Problem Set 1

Carlos Enrique Lezama Jacinto

Instituto Tecnológico Autónomo de México

The Model

Assume that the economic model generating the data for potential outcomes is of the form:

$$
\begin{aligned}
Y_1 &= \alpha + \varphi + U_1, \\
Y_0 &= \alpha + U_0,
\end{aligned}
$$

where U_1 and U_0 represent the unobservables in the potential outcome equations and φ represents the benefit associated with the treatment ($D = 1$). Individuals decide whether or not to receive the treatment ($D = 1$ or $D = 0$) based on a latent variable I:

$$
I = Z\gamma + V,
$$

where Z and V represent observables and unobservables, respectively. Thus, we can define a binary variable D indicating treatment status,

$$
D = \mathbb{1}\left[I \geq 0\right].
$$

Finally, we assume that the error terms in the model are not independent even conditioning on the observables, i.e. $U_1 \not\perp\!\!\!\perp U_0 \not\perp\!\!\!\perp V \mid Z$, but $(U_1, U_0, V) \perp\!\!\!\perp Z$.

1. Treatment Parameters

Let $\beta = Y_1 - Y_0 = \varphi + U_1 - U_0$ such that, in regression notation, $Y = \alpha + \beta D + \varepsilon$ where $\varepsilon = U_0$.

Average Treatment Effect.

$$
\begin{aligned}
ATE &= E\left[\beta\right] \\
&= E\left[\varphi + U_1 - U_0\right] \\
&= \varphi
\end{aligned}
$$

Treatment on the Treated.

$$
\begin{aligned}
TT &= E\left[\beta \mid D = 1\right] \\
&= E\left[\varphi + U_1 - U_0 \mid Z\gamma + V \geq 0\right] \\
&= \varphi + E\left[U_1 - U_0 \mid V \geq -Z\gamma\right]
\end{aligned}
$$

Treatment on the Untreated.

$$
\begin{aligned}
TUT &= E\left[\beta \mid D = 0\right] \\
&= E\left[\varphi + U_1 - U_0 \mid Z\gamma + V < 0\right] \\
&= \varphi + E\left[U_1 - U_0 \mid V < -Z\gamma\right]
\end{aligned}
$$

Marginal Treatment Effect.

$$
\begin{aligned}
MTE &= E\left[\beta \mid I = 0,\ V = v\right] \\
&= E\left[\varphi + U_1 - U_0 \mid I = 0,\ V = v\right] \\
&= \varphi + E\left[U_1 - U_0 \mid Z\gamma = -V,\ V = v\right]
\end{aligned}
$$

Instrumental Variables.

$$
\hat{\beta}_{\mathrm{IV}}(J(Z)) = \frac{\mathrm{Cov}(J(Z), Y)}{\mathrm{Cov}(J(Z), D)} \xrightarrow{p} \beta
$$

My solutions to the first problem set in Advanced Microeconometrics (ECO – 20513).

Ordinary Least Squares.

$$\hat{\beta}_{\text{OLS}} = \frac{\text{Cov}(Y,D)}{\text{Var}(D)} \implies \hat{\beta}_{\text{OLS}} = \left(D^T D\right)^{-1} D^T Y$$

Local Average Treatment Effect.

$$\begin{aligned} LATE &= E\left[\beta \mid D(z)=0, D(z')=1\right] \\ &= E\left[\beta \mid z\gamma < -V \leq z'\gamma\right] \\ &= \varphi + E\left[U_1 - U_0 \mid -z'\gamma \leq V < -z\gamma\right] \end{aligned}$$

2. Some closed form expressions

Suppose that the error terms in the model have the following structure:

$$\begin{aligned} U_1 &= \sigma_1 \epsilon, \\ U_0 &= \sigma_0 \epsilon, \\ V &= \sigma_V^* \epsilon, \\ \epsilon &\sim \mathcal{N}(0,1). \end{aligned}$$

So, we can say that

$$\begin{aligned} TT &= \varphi + E\left[U_1 - U_0 \mid V \geq -Z\gamma\right] \\ &= \varphi + E\left[\epsilon(\sigma_1 - \sigma_0) \mid \epsilon \geq \frac{-Z\gamma}{\sigma_V^*}\right] \\ &= \varphi + (\sigma_1 - \sigma_0) E\left[\epsilon \mid \epsilon \geq \frac{-Z\gamma}{\sigma_V^*}\right] \\ &= \varphi + (\sigma_1 - \sigma_0) \frac{\phi\left(-Z\gamma/\sigma_V^*\right)}{1 - \Phi\left(-Z\gamma/\sigma_V^*\right)}, \end{aligned}$$

and

$$\begin{aligned} TUT &= \varphi + E\left[U_1 - U_0 \mid V < -Z\gamma\right] \\ &= \varphi + E\left[\epsilon(\sigma_1 - \sigma_0) \mid \epsilon < \frac{-Z\gamma}{\sigma_V^*}\right] \\ &= \varphi + (\sigma_1 - \sigma_0) E\left[\epsilon \mid \epsilon < \frac{-Z\gamma}{\sigma_V^*}\right] \\ &= \varphi - (\sigma_1 - \sigma_0) \frac{\phi\left(-Z\gamma/\sigma_V^*\right)}{\Phi\left(-Z\gamma/\sigma_V^*\right)} \\ &= \varphi + (\sigma_0 - \sigma_1) \frac{\phi\left(-Z\gamma/\sigma_V^*\right)}{\Phi\left(-Z\gamma/\sigma_V^*\right)}. \end{aligned}$$

3. Parametrization 1

Now, to add more structure to the problem, suppose that $Z = (1, Z_1, Z_2)$, and $\gamma = (\gamma_0, \gamma_1, \gamma_2)$. Also, suppose that

$$\begin{aligned} \gamma_0 &= 0.2, & \gamma_1 &= 0.3, & \gamma_2 &= 0.1, \\ \sigma_1 &= 0.012, & \sigma_0 &= 0.05, & \sigma_V^* &= 1, \\ \alpha &= 0.02, & \varphi &= 0.2. \end{aligned}$$

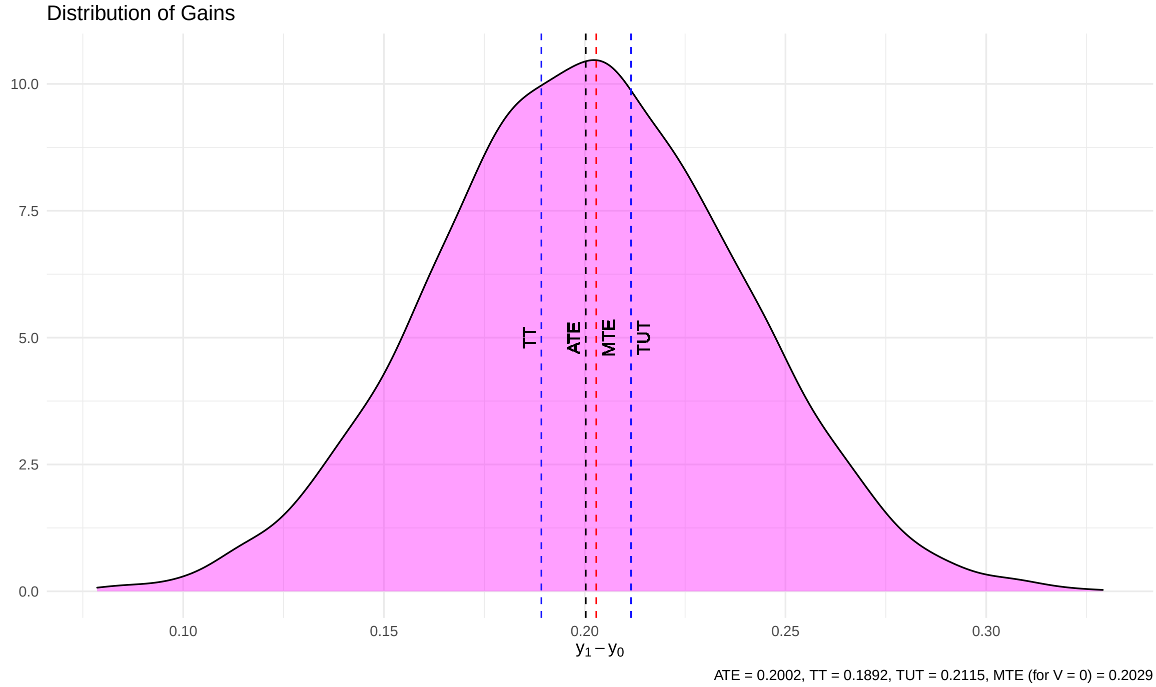
Finally,

$$\begin{aligned} Z_1 &\sim \mathcal{N}(-1, 9), \\ Z_2 &\sim \mathcal{N}(1, 9), \end{aligned}$$

where $Z_1 \perp\!\!\!\perp Z_2$. Since we only observe either Y_1 or Y_0, the econometrician observes the outcome described as follows:

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

```r
set.seed(1234)

N <- 5000 # random sample observations

alpha <- 0.02
phi <- 0.2

g <- matrix(c(0.2, 0.3, 0.1))
sigma1 <- 0.012
sigma0 <- 0.05
sigmaV <- 1

eps <- rnorm(N)

U1 <- sigma1 * eps
U0 <- sigma0 * eps
V <- sigmaV * eps

Z <- cbind(rep(1, N),
           rnorm(N, -1, 9),
           rnorm(N, 1, 9))

I <- (Z %*% g) + V
D <- ifelse(I >= 0, 1, 0)

Y1 <- alpha + phi + U1
Y0 <- alpha + U0

Y <- (D * Y1) + ((1 - D) * Y0)

df <- data.frame('beta' = Y1 - Y0,
                 'decision' = D,
                 'y1' = Y1,
                 'y0' = Y0,
                 'net utility' = I)

ATE <- df %>%
    pull(beta) %>%
    mean()

TT <- df %>%
    filter(decision == 1) %>%
    pull(beta) %>%
    mean()

TUT <- df %>%
    filter(decision == 0) %>%
    pull(beta) %>%
    mean()

MTE <- df %>%
    filter(abs(net.utility - V) <= 0.01) %>%
    pull(beta) %>%
    mean()
```



4. Parametrization 2

A. $LATE(z_1 = -2, z_1' = 1)$.

```r
print(1234)
```

```
## [1] 1234
```

B. $LATE(z_2 = 0, z_2' = 2)$.

```r
print(1234)
```

```
## [1] 1234
```

C. $IV(Z_1)$.

```r
beta.IV.Z1 <- cov(Z[, 2], Y) / cov(Z[, 2], D)
```

$\hat{\beta}_{\text{IV}}(Z_1) = 0.2032.$

D. $IV(Z_2)$.

```r
beta.IV.Z2 <- cov(Z[, 3], Y) / cov(Z[, 3], D)
```

$\hat{\beta}_{\text{IV}}(Z_2) = 0.2046.$

E. OLS.

```r
beta.OLS <- cov(Y, D) / var(D)
```

$\hat{\beta}_{\text{OLS}} = 0.2186.$

5. GMM

6. Likelihood Function

7. Discussion

8. Maximum Likelihood