

## Econometrics II, Problem Set 3

Eyayaw Beze

May 19, 2020

### Exercise 1: Panel data model

Consider the following between model with a single explanatory variable  $\bar{x}_i$ :

$$\bar{y}_i = \beta_0 + \beta_1 \bar{x}_i + a_i + \bar{u}_i \quad (1)$$

$a_i + \bar{u}_i$  are unobserved.  $E(a_i) = 0$ ,  $E(\bar{u}_i) = 0$ ,  $E(\bar{x}_i \bar{u}_i) = 0$ , and  $Cov(x_{it}, a_{it}) = \sigma_{xa} \neq 0$

(a) Show that the between estimator  $\hat{\beta}_1^B$  satisfies the following probability limit property:

$$\begin{aligned} plim \hat{\beta}_1^B &= \beta_1 + \frac{\sigma_{xa}}{Var(\bar{x}_i)} \quad \text{if } N \rightarrow \infty \\ \\ p \lim_{n \rightarrow \infty} \hat{\beta}_1^B &= \frac{Cov(\bar{y}_i, \bar{x}_i)}{Var(\bar{x}_i)} = \frac{Cov(\beta_0 + \beta_1 \bar{x}_i + a_i + \bar{u}_i, \bar{x}_i)}{Var(\bar{x}_i)} \\ &\quad \quad \quad = 0 \text{ by orthogonality} \\ &= \frac{Cov(\beta_0 + \beta_1 \bar{x}_i, \bar{x}_i) + Cov(a_i, \bar{x}_i) + \overbrace{Cov(\bar{u}_i, \bar{x}_i)}}{Var(x)} \\ &= \frac{Cov(\beta_0 + \beta_1 \bar{x}_i, \bar{x}_i) + \sigma_{xa}}{Var(x)} \\ &= \frac{\beta_0 E(\bar{x}_i) + \beta_1 E(\bar{x}_i^2) - E(\bar{x}_i)[\beta_0 + \beta_1 E(\bar{x}_i)] + \sigma_{xa}}{Var(\bar{x}_i)} \\ &= \frac{\beta_1 [E(\bar{x}_i^2) - [E(\bar{x}_i)]^2] + \sigma_{xa}}{Var(\bar{x}_i)} \\ &= \beta_1 + \frac{\sigma_{xa}}{Var(\bar{x}_i)} \end{aligned}$$

(b)

$$\begin{aligned}
Var(\bar{x}_i) &= Var\left(\frac{1}{T} \sum_{t=1}^T x_{it}\right) \\
&= \frac{1}{T^2} Var\left(\sum_{t=1}^T x_{it}\right) \\
&= \frac{1}{T^2} \left(\sum_{t=1}^T Var(x_{it})\right) \\
&= \frac{1}{T^2} T Var(x_{it}) \\
&= \frac{\sigma_x^2}{T}
\end{aligned}$$

Then,

$$\begin{aligned}
p \lim_{n \rightarrow \infty} \hat{\beta}_1^B &= \beta_1 + \frac{\sigma_{xa}}{Var(\bar{x}_i)} \\
&= \beta_1 + \frac{\sigma_{xa}}{\frac{\sigma_x^2}{T}} \\
&= \beta_1 + T \frac{\sigma_{xa}}{\sigma_x^2}
\end{aligned}$$

(c) When T grows  $Var(\bar{x}_i)$  goes down and consequently,  $\beta_1$  would be overestimated.

The assumption that the  $x_{it}$ 's are uncorrelated overtime is not necessarily reasonable. Because  $x_{it}$ 's could be serially correlated overtime for some i's or for some  $x_i$ 's such as income, education, knowledge etc. Income of individuals is likely to be correlated overtime.

## Exercise 2: FE vs LDV

$$y_{it} = \beta_1 y_{it-1} + \beta_2 X_{it} + \delta D_{it} + \alpha_{it} + \tau_t + u_{it} \quad \text{for } t = 1, 2, 3, 4, 5 \quad (2)$$

(a) Consistent estimation of  $\delta$  through OLS in equation (2) requires the usual Gauss-Markov set of assumptions to be satisfied. Particularly, no (perfect) correlation among the regressors; the lagged dependent variable  $y_{it-1}$ ,  $X_{it}$ , the treatment  $D_{it}$ , the year dummies  $\tau_t$ , and individual fixed effects  $\alpha_i$  should be orthogonal to the error term  $u_{it}$ .

However, if the treatment is correlated with an unobserved individual effect,  $\alpha_i$ , we may try to kill it (i.e, the fixed effect) by differencing,

$$\Delta y_{it} = \beta_1 \Delta y_{it-1} + \beta_2 \Delta X_{it} + \delta \Delta D_{it} + \Delta \tau_t + \Delta u_{it} \quad \text{for } t = 2, 3, 4, 5 \quad (3)$$

This may not result in consistent estimator for the treatment effect. The problem here is that the differenced residual,  $\Delta u_{it}$ , is necessarily correlated with the lagged dependent variable,  $\Delta y_{it-1}$ , because both are a function of  $u_{it-1}$ . Consequently, the OLS estimates of the first differenced equation (3) are not consistent for the parameters in equation (2) (Angrist and Pischke 2009, 182–85).

Since we have more than three time periods, one solution to the above problem is to use  $y_{it-2}$  as an instrument for  $\Delta y_{it-1}$  in (3). But this requires that  $y_{it-2}$  be uncorrelated with the differenced residuals,  $\Delta u_{it}$ . However, if  $u_{it}$  is serially correlated, there may be no consistent estimator for (3).

- (b) An empirical guy can check the robustness of his or her findings using alternative identifying assumptions and should expect broadly similar results from using both models, fixed effects or lagged dependent variables. In the particular setup of (2), the estimates of  $\delta$  should be bounded by the estimates of fixed effects and lagged dependent variables (bracketing property) (Angrist and Pischke 2009, 182–85).

Mistakenly using fixed effects or lagged dependent variable model to estimate the causal effect, while the above pooled OLS set up is correct, may result in either over or underestimating the estimates of the treatment effect.

## Exercise 3: First Differences

$$\log(uclms_{it}) = \beta_1 ez_{it} + \alpha_i + \tau_t + u_{it} \quad (4)$$

(a)  $\hat{\beta}_1 = -0.18188$

The interpretation of the estimate of the treatment effect  $\beta_1$  would be: having enterprise zone in a city would lower (growth of) unemployment claims by about 18.19 percent, and the effect is statistically significant at 5 % level.

The differenced residual,  $\Delta u_{it}$ , could be correlated with the lagged differenced residual  $u_{it-1}$ . So in our estimation we should account for this serial correlation in calculating standard errors and test statistics.

$$\log(uclms_{it}) = \beta_1 ez_{it} + \alpha_i + \alpha_i * year_t + \tau_t + u_{it} \quad (5)$$

(b)  $\hat{\beta}_1 = -0.23834$

The estimate of  $\beta_1$  is slightly higher in this case. Thus, having enterprise zone reduces unemployment claims by about 23.83 percent while controlling for city and year (as interaction), but not for year dummies.

$$\log(uclms_{it}) = \beta_1 ez_{it} + \alpha_i + \alpha_i * year_t + \tau_t + u_{it} \quad (6)$$

(c)  $\hat{\beta}_1 = -0.18133$

The magnitude of effect of enterprise zone (about 18.13 percent) is the same as in (4) (where we control for full set of year dummies), adding an interaction between year and city fixed effects (like in (5)) does not seem to add effect-as compared to (4).

Please refer to the do.file: [do\\_problem\\_set\\_3.do](#)

## References

Angrist, Joshua, and Joern-Steffen Pischke. 2009. “Mostly Harmless Econometrics: An Empiricist’s Companion,” January. <https://press.princeton.edu/books/paperback/9780691120355/mostly-harmless-econometrics>.