ECMA 31100: Problem Set 3

Due March 12 by 11:59PM

Question 1 Consider the heterogeneous IV setting from the Week 5/6 notes. Consider the IV estimand produced by estimating the linear regression model (1) using 2SLS and first stage (2):

$$Y = \beta_0 + \beta_1 D + X'\beta_2 + U \tag{1}$$

$$D = \pi_0 + \pi_1 Z + X'\pi_2 + V; \quad \mathrm{E}\left([1, Z, X']' V\right) = 0. \tag{2}$$

Suppose that $D \in \{0,1\}, Z \in \{0,1\}$, and X is a vector of regressors containing functions of the covariates W, e.g. $X = W$ or $X = (\mathbf{1}(W = w_2), \ldots, \mathbf{1}(W = w_k))$, where k is the number of distinct values W can take, and $\mathbf{1}(W = w_1)$ is omitted to avoid perfect collinearity. The instrument Z is exogenous conditional on W and the monotonicity assumption conditional on W holds. In this question we will consider various specifications of X that alter the interpretation of β_1.

a) Write a formula for the IV estimand β_1^{IV} using the specification above. Decompose it into a weighted average of complier average treatment effects (conditional on W) and always-taker treatment effects (conditional on W), assuming $\mathrm{E}(y_0|W)$ is linear in W. Note that in general the term $\mathrm{E}\left(\mathrm{E}\left(\tilde{Z}|W\right)\mathrm{E}(y_0|W)\right)$ will not disappear like it did in the notes, except, for example, when W is included linearly or in a saturated fashion (see parts c and d). Leave this term in your formula (separated from the two terms weighting complier and always-taker treatment effects) and add in an explanation for why it's there.

bi) Let $\mathrm{BLP}(Z|X)$ denote the best linear predictor of Z given X (including a constant). Show that the numerator of weights on the complier ATEs are negative iff $\mathrm{BLP}(Z|X) > 1$. Give an example of Z and X that produces $\mathrm{BLP}(Z|X) > 1$ for some realization of X.

bii) Show that the weights on the always-taker ATEs must take positive and negative values unless $\mathrm{E}(Z|W) = \mathrm{BLP}(Z|X)$.

c) Suppose $X = (\mathbf{1}(W = w_2), \ldots, \mathbf{1}(W = w_k))$. Use part a) to argue that β^{IV} is a positive weighted average of complier average treatment effects (the average is over the distribution of W). Do this in 2 steps:

- Argue that the numerator of the weights equals $Cov(D, Z|X)$, which is non-negative. What must be true about the proportion of compliers and proportion of those receiving $Z = 1$ and $Z = 0$ at each X for this quantity to be strictly positive at each X?
- Argue that the denominator equals $\mathrm{E}(Cov(D, Z|X))$, which is non-negative. Under what condition are the weights well-defined?

d) Now suppose $X = W$. Show by example that the denominator of the weights can be strictly negative even though monotonicity conditional on W holds.

Hint: Write $\mathrm{E}\left(D\tilde{Z}\right) = \mathrm{E}\left(\mathrm{E}\left(D\tilde{Z}|W\right)\right)$, where $\tilde{Z} = Z - \mathrm{BLP}\left(Z|W\right)$ and decompose $\mathrm{E}\left(D\tilde{Z}|W\right)$ as you did for $\mathrm{E}\left(Y\tilde{Z}|W\right)$ in part a). Now let $W \in \{0, 1, 2\}$, where each value occurs with probability $1/3$, and, for some small $\epsilon > 0$, let

$$\mathrm{E}\left(Z|W\right) = \begin{cases} \epsilon & W = 0 \\ 1 - \epsilon & W \in \{1, 2\} \end{cases},$$

which ensures $1 - \mathrm{BLP}\left(Z|W\right) < 0$ for $W = 2$. Now let the proportion of never takers vary with W.

e) Now suppose there are no covariates (X is empty) but instead of a binary instrument we have a multi-valued instrument Z which satisfies the monotonicity condition. Suppose the first stage regression is estimated using dummy variables for each possible value of Z (without the first value, so we can keep the constant). Show that the IV estimand is a weighted average of LATEs, and interpret the weights. Hint: Imbens Angrist 1994.

f) Now suppose there are no covariates (X is empty) but instead of a binary treatment we have a multi-valued treatment $D \in \{0, 1, \ldots, K\}$ and binary instrument. Note that there are now $K + 1$ potential outcomes $y_0, \ldots, y_K$. Show that

$$\beta_{IV} = \sum_{k=1}^{K} \left[\frac{\mathrm{P}\left(D_1 \geq k > D_0\right)}{\sum_{m=1}^{K} \mathrm{P}\left(D_1 \geq m > D_0\right)} \right] \mathrm{E}\left(Y_k - Y_{k-1}|D_1 \geq k > D_0\right).$$

How do you interpret the weights, expectations and estimand as a whole?

Question 2 Read Abadie (2003) (here). The data used in Section 6 is on canvas. Reproduce columns 1-3 of Table 2. Explain how you would use Abadie's Kappa with a linear local average response function (defined in section 4.2.1 of that paper) to reproduce column 4. Implement this using logistic regression to estimate $\mathrm{P}\left(Z = 1|X\right)$. Are your parameter estimates much different? Note: You don't need to write your own code for the logistic regression or produce standard errors for the estimates constructed using Abadie's Kappa.

Question 3 a) What does a triple difference in difference argument allow for that an ordinary difference in difference argument doesn't? Give an example where you might use a triple diff in diff rather than a double difference, and an example where a triple difference would actually lead to a biased ATT estimate but a double difference wouldn't.

b) Suppose you use a double difference argument and have one treatment and one control group ($g = 0, 1$), one post-treatment date t_0 and several pre-treatment dates $t = 0, \ldots t_0 - 1$. You don't believe the common trends assumption, but still want to identify the ATT for the treated group in the post-period. You make a parametric trends assumption:

$$\mathrm{E}\left(y\left(0\right)|G, T\right) = \alpha + \sum_{t=1}^{t_0} \beta_t \mathbf{1}\left(T = t\right) + G\left(\gamma + \delta_1 T + \delta_2 T^2\right).$$

Show that this assumption relaxes common trends, but then specify the regression which identifies ATT_{t_0}. Is there an issue with collinearity here? (Think about how large t_0 must be to avoid it).

Question 4 Suppose we have a panel $\{(y_{it}, x_{it}) : t = 1, \ldots, T\}_{i=1}^{N}$. Show that the OLS estimates of β are numerically equivalent in the following specifications:

$$
\begin{aligned}
y_{it} &= x_{it}'\beta + \gamma_i + u_{it}; \\
y_{it} - \bar{y}_i &= (x_{it} - \bar{x}_i)'\,\beta + \epsilon_{it},
\end{aligned}
$$

where γ_i is an individual fixed effect (adds a dummy for each i into the specification) and $\bar{y}_i = \frac{1}{T}\sum_{i=1}^{T} y_{it}$ and $\bar{x}_i = \frac{1}{T}\sum_{i=1}^{T} x_{it}$. x_{it} does not contain a constant to avoid collinearity.