c_h|R = r)E[Y(1) - Y(0)|R = r, C = c_h] \\
&+ P(C = c_l|R = r)E[Y(1) - Y(0)|R = r, C = c_l] \\
&= \pi_h(g_1(r) + \alpha_1 - g_0(r) - \alpha_0) + \pi_r(g_1(r) - g_0(r)) \\
&= g_1(r) - g_0(r) + \pi_h(\alpha_1 - \alpha_0)
\end{aligned}$$

We observe π_h in the data, so we are left to identify $g_1(r), g_0(r), \alpha_1$, and α_0.

There are numerous categories r can fall into: $r < c_l$, $r = c_l$, $r \in (c_l, c_h)$, and $r \geq c_h$. We can use information from these different intervals to point identify what we need.

1. For $r < c_l$, we observe

$$E\left[Y(0) \mid R = r, C = c_l\right] = g_0(r)$$

2. for $r = c_l$, we observe

$$E\left[Y(1) \mid R = c_l\right] = g_1(r)$$

3. For $r \in (c_l, c_h)$, we observe

$$E[Y(0)|R = r, C = c_h] = g_0(r) + \alpha_0$$

4. For $r = c_h$, we observe

$$E\left[Y(1) \mid R = c_h\right] = g_1(r) + \alpha_1$$

Combining observed quantities, we have

$$\begin{aligned}
\alpha_1 &= E[Y(1)|R = c_h] - E[Y(1)|R = c_l] \\
\alpha_0 &= E[Y(0)|R = r, C = C_h] - E[Y(0)|R = r, C = c_l]
\end{aligned}$$

and everything has been identified.

In the usual sharp RD designs, there is only one cutoff and thus you are either treated or untreated. Here, when $r \in (c_l, c_h)$, these observations are treated for one cutoff, but untreated for the higher cutoff. In order to pool information together from two different cutoff points, this problem assumes that the functional form, $g_d(r)$, is the same across the two cutoffs, and the only differences is in the intercept.

Problem 5

Parameterizing the MTE

We can run simple OLS regressions to recover $E[Y \mid D=1, X=x, Z=z]$, and $E[Y \mid D=0, X=x, Z=z]$. We can then recover estimates for the $m(d|u,x)$ specifications in the following way:

$$\begin{aligned}
E[Y \mid D=1, X=x, Z=z] &= \frac{1}{p}\int_0^p E[Y \mid D=d, X=x, Z=z, U=u]\mathrm{d}u \\
&= \frac{1}{p}\int_0^p E[Y(d) \mid X=x, U=u]\mathrm{d}u \\
&= \frac{1}{p}\int_0^p m(1 \mid u, x)\mathrm{d}u \\
E[Y \mid D=0, X=x, Z=z] &= \frac{1}{1-p}\int_p^1 m(0 \mid u, x)\mathrm{d}u
\end{aligned}$$

I apply these derivations to all our specifications, and illustrate how we can recover the parameters of interest from the OLS regressions.

Spec 1: $m(d \mid u, x) = \alpha_d + \beta_d u + \gamma_d' x$

$$\begin{aligned}
E[Y \mid D=1, P=p, X=x] &= \frac{1}{p}\left[\alpha_1 p + \frac{\beta_1}{2}p^2 + \gamma_1' x p\right] \\
&= \alpha_1 + \frac{\beta_1}{2}p + \gamma_1' x \\
E[Y \mid D=0, P=p, X=x] &= \frac{1}{1-p}\left(\alpha_0 + \frac{1}{2}\beta_0 + x'\gamma_0\right) - \frac{1}{1-p}\left(\alpha_0 p + \frac{1}{2}\beta_0 p^2 + x'\gamma_0 p\right) \\
&= \alpha_0 + \frac{1}{2}\beta_0 + \frac{1}{2}\beta_0 p + x'\gamma_0
\end{aligned}$$

$$\begin{aligned}
\alpha_1 &= \hat{\alpha}_1 \\
\beta_1 &= 2\hat{\beta}_1 \\
\alpha_0 &= \hat{\alpha}_0 - \hat{\beta}_0 \\
\beta_0 &= 2\hat{\beta}_0
\end{aligned}$$

Spec 2: $m(d \mid u, x) = \alpha_d + \beta_d u + \gamma' x$

The derivation is the same as for Spec 1, except the coefficients on x are the same. I implement this by running one regression to estimate all the parameters.

Spec 3: $m(d \mid u, x) = \alpha_d + \beta_d u + \gamma_d' x + \delta_d' x u$

$$\begin{aligned}
E[Y|D=1, P=p, X=x] &= \alpha_1 + \frac{1}{2}\beta_1 p + \gamma_1' x + \frac{1}{2}\delta_1' x p \\
E[Y|D=0, P=p, X=x] &= \frac{1}{1-p}\left(\alpha_0 + \frac{1}{2}\beta_0 + x'\gamma_0 + \delta_0' x\right) \\
&\quad - \frac{1}{1-p}\left(\alpha_0 p + \frac{1}{2}\beta_0 p^2 + x'\gamma_0 + \frac{1}{2}\delta_0' x p^2\right) \\
&= \alpha_0 + \frac{1}{2}\beta_0 + \frac{1}{2}\beta_0 p + x'\gamma_0 + \frac{1}{2}x'\delta_0 + \frac{1}{2}p x'\delta_0
\end{aligned}$$

$$\alpha_1 = \hat{\alpha}_1$$
$$\beta_1 = 2\hat{\beta}_1$$
$$\gamma_1 = \hat{\gamma}_1$$
$$\delta_1 = 2\hat{\delta}_1$$
$$\alpha_0 = \hat{\alpha}_0 - \hat{\beta}_0$$
$$\beta_0 = 2\hat{\beta}_0$$
$$\gamma_0 = \hat{\gamma}_0 - \hat{\delta}_0$$
$$\delta_0 = 2\hat{\delta}_0$$

Spec 4: $m(d \mid u, x) = \alpha_d + \beta_{d1} u + \beta_{d2} u^2 + \gamma_d' x$

$$E[Y|D=1, P=p, X=x] = \alpha_1 + \frac{1}{2}\beta_{11} p + \frac{1}{3}\beta_{12} p^2 + x'\gamma_1$$
$$E[Y|D=0, P=p, X=x] = \left(\alpha_0 + \frac{1}{2}\beta_{01} + \frac{1}{3}\beta_{02}\right) + \left(\frac{1}{2}\beta_{01} + \frac{1}{3}\beta_{02}\right) p + \frac{1}{3}\beta_{02} p^2 + x'\gamma_0$$

$$\alpha_1 = \hat{\alpha}_1$$
$$\beta_{11} = 2\hat{\beta}_{11}$$
$$\beta_{12} = 3\hat{\beta}_{12}$$
$$\gamma_1 = \hat{\gamma}_1$$
$$\alpha_0 = \hat{\alpha}_0 - \hat{\beta}_{01}$$
$$\beta_{01} = 2(\hat{\beta}_{01} - \hat{\beta}_{02})$$
$$\beta_{02} = 3\hat{\beta}_{02}$$
$$\gamma_0 = \hat{\gamma}_0$$

Spec 5: $m(d \mid u, x) = \alpha_d + \beta_{d1} u + \beta_{d2} u^2 + \beta_{d3} u^3 + \gamma_d' x$

$$E[Y|D=1, P=p, X=x] = \alpha_1 + \frac{1}{2}\beta_{11} u + \frac{1}{3}\beta_{12} u^2 + \frac{1}{4}\beta_{13} u^3 + x'\gamma_1$$
$$\begin{aligned} E[Y|D=0, P=p, X=x] = &\left(\alpha_0 + \frac{1}{2}\beta_{01} + \frac{1}{3}\beta_{02} + \frac{1}{4}\beta_{03}\right) + \left(\frac{1}{2}\beta_{01} + \frac{1}{3}\beta_{02} + \frac{1}{4}\beta_{03}\right) u \\ &+ \left(\frac{1}{3}\beta_{02} + \frac{1}{4}\beta_{03}\right) u^2 + \frac{1}{4}\beta_{03} u^3 + x'\gamma_0 \end{aligned}$$

$$\alpha_1 = \hat{\alpha}_1$$
$$\beta_{11} = 2\hat{\beta}_{11}$$
$$\beta_{12} = 3\hat{\beta}_{12}$$
$$\beta_{13} = 4\hat{\beta}_{13}$$
$$\gamma_1 = \hat{\gamma}_1$$
$$\alpha_0 = \hat{\alpha}_0 - \hat{\beta}_{01}$$
$$\beta_{01} = 2(\hat{\beta}_{01} - \hat{\beta}_{02})$$
$$\beta_{02} = 3(\hat{\beta}_{02} - \hat{\beta}_{03})$$
$$\beta_{03} = 4\hat{\beta}_{03}$$
$$\gamma_0 = \hat{\gamma}_0 - \hat{\delta}_0$$

Table 1: Treatment Effects

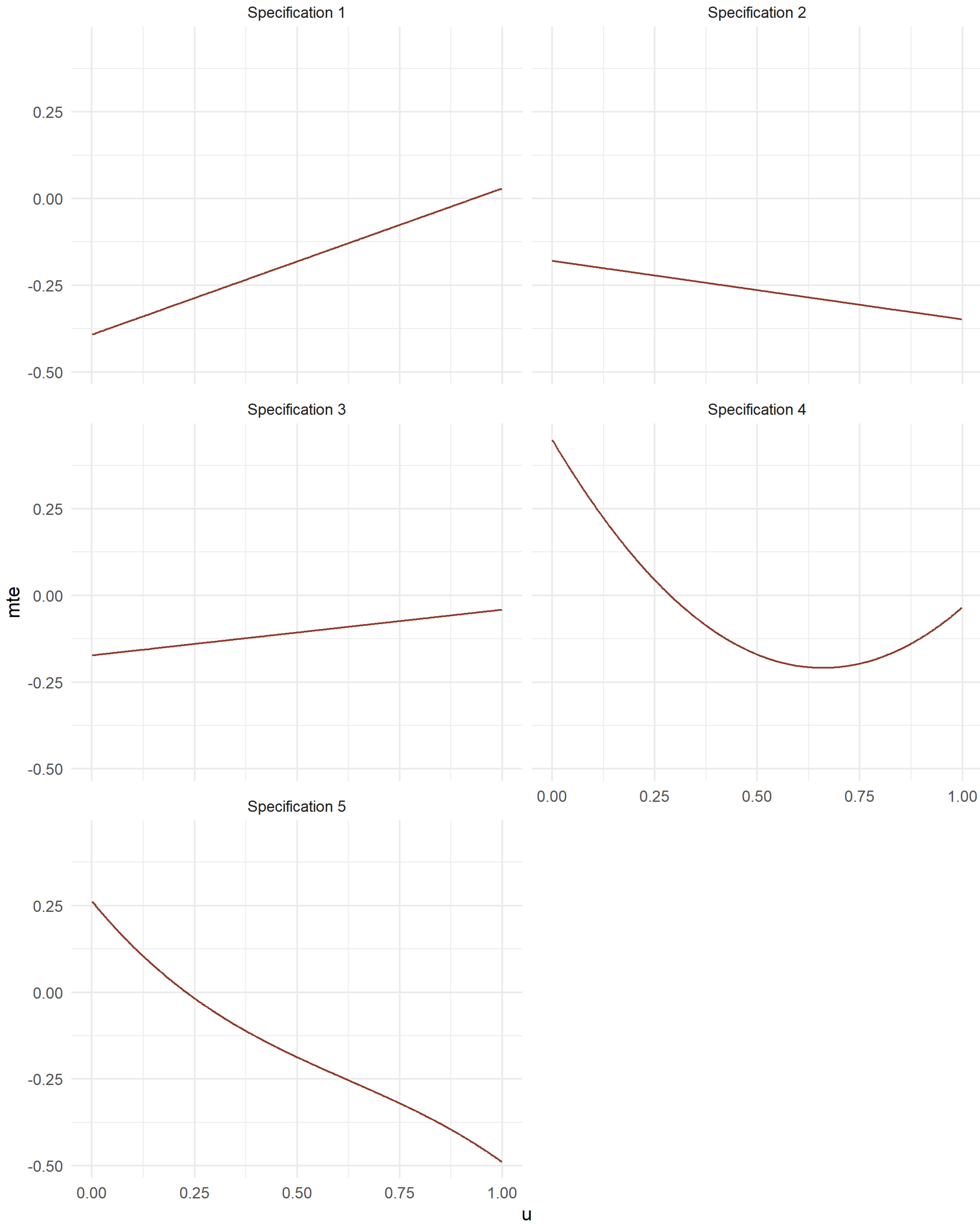
Estimator	Specification 1	Specification 2	Specification 3	Specification 4	Specification 5
Same Sex Instrument					
ate	-0.189	-0.111	-0.187	-0.187	-0.187
att	-0.196	-0.035	-0.189	-0.189	-0.188
atu	-0.184	-0.162	-0.185	-0.186	-0.185
late	-0.196	-0.101	-0.188	-0.189	-0.188
Twins Instrument					
ate	-0.188	-0.113	-0.188	-0.187	-0.185
att	-0.190	-0.030	-0.190	-0.186	-0.182
atu	-0.186	-0.168	-0.187	-0.188	-0.188
late	-0.190	-0.100	-0.188	-0.185	-0.181
Both Instruments					
ate	-0.189	-0.111	-0.187	-0.188	-0.187
att	-0.196	-0.035	-0.189	-0.189	-0.189
atu	-0.184	-0.162	-0.185	-0.187	-0.186
late	-0.197	-0.101	-0.188	-0.189	-0.188

The TSLS estimators shown here are very similar to those computed in Table 11 of the paper. We see that the magnitude of the treatment effect for different estimators is bigger when estimated using MTEs rather than the TSLS, suggesting that the effect is larger when extrapolated beyond the effect for compliers.

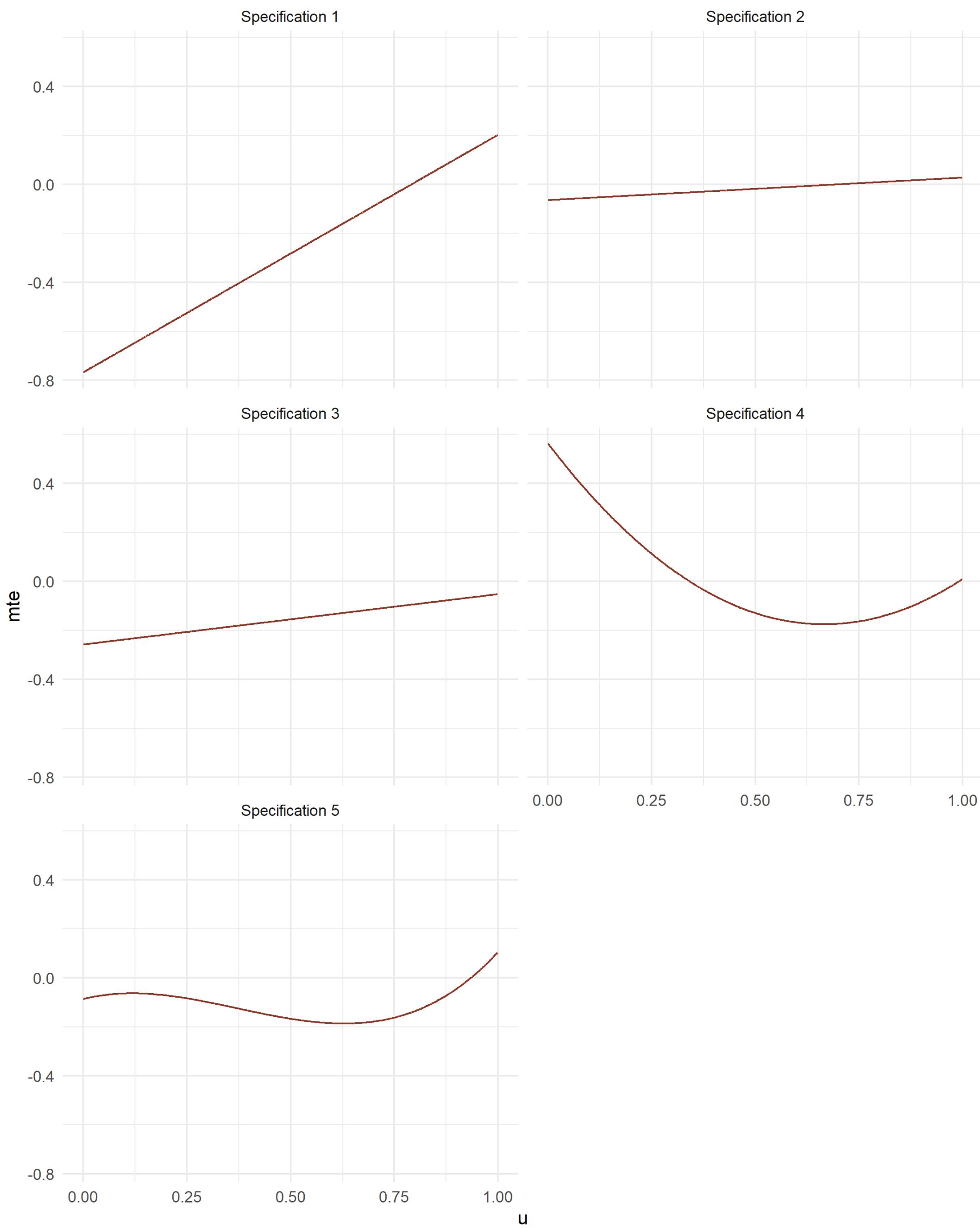
Table 2: TSLS Estimators

Instrument	Estimate
Same Sex	-0.118
Twins	-0.083
Both	-0.113

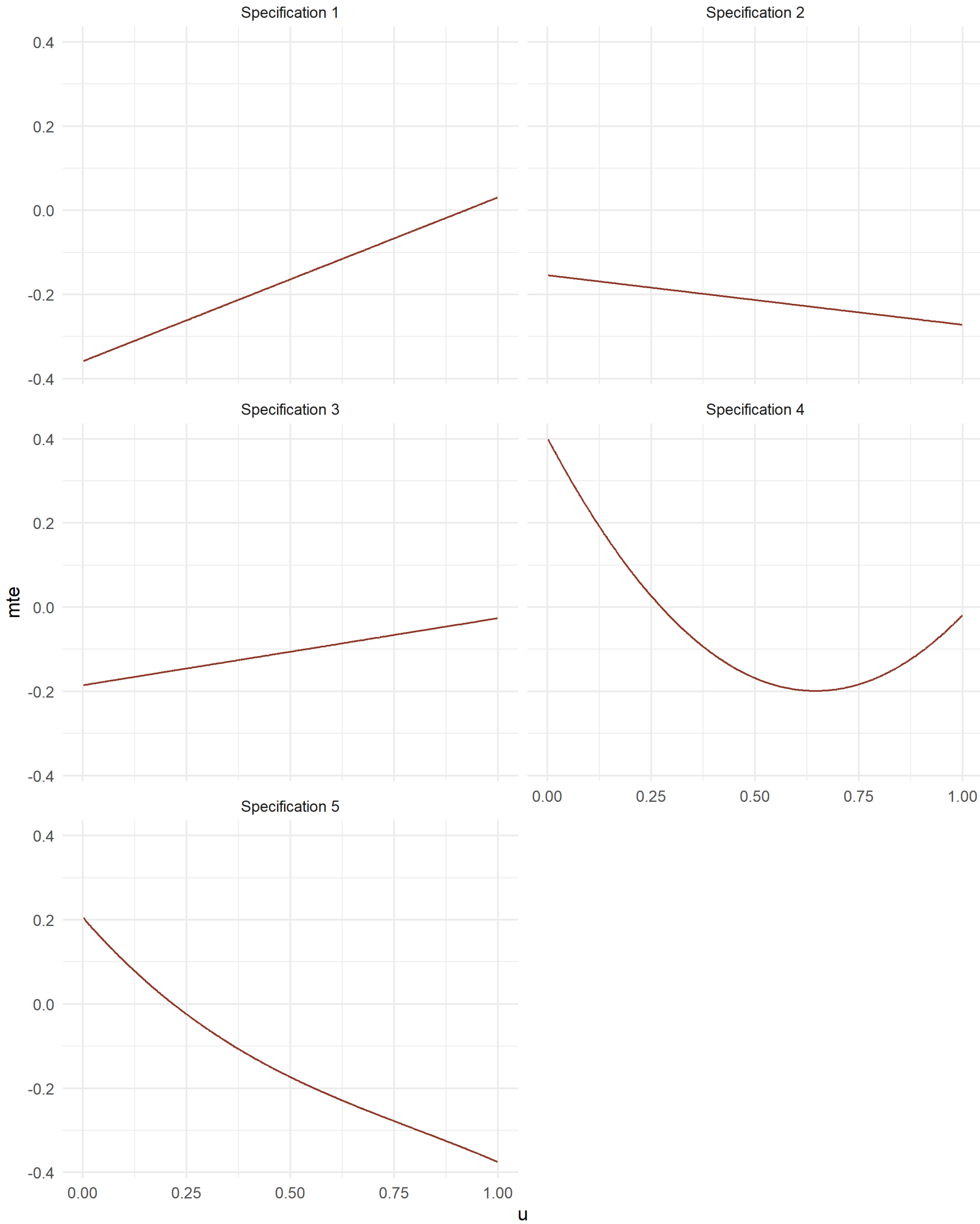
Part (a) Same Sex Instrument



Part (b) Twins Instrument



Part (c) Both Instruments



Problem 6

Maximization Problem:

$$\bar{\beta}^\star = \max_{\theta \in \Theta} \sum_{d \in \{0,1\}} \sum_{k=0}^{K_d} \gamma_{dk}^\star \theta_{dk} \quad \text{subject to} \quad \sum_{d \in \{0,1\}} \sum_{k=0}^{K_d} \gamma_{sdk} \theta_{dk} = \beta_s \text{ for all } s \in \mathcal{S}$$

$$\gamma_{dk}^\star \equiv E\left[\int_0^1 b_{dk}(u, X)\omega_d^\star(u, X, Z)\mathrm{d}u\right]$$

$$\gamma_{sdk} \equiv E\left[\int_0^1 b_{dk}(u, X)\omega_{ds}(u, X, Z)\mathrm{d}u\right]$$

The ATT weights are:

$$\omega_1^\star(u, x, z) = \frac{\mathbb{1}[u \leq p]}{P[D = 1]}$$

$$\omega_0^\star(u, x, z) = -\frac{\mathbb{1}[u \leq p]}{P[D = 1]}$$

Given

$$p(1) = 0.12, \quad p(2) = 0.29, \quad p(3) = 0.48, \quad \text{and} \quad p(4) = 0.78$$

$$P[Z = z] = 1/4$$

$$m_0(u) = 0.9 - 1.1u + 0.3u^2$$

$$m_1(u) = 0.35 - 0.3u - 0.05u^2$$

IV Slope: I need to calculate both sides of the constraint. I start with the right hand side,

$$\beta_s = E\left[\int_0^1 m_0(u, X)\omega_{0s}(u, X, Z)\mathrm{d}u\right] + E\left[\int_0^1 m_1(u, X)\omega_{1s}(u, X, Z)\mathrm{d}u\right]$$

$$\text{where } \omega_{0s}(u, x, z) \equiv s(0, x, z)\mathbb{1}[u > p]$$

$$\text{and } \omega_{1s}(u, x, z) \equiv s(1, x, z)\mathbb{1}[u \leq p]$$

Know:

$$E[Z] = \frac{1 + 2 + 3 + 4}{4} = 2.5$$

$$E[DZ] = \frac{1}{4}(1 \times .12 + 2 \times .29 + 3 \times .48 + 4 \times .78) = 1.315$$

$$E[D] = \frac{.12 + .29 + .48 + .78}{4} = .4175$$

Which gives us

$$s(d, x, z) = \frac{z - E[Z]}{\mathrm{Cov}(D, Z)}$$

$$= \frac{z - 2.5}{0.27125}$$

We can now integrate over our weights and MTR functions, which gives us the left hand side, $\beta_s = -0.275$

To calculate the right hand side, $\gamma_{sdk} \equiv E\left[\int_0^1 b_{dk}(u,X)\omega_{ds}(u,X,Z)\mathrm{d}u\right]$, we use the Bernstein polynomial

$$b_k^K(z) \equiv \left(\begin{array}{c} K \\ k \end{array} \right) u^k(1-u)^{K-k}$$

<u>TSLS Slope</u>
Now, our weights are

$$\begin{aligned} s(d,z) &= \left(\Pi E\left[\widetilde{Z}\tilde{X}'\right]\right)^{-1}\Pi\widetilde{Z} \\ \Pi &\equiv E\left[\widetilde{X}\widetilde{Z}'\right]E\left[\widetilde{Z}\widetilde{Z}'\right]^{-1} \\ \widetilde{X} &\equiv [1,D]' \\ \widetilde{Z} &\equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

which gives us that $\beta_{tsls} = -0.2687$. Following the previous section, we construct the right hand side of the constraint using Bernstein polynomials weighted by this TSLS $s(d,z)$ function.

<u>Monotonicity Constraints</u>
Following Mogstad, Santos & Torgovitsky (2018), we can add linear constraints to our maximization problem in order to impose our monotonicity shape constraint. To implement, we add $K \times 2$ constraints,

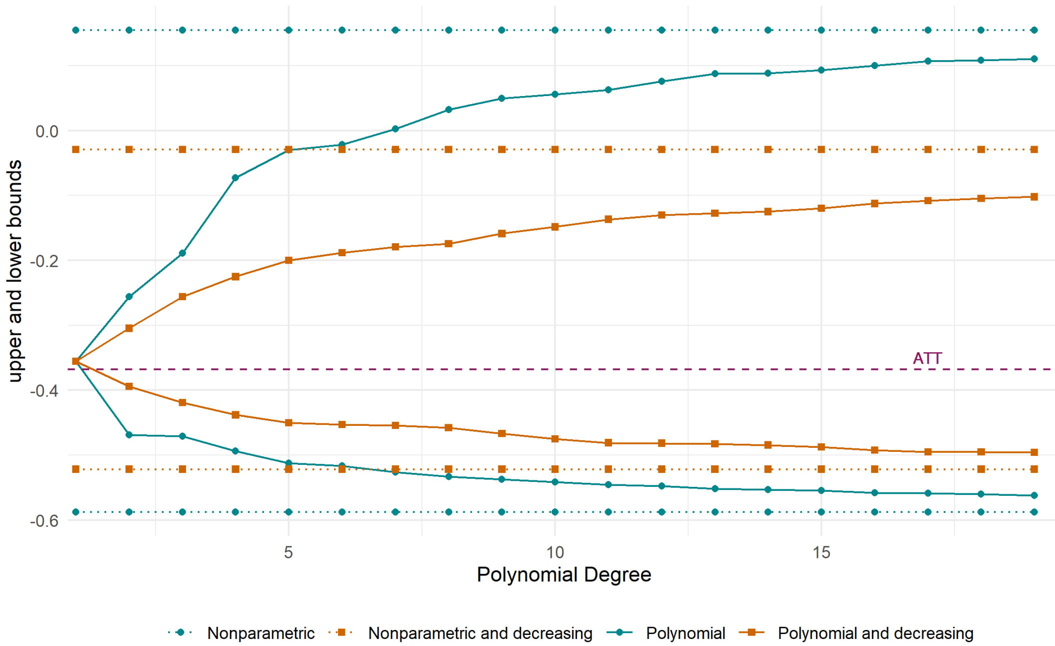
$$\theta_{d,k} - \theta_{d,k+1} < 0, \quad d \in \{0,1\}, \forall k = 0,...,K$$

<u>Nonparametric Bounds</u>
Again, following Mogstad, Santos & Torgovitsky (2018), we estimate the nonparametric bounds by using the following basis

$$b_j(u,x) \equiv \mathbb{1}\left[u \in \mathcal{U}_j\right] \quad \text{for } 1 \leq j \leq J$$

where $\mathcal{U}_j = \{[0,.12],[.12,.29],[.29,.48],[.48,.78],[.78,1]\}$

Figure 6 Replication



References

Mogstad, Magne, Andres Santos, and Alexander Torgovitsky. "Supplement to 'Using Instrumental Variables for Inference About Policy Relevant Treatment Parameters,'" n.d., 10.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R Walters. "The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables," n.d., 51.