Homework #3

Professor: Pat Kline

Students: Christina Brown, Sam Leone, Peter McCrory, Preston Mui

Reweighting

Matching

(a) To run OLS on $Y_{i2} - Y_{i1} = \gamma_2 - \gamma_1 + \beta D_{i2} + \epsilon_{i2} - \epsilon_{i1}$, the identifying restriction is that $E\left[D_{i2}(\epsilon_{i2} - \epsilon_{i1})\right] = 0$; that is, program participation is uncorrelated with the change in error terms $\epsilon_{i2} - \epsilon_{i1}$.

(b) $Y_{i2} = Y_{i1} + (\gamma_2 - \gamma_1) + \beta D_{i2} + \epsilon_{i2} - \epsilon_{i1}$

(c) Taking expectations,

$$\begin{aligned} E\left[Y_{i2}|Y_{i1}, D_{i2} = 1\right] &= Y_{i1} + (\gamma_2 - \gamma_1) + \beta + E[\epsilon_{i2} - \epsilon_{i1}|Y_{i1}, D_{i1} = 1] \\ E\left[Y_{i2}|Y_{i1}, D_{i2} = 0\right] &= Y_{i1} + (\gamma_2 - \gamma_1) + E[\epsilon_{i2} - \epsilon_{i1}|Y_{i1}, D_{i1} = 0] \end{aligned}$$

the difference of these two identifies β if

$$\begin{aligned} E[\epsilon_{i2} - \epsilon_{i1}|Y_{i1}, D_{i2} = 1] - E[\epsilon_{i2} - \epsilon_{i1}|Y_{i1}, D_{i2} = 0] &= 0 \\ E\left[E[\epsilon_{i2} - \epsilon_{i1}|D_{i2} = 1] - E[\epsilon_{i2} - \epsilon_{i1}|D_{i2} = 0]\middle|Y_{i1}\right] &= 0 \\ E\left[D_{i2}(\epsilon_{i2} - \epsilon_{i1})\middle|Y_{i1}\right] &= 0 \end{aligned}$$

(d) This assumption is a bit stricter than that of the assumption in part (a); it is the same restriction in part (a), but conditional on Y_{i1}.

2SLS

(a) The moment conditions defining the 2SLS estimator are

$$\begin{aligned} E\left[Z_i \epsilon_i\right] &= 0 \\ E\left[Z_i^2 \epsilon_i\right] &= 0 \end{aligned}$$

they imply each other, since functions of uncorrelated variables are uncorrelated with each other.

(b) The moment conditions defining the control function estimator are

$$
\begin{aligned}
E\left[Z_i\epsilon_i\right] &= 0 \\
E\left[Z_i v_i\right] &= 0 \\
E\left[\epsilon_i | v_i\right] &= \rho\frac{\sigma_\epsilon}{\sigma_v}v_i \\
E\left[(Y_i - \beta_0 - \beta_1 X_i - \beta_2 X_i^2 - \rho\frac{\sigma_\epsilon}{\sigma_v}v_i)X_i\right] &= 0 \\
E\left[(Y_i - \beta_0 - \beta_1 X_i - \beta_2 X_i^2 - \rho\frac{\sigma_\epsilon}{\sigma_v}v_i)X_i^2\right] &= 0 \\
E\left[(Y_i - \beta_0 - \beta_1 X_i - \beta_2 X_i^2 - \rho\frac{\sigma_\epsilon}{\sigma_v}v_i)v_i\right] &= 0
\end{aligned}
$$

(c) The control function estimator relies on stronger conditions than the 2SLS estimator. The only condition in the 2SLS estimator is that the instrument, Z, is uncorrelated with the error term ϵ. However, the control function estimator relies on specifying the exogenous relationship between X and Z, in particular that the conditional expection of X given Z is linear.

(d) The control function would work better when Z is weak, because it brings to bear more moment restrictions than the 2SLS estimator, whereas the 2SLS estimator only relies on the exogeneity of Z.

(e) The table below contains the estimates for β_1 and β_2 under the 2SLS and control function methods.

  (i) Overall, for all values of γ, the control function estimate appears to be more efficient (the spread of the estimates is much lower), because the control function's "first-stage" is correctly specified, and the estimation involves more moment restrictions.

  (ii) As γ_1 shrinks, the mean 2sls estimate falls (away from the true value of 1.0), although the median appears to remain unbiased. The interquartile range of both estimators increases, although the IQR for the 2sls estimate increases much faster than that of the control function. The control function estimates appear to remain both median and mean-unbiased.

  (iii) If one wanted to test the hypothesis that $\rho\frac{\sigma_\epsilon}{\sigma_v} = 0$, one would need an additional term in the variance of the coefficient on v_i. In particular, by using the two-stage estimate formulas for the asymptotic variance of the estimate of $\rho\frac{\sigma_\epsilon}{\sigma_v}$, one would see that in the "sandwich" estimator, one would replace the term

$$E[s_i(\theta_0, \gamma^*)s_i(\theta_0, \gamma^*)']$$

  where $\theta_0 = (\beta_0, \beta_1, \beta_2, \rho\frac{\sigma_\epsilon}{\sigma_v})$ and $\gamma^* = (\gamma_0, \gamma_1)$, and s_i being the score function (gradient of the squared residual with respect to θ_0), with this term

$$
\begin{gathered}
E[g_i(\theta_0, \gamma^*)g_i(\theta_0, \gamma^*)'] \\
\text{s. t. } g_i(\theta_0, \gamma^*) \equiv s_i(\theta_0, \gamma^*) + E\left[\nabla_\gamma s_i(\theta_0, \gamma^*)\right] r_i(\theta_0, \gamma^*)
\end{gathered}
$$

where r_i is the asymptotic variance of $\sqrt{N}(\hat{\gamma}-\gamma^*)$ and can be obtained from the "first-stage" regression. To construct an estimate of the standard error, one can simply plug in the population analogues into this sandwich estimator and run a t-test with the modified standard error.

Statistic	β_1^{2sls}	β_2^{2sls}	β_1^{cf}	β_2^{cf}
$\gamma = 0.3$				
mean	1.003234	1.076318	.9938653	1.000297
sd	.3412886	3.060352	.1081107	.0204986
p50	.9992083	.9904965	.9961268	1.000156
iqr	.1585197	.3704991	.1449331	.0277689
$\gamma = 0.2$				
mean	.7418318	2.253272	.9936412	.999546
sd	8.550667	47.84852	.1688314	.0207697
p50	.9829823	1.005342	.9991214	.9994298
iqr	.315193	.7726767	.2297587	.0277564
$\gamma = 0.1$				
mean	.8100506	.839673	.9693889	.9988507
sd	7.576605	17.05286	.5007757	.0224328
p50	1.029302	1.044092	1.005126	.9992241
iqr	.7045021	1.448808	.4600121	.0299107

Bootstrap OLS

Bootstrap Probit