Problem Set #6 Sarah Bass

1a) $$\mu_0 = \frac{1}{\sum_{i=1}^n T_i} \sum_{i=1}^{n} \sum_{t=1}^{T_i} y_{it}$$ average across all observations, all time periods

$$= \frac{\sum_{i=1}^n 1_i' Y_i}{\sum_{i=1}^n 1_i' 1_i} = \hat{\mu}_{OLS}$$

1b) $$\hat{u}_{IV} = \frac{\sum_{i=1}^n z_i' Y_i}{\sum_{i=1}^n z_i' 1_i}$$

$$= \frac{\sum_{i=1}^n z_i'(\mu_0 + \alpha_i 1_i + \varepsilon_i)}{\sum_{i=1}^n z_i' 1_i}$$

$$= \mu_0 + \frac{\sum_{i=1}^n z_i'(\alpha_i 1_i + \varepsilon_i)}{\sum_{i=1}^n z_i' 1_i}$$

$$var(\hat{\mu}_{IV}) = var\left(\frac{\sum_{i=1}^n z_i'(\alpha_i 1_i + \varepsilon_i)}{\sum_{i=1}^n z_i' 1_i}\right)$$

$$= \frac{\sum_{i=1}^n z_i' \, var(\alpha_i 1_i + \varepsilon_i) \, z_i}{(\sum_{i=1}^n z_i' 1_i)^2}$$

$$= \frac{\sum_{i=1}^n z_i' \Omega_i z_i}{(\sum_{i=1}^n z_i' 1_i)^2},$$ where

$$\Omega_i = var(\alpha_i 1_i + \varepsilon_i)$$

$$= var(\alpha_i 1_i) + var(\varepsilon_i)$$

$$= 1_i 1_i' \sigma_\alpha^2 + \sigma^2 T_i$$

1c) Using Cauchy Schwarz:

$$z_i' 1_i = z_i' \Omega_i^{1/2} \Omega_i^{-1/2} 1_i \le \| \Omega_i^{1/2} z_i \| \, \| \Omega_i^{-1/2} 1_i \|$$

and

$$\left(\sum_{i=1}^n z_i' 1_i\right)^2 \le \left(\sum_{i=1}^n \| \Omega_i^{1/2} z_i \| \, \| \Omega_i^{-1/2} 1_i \|\right)^2$$
$$\le \sum_{i=1}^n z_i' \Omega_i z_i \cdot \sum_{i=1}^n 1_i' \Omega_i^{-1} 1_i$$

So,

$$\mathrm{var}(\hat{\mu}_{IV}) = \frac{\sum_{i=1}^n z_i' \Omega_i z_i}{\left(\sum_{i=1}^n z_i' 1_i\right)^2}$$
$$\ge \frac{\sum_{i=1}^n z_i' \Omega_i z_i}{\sum_{i=1}^n z_i' \Omega_i z_i \cdot \sum_{i=1}^n 1_i' \Omega_i^{-1} 1_i}$$
$$= \frac{1}{\sum_{i=1}^n 1_i' \Omega_i^{-1} 1_i}$$

Let $\tilde{z} = \Omega_i^{-1} 1_i$. Then the instrument has the variance

$$\frac{\sum_{i=1}^n \tilde{z}_i' \Omega_i \tilde{z}_i}{\left(\sum_{i=1}^n \tilde{z}_i' 1_i\right)^2} = \frac{\sum_{i=1}^n 1_i' \Omega_i^{-1} \Omega_i \Omega_i^{-1} 1_i}{\left(\sum_{i=1}^n 1_i' \Omega_i^{-1} 1_i\right)^2} = \frac{1}{\sum_{i=1}^n 1_i' \Omega_i^{-1} 1_i}$$

1d) Let $T_i = T$ for all i. Using the Sherman Morrison formula,

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

So, $$\tilde{z}_i = \Omega_i^{-1} 1_i = \frac{1_i}{\sigma^2}\left(1 - \frac{T_i \sigma_\alpha^2}{T_i \sigma_\alpha^2 + \sigma^2}\right) = \frac{1_i}{T_i \sigma_\alpha^2 + \sigma^2} = \frac{1_i}{T\sigma_\alpha^2 + \sigma^2} \text{ for } T_i = T$$

Then $$\hat{\mu}_{GLS} = \frac{\sum_{i=1}^n \tilde{z}_i' y_i}{\sum_{i=1}^n \tilde{z}_i' 1_i}$$
$$= \frac{\sum_{i=1}^n \frac{1}{T\sigma_\alpha^2 + \sigma^2} 1_i' y_i}{\sum_{i=1}^n \frac{1}{T\sigma_\alpha^2 + \sigma^2} 1_i' 1_i}$$
$$= \frac{\sum_{i=1}^n 1_i' y_i}{\sum_{i=1}^n 1_i' 1_i} = \hat{\mu}_{OLS}$$

Thus GLS and OLS are the same estimators when $T_i = T$ for all i, so GLS is not more efficient than OLS.

1e) Let $\bar{\varepsilon}_i = \frac{1}{T_i}\sum_{t=1}^{T_i} \varepsilon_{it}$. Then,

$$\hat{\sigma}_i^2 = \frac{1}{T_i - 1}\sum_{t=1}^{T_i} (y_{it} - \bar{y}_i)^2$$

$$= \frac{1}{T_i - 1}\sum_{t=1}^{T_i} (\varepsilon_{it} - \bar{\varepsilon}_i)^2$$

$$= \frac{1}{T_i - 1}\sum_{t=1}^{T_i} (\varepsilon_{it} - \bar{\varepsilon}_i)\varepsilon_{it}$$

$$= \frac{1}{T_i - 1}\sum_{t=1}^{T_i} \varepsilon_{it}^2 - \frac{1}{T_i(T_i-1)}\sum_{t=1}^{T_i}\sum_{s=1}^{T_i} \varepsilon_{is}\varepsilon_{it}$$

$$= \frac{1}{T_i}\sum_{t=1}^{T_i} \varepsilon_{it}^2 - \frac{1}{T_i(T_i-1)}\sum_{t=1}^{T_i}\sum_{s=1,\, s\neq t}^{T_i} \varepsilon_{is}\varepsilon_{it}$$

Note $E[\varepsilon_{is}\varepsilon_{it}] = \sigma^2 1(s=t)$ is unbiased because

$$E[\hat{\sigma}_i^2] = E\left[\frac{1}{T_i}\sum_{t=1}^{T_i} \varepsilon_{it}^2 - \frac{1}{T_i(T_i-1)}\sum_{t=1}^{T_i}\sum_{s=1,\, s\neq t}^{T_i} \varepsilon_{is}\varepsilon_{it}\right]$$

$$= \frac{1}{T_i}\sum_{t=1}^{T_i} \sigma^2 - \frac{1}{T_i(T_i-1)}\sum_{t=1}^{T_i}\sum_{s=1,\, s\neq t}^{T_i} 0 = \sigma^2$$

$\hat{\sigma}^2$ is unbiased since it is an average of independent unbiased estimators. Further, since $\hat{\sigma}^2$ is an average of independent random variables, $\text{var}(\hat{\sigma}_i^2)$ is bounded and consistent.

1f) $$\hat{\sigma}_{\alpha,i}^2(\mu_0) + \hat{\sigma}_i^2 = \frac{1}{T_i}\sum_{t=1}^{T_i} (y_{it} - \mu_0)^2$$

$$= \frac{1}{T_i}\sum_{t=1}^{T_i} (\alpha_i + \varepsilon_{it})^2$$

$$= \alpha_i^2 + \frac{1}{T_i}\sum_{t=1}^{T_i} (2\alpha_i\varepsilon_{it} + \varepsilon_{it}^2)$$

continued on next page.

Note, $E[\alpha_i \varepsilon_{it}] = 0$. So,

$$E\left[\hat{\sigma}^2_{\alpha,i}(\mu_0) + \hat{\sigma}^2_{\varepsilon,i}\right] = E\left[\alpha_i^2 + \frac{1}{T_i}\sum_{t=1}^{T_i}\left(2\alpha_i \varepsilon_{it} + \varepsilon_{it}^2\right)\right]$$

$$= \sigma_\alpha^2 + \sigma^2$$

Along with part e, this shows that $\hat{\sigma}^2_{\alpha,i}(\mu_0)$ is unbiased.

1g) Note, FGLS has the same asymptotic variance as GLS

Let $\hat{V} = \left(\sum_{i=1}^{n} \frac{T_i}{T_i \hat{\sigma}_\alpha^2 + \hat{\sigma}^2}\right)^{-1}$ be an estimator for the GLS variance.

2a) $$\hat{\beta}_{FE} \to_p B_0 + \frac{E\left[\sum_{t=1}^T (x_{it} - \bar{x}_i)\varepsilon_{it}\right]}{E\left[\sum_{t=1}^T (x_{it} - \bar{x}_i)^2\right]}$$

$$= B_0 + \frac{E\left[\sum_{t=1}^T x_{it}\varepsilon_{it}\right] - E\left[\sum_{t=1}^T \bar{x}_i \varepsilon_{it}\right]}{(T-1)\sigma_x^2}$$

$$= B_0 - \frac{E\left[\sum_{t=1}^T \bar{x}_i \varepsilon_{it}\right]}{(T-1)\sigma_x^2}$$

Note $\bar{x} = \frac{\sum_{s=1}^T x_{is}}{T}$

$$= B_0 - \frac{E\left[\sum_{t=1}^T \sum_{s=1}^T x_{is}\varepsilon_{it}\right]}{T(T-1)\sigma_x^2}$$

$x_{it} = \delta \varepsilon_{i,t-1} + u_{it}$

$E[x_{is}\varepsilon_{it}] = \delta \sigma_x^2 \mathbb{1}\{s = t+1\}$

$$= B_0 - \frac{(T-1)\delta\sigma_x^2}{T(T-1)\sigma_x^2}$$

$$= B_0 - \frac{\delta}{T}$$

continued on next page.

2a) $\hat{\beta}_{FD} \rightarrow_p \beta_0 + \dfrac{E\left[\sum_{t=2}^{T} (X_{it} - X_{i,t-1})(\varepsilon_{it} - \varepsilon_{i,t-1})\right]}{E\left[\sum_{t=2}^{T} (X_{it} - X_{i,t-1})^2\right]}$

$= \beta_0 + \dfrac{E\left[\sum_{t=2}^{T} X_{it}\varepsilon_{it} - X_{it}\varepsilon_{i,t-1} - X_{i,t-1}\varepsilon_{it} + \varepsilon_{it}\varepsilon_{i,t-1}\right]}{E\left[\sum_{t=2}^{T} X_{it}^2 - 2X_{it}X_{i,t-1} + X_{i,t-1}^2\right]}$

$= \beta_0 - \dfrac{(T-1)\delta\sigma_X^2}{2(T-1)\sigma_X^2}$

$= \beta_0 - \dfrac{\delta}{2}$

2b) The asymptotic biases are the same if $T=2$

Question 3

Part A

	FE	OLS	δ_2	δ_3	δ_4	Robust LB	Robust UB	Cluster LB	Cluster UB
$\phi = 0, n = 40$	1.2214	2.3122	1.2101	0.73138	0.76984	0.68019	1.7627	0.72118	1.7217
$\phi = 0.8, n = 40$	1.3523	2.8034	0.77108	0.66489	0.60957	0.79652	1.908	0.49682	2.2077
$\phi = 0, n = 70$	0.66709	2.5265	0.95294	1.0104	1.0177	0.21289	1.1213	0.077479	1.2567
$\phi = 0.8, n = 70$	1.7096	3.7238	0.87402	0.62422	0.77745	1.2558	2.1633	1.0825	2.3366
$\phi = 0, n = 100$	0.84789	2.7105	0.98069	0.93457	1.1414	0.40468	1.2911	0.3488	1.347
$\phi = 0.8, n = 100$	0.97849	2.787	0.88999	0.87126	0.87842	0.66157	1.2954	0.49966	1.4573

Part B Simulations

	FE	OLS	Robust Coverage	Cluster Coverage
$\phi = 0, n = 40$	1.019	2.6498	0.92	0.95
$\phi = 0.8, n = 40$	0.93012	2.5386	0.76	0.92
$\phi = 0, n = 70$	1.0452	2.7123	0.92	0.95
$\phi = 0.8, n = 70$	0.93225	2.729	0.87	0.96
$\phi = 0, n = 100$	0.97549	2.6976	0.85	0.89
$\phi = 0.8, n = 100$	1.0115	2.6827	0.81	0.92

3c)

As we can see, the fixed effects estimator is unbiased, however the OLS estimates are severely biased upwards. We can also see the clustered standard errors have higher coverage, especially for higher values of the autoregressive parameter.