Problem Set #6

Danny Edgel
Econ 710: Economic Statistics and Econometrics II
Spring 2021

March 2, 2021

Discussed and/or compared answers with Sarah Bass, Emily Case, Katherine Kwok, Michael Nattinger, and Alex Von Hafften

Question 1

(i) The direct representation of the sample average is:

$$\mu_0 = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{T_i}\sum_{t=1}^{T_i}Y_{it}$$

Since 1_i contains T_i elements, $1_i'1_i = T_i$, and $1_i'Y_i = \sum_{t=1}^{T_i} Y_i$. It is clear, then, that

$$\mathbb{E}\left[\hat{\mu}_{OLS}\right] = \frac{\sum_{i=1}^{n} 1_i'Y_i}{\sum_{i=1}^{n} 1_i'1_i}$$

(ii) We can solve for the variance of $\hat{\mu}_{IV}$ as follows:

$$Var(\hat{\mu}_{IV}) = Var\left(\frac{\sum_{i=1}^{n} Z_i'Y_i}{\sum_{i=1}^{n} Z_i'1_i}\right) = \frac{\sum_{i=1}^{n} Z_i'Var(Y_i)Z_i}{\left(\sum_{i=1}^{n} Z_i'1_i\right)^2}$$

Where:

$$Var(Y_i) = Var(\mu_0 + \alpha_i + \varepsilon_{it}) = Var(\alpha_i) + Var(\varepsilon_{it}) + 2Cov(\alpha_i, \varepsilon_{it})$$
$$\Rightarrow \Omega_i = \sigma_\alpha^2 1_i 1_i' + \sigma^2 I_{T_i}$$

(iii) To determine how we can show that $Var(\hat{\mu}_{IV} \geq \left(\sum_{i=1}^n 1_i'\Omega_i^{-1}1_i\right)^{-1}$, we simply need to find that the following inequality holds:

$$\left(\sum_{i=1}^n Z_i'1_i\right)^2 \leq \left(\sum_{i=1}^n Z_i'\Omega_i Z_i\right)\left(\sum_{i=1}^n 1_i'\Omega_i^{-1}1_i\right)$$

Once this inequality is established, it follows that:

$$Var(\hat{\mu}_{IV}) = \frac{\sum_{i=1}^n Z_i'Var(Y_i)Z_i}{\left(\sum_{i=1}^n Z_i'1_i\right)^2} \geq \frac{\sum_{i=1}^n Z_i'Var(Y_i)Z_i}{\left(\sum_{i=1}^n Z_i'\Omega_i Z_i\right)\left(\sum_{i=1}^n 1_i'\Omega_i^{-1}1_i\right)} = \left(\sum_{i=1}^n 1_i'\Omega_i^{-1}1_i\right)^{-1}$$

We can establish the inquality using the Cauchy-Schwarz inequality:

$$\left(\sum_{i=1}^n Z_i'1_i\right)^2 = \left(\sum_{i=1}^n Z_i'\Omega^{1/2}\Omega^{-1/2}1_i\right)^2 \leq \left(\sum_{i=1}^n Z_i'\Omega_i Z_i\right)\left(\sum_{i=1}^n 1_i'\Omega_i^{-1}1_i\right)$$

This variance is achieved by $\overline{Z}_i = \Omega_i^{-1}1_i$

(iv) If $T_i = T$ for all i (i.e., the panel is balanced), then the GLS estimator weights by the entries of Σ^{-1}, where:

$$\Sigma^{-1} = \frac{1}{\sigma^2}\left(I_T - \frac{\sigma_\alpha^2 T}{\sigma^2 + \sigma_T^2}\frac{1_i 1_i'}{T}\right)$$

Then, the optimal instrument for GLS is:

$$\overline{Z}_i = \Sigma^{-1}1_i = \frac{1_i}{\sigma^2}\left(1 - \frac{\sigma_\alpha^2 T}{\sigma^2 + \sigma_T^2}\right) = \frac{1_i}{\sigma^2 + \sigma^2 T}$$

This instrument cancels out in the estimator for μ_0, yielding the OLS estimator:

$$\hat{\mu}_{GLS} = \frac{\sum_{i=1}^n \overline{Z}_i'Y_i}{\sum_{i=1}^n \overline{Z}_i'1_i} = \frac{\sum_{i=1}^n \frac{1}{\sigma^2+\sigma_T^2}1_i'Y_i}{\sum_{i=1}^n \frac{1}{\sigma^2+\sigma_T^2}1_i'1_i} = \frac{\sum_{i=1}^n 1_i'Y_i}{\sum_{i=1}^n 1_i'1_i} = \hat{\mu}_{OLS}$$

Thus, if the panel is balanced, OLS and GLS are identical.

(v) First, note that:

$$\overline{Y} = \frac{1}{T_i}\sum t = 1^{T_i}\mu_0 + \alpha_i + \varepsilon_{it} = \mu_0 + \alpha_i + \frac{1}{T_i}\sum t = 1^{T_i}\varepsilon_{it}$$

And let $\overline{\varepsilon} = \frac{1}{T_i} \sum t = 1^{T_i} \varepsilon_{it}$. Then,

$$\begin{aligned}
\mathbb{E}\left[\hat{\sigma}_i^2\right] &= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[(\mu_0 + \alpha_i + \varepsilon_{it} - \mu_0 - \alpha_i - \overline{\varepsilon})^2\right] \\
&= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[(\varepsilon_{it} - \overline{\varepsilon})^2\right] \\
&= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[(\varepsilon_{it} - \overline{\varepsilon})\varepsilon_{it}\right] \\
&= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[\varepsilon_{it}^2\right] - \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[\varepsilon_{it}\overline{\varepsilon})\right] \\
&= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[\varepsilon_{it}^2\right] - \frac{1}{T_i(T_i - 1)} \sum_{t=1}^{T_i} \mathbb{E}\left[\varepsilon_{it}^2\right] - \frac{1}{T_i(T_i - 1)} \sum_{t=1}^{T_i} \sum_{s \neq t}^{T_i} \mathbb{E}\left[\varepsilon_{it}\varepsilon_{is}\right] \\
&= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \sigma^2 - \frac{1}{T_i(T_i - 1)} \sum_{t=1}^{T_i} \sigma^2 \\
&= \sigma^2
\end{aligned}$$

(vi) From (v) and letting $\mu = \mu_0$, we can derive,

$$\begin{aligned}
\mathbb{E}\left[\hat{\sigma}_{\alpha,i}^2\right] = \mathbb{E}\left[\frac{1}{T_i - 1} \sum_{t=1}^{T_i} (Y_{it} - \mu_0)^2 - \hat{\sigma}_i^2\right] &= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[(Y_{it} - \mu_0)^2\right] - \sigma_i^2 \\
&= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[Y_{it}^2\right] - 2\mu_0 \mathbb{E}\left[Y_{it}\right] + \mu_0^2 - \sigma_i^2 \\
&= \frac{1}{T_i - 1} \sum_{t=1}^{T_i} \mathbb{E}\left[(\mu_0 + \alpha_i + \varepsilon_{it})^2\right] - 2\mu_0^2 + \mu_0^2 - \sigma_i^2 \\
&= \mathbb{E}\left[(\mu_0 + \alpha_i)^2\right] - \mu_0^2 - \sigma_i^2 + \frac{1}{T_i - 1} \sum_{t=1}^{T_i} -2\mu_0 \mathbb{E}\left[\varepsilon_{it}\right] - 2\mathbb{E}\left[\alpha_i \varepsilon_{it}\right] + \mathbb{E}\left[\varepsilon_{it}^2\right] \\
&= \mu_0^2 + 2\mu_0 \mathbb{E}\left[\alpha_i\right] + \mathbb{E}\left[\alpha_i^2\right] - \mu_0^2 - \sigma_i^2 + \sigma_i^2 \\
&= \sigma_\alpha^2
\end{aligned}$$

This shows that $\hat{\sigma}_{\alpha,i}^2$ is an unbiased estimator of σ_α^2 for each i. Then $\frac{1}{N} \sum_{n=1}^{N} \hat{\sigma}_{\alpha,i}(\hat{\mu}_{OLS})$ must be an unbiased estimator of σ_α^2, as well.

(vii) We can use the fact that FGLS and GLS have the same asymptotic variance to construct a variance estimator for FGLS. The asymptotic variance for GLS is:

$$V = \left(\sum_{i=1}^{n} 1_i' \Omega_i^{-1} 1_i \right)^{-1} = \left(\sum_{i=1}^{n} \frac{T_i}{T_i \sigma_\alpha^2 + \sigma^2} \right)^{-1}$$

Then, an estimator for the asymptotic variance for FGLS is:

$$\hat{V} = \left(\sum_{i=1}^{n} \frac{T_i}{T_i \hat{\sigma}_\alpha^2 + \hat{\sigma}^2} \right)^{-1}$$

Question 2

(i) The asymptotic bias of the fixed effects estimator can be deduced from the probability limit of $\hat{\beta}_{FE}$:

$$\hat{\beta}_{FE} \to_p \beta_0 + \frac{\mathbb{E}\left[\sum_{t=1}^{T} (X_{it} - \overline{X}_i)\varepsilon_{it}\right]}{\mathbb{E}\left[\sum_{t=1}^{T} (X_{it} - \overline{X}_i)^2\right]}$$

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} (X_{it} - \overline{X}_i)\varepsilon_{it}\right] &= \sum_{t=1}^{T} \mathbb{E}\left[X_{it}\varepsilon_{it}\right] - \mathbb{E}\left[\overline{X}_i \varepsilon_{it}\right] \\
&= -\sum_{t=1}^{T} \mathbb{E}\left[\frac{1}{T} \sum_{s=1}^{T} X_{is}\varepsilon_{it}\right] \\
&= -\sum_{t=1}^{T} \mathbb{E}\left[X_{it+1}\varepsilon_{it}\right] = -\sum_{t=1}^{T} \delta \sigma_X^2 \\
\Rightarrow \hat{\beta}_{FE} \to_p \beta_0 &- \frac{\sum_{t=1}^{T} \delta \sigma_X^2}{\mathbb{E}\left[\sum_{t=1}^{T} (X_{it} - \overline{X}_i)^2\right]}
\end{aligned}$$

The asymptotic bias of the first differences estimator can be similarly derived:

$$\hat{\beta}_{FD} \to_p \beta_0 + \frac{\mathbb{E}\left[\sum_{t=2}^{T} (X_{it} - X_{it-1})(\varepsilon_{it} - \varepsilon_{it-1})\right]}{\mathbb{E}\left[\sum_{t=2}^{T} (X_{it} - X_{it-1})^2\right]}$$

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=2}^{T} (X_{it} - X_{it-1})(\varepsilon_{it} - \varepsilon_{it-1})\right] &= \sum_{t=2}^{T} \mathbb{E}\left[X_{it}\varepsilon_{it}\right] - \mathbb{E}\left[X_{it}\varepsilon_{it-1}\right] - \mathbb{E}\left[X_{it-1}\varepsilon_{it}\right] + \mathbb{E}\left[X_{it-1}\varepsilon_{it-1}\right] \\
&= -\sum_{t=2}^{T} \mathbb{E}\left[X_{it}\varepsilon_{it-1}\right] = -\sum_{t=2}^{T} \delta \sigma_X^2 \\
\Rightarrow \hat{\beta}_{FD} \to_p \beta_0 &- \frac{\sum_{t=2}^{T} \delta \sigma_X^2}{\mathbb{E}\left[\sum_{t=2}^{T} (X_{it} - X_{it-1})^2\right]}
\end{aligned}$$

(ii) Yes; if $T = 2$, then the bias of these two estimators is identical.

Question 3

The table below presents the results of the simulations.[1] As you can see, the pooled OLS estimator of β_0 is substantially biased upward. An intuitive but less clearly apparent result is that both the fixed effect estimation of β_0 is generally efficient but much more efficient for larger n and no autocorrelation. However, the FE estimate is still biased upward in all specifications, though the bias decreases in the FE model for larger n (this is not true for OLS).

The coverage results are more muddled. The confidence intervals are more accurate with heteroskedasticity robust SEs when there is no autocorrelation, but the cluster robust SEs provide more accurate confidence intervals in the presence of autocorrelation. The latter result is consistent with theory, since the autocorrelation only occurs within the groups which get clustered in the cluster robust SEs.

		OLS		FE, Robust		FE, Cluster	
n	ϕ	Mean	Coverage	Mean	Coverage	Mean	Coverage
40	0	2.544	0.909	1.176	0.845	1.176	0.776
40	.8	2.570	0.812	1.198	0.772	1.198	0.789
70	0	2.663	0.956	1.161	0.922	1.161	0.893
70	.8	2.668	0.862	1.149	0.851	1.149	0.896
100	0	2.678	0.960	1.122	0.943	1.122	0.936
100	.8	2.660	0.859	1.111	0.875	1.111	0.931

[1] See the attached .do file for the code used to generate this table.