Econometrics HW6

Michael B. Nattinger*

March 3, 2021

1 Question 1

1.1 Part i

$\hat{\mu}_{ols}$ can be computed as the average of all observations in the sample:

$$\begin{aligned}\hat{\mu}_{ols} &= \frac{1}{\sum_{i=1}^n T_i}\sum_{i=1}^n\sum_{t=1}^{T_i} Y_{it} \\ &= \frac{\sum_{i=1}^n 1_i' Y_i}{\sum_{i=1}^n 1_i' 1_i}.\end{aligned}$$

1.2 Part ii

We write the estimator as signal plus noise:

$$\begin{aligned}\hat{\mu}_{iv} &= \frac{\sum_{i=1}^n Z_i' Y_i}{\sum_{i=1}^n Z_i' 1_i} \\ &= \frac{\sum_{i=1}^n Z_i'(\mu_0 1_i + \alpha_i 1_i + \epsilon_i)}{\sum_{i=1}^n Z_i' 1_i} \\ &= \mu_0 + \frac{\sum_{i=1}^n Z_i'(\alpha_i 1_i + \epsilon_i)}{\sum_{i=1}^n Z_i' 1_i}.\end{aligned}$$

Now, we can find the variance:

$$\begin{aligned}Var(\hat{\mu}_{iv}) &= Var\left(\frac{\sum_{i=1}^n Z_i'(\alpha_i 1_i + \epsilon_i)}{\sum_{i=1}^n Z_i' 1_i}\right) \\ &= \frac{\sum_{i=1}^n Z_i' Var(\alpha_i 1_i + \epsilon_i) Z_i}{(\sum_{i=1}^n Z_i' 1_i)^2} \\ &= \frac{\sum_{i=1}^n Z_i' \Omega_i Z_i}{(\sum_{i=1}^n Z_i' 1_i)^2},\end{aligned}$$

where

$$\begin{aligned}\Omega_i &= Var(\alpha_i 1_i + \epsilon_i) \\ &= \sigma_\alpha^2 1_i 1_i' + \sigma^2 I_{T_i} \\ &= \sigma_\alpha^2 1_i 1_i' + \sigma^2 I_{T_i}.\end{aligned}$$

*I worked on this assignment with my study group: Alex von Hafften, Andrew Smith, and Ryan Mather. I have also discussed problem(s) with Emily Case, Sarah Bass, Katherine Kwok, and Danny Edgel.

1.3 Part iii

We will apply Cauchy-Schwarz.

$$Z_i'1_i = Z_i'\Omega_i^{1/2}\Omega_i^{-1/2} \leq ||Z_i\Omega_i^{1/2}||||\Omega_i^{-1/2}1_i||,$$

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

We then apply the latter to the formula for variance and get:

$$Var(\hat{\mu}_{iv}) \geq \frac{\sum_{i=1}^{n} Z_i'\Omega_i Z_i}{\sum_{i=1}^{n} Z_i'\Omega_i Z_i \sum_{i=1}^{n} Z_i'\Omega_i^{-1} Z_i}$$

If we use $\tilde{Z}_i = \Omega_i^{-1}1_i$ as an instrument then that has variance:

$$\frac{\sum_{i=1}^{n} \tilde{Z}_i\Omega_i\tilde{Z}_i}{(\sum_{i=1}^{n} \tilde{Z}_i'1_i)^2} = \frac{1}{\sum_{i=1}^{n} 1_i'\Omega_i'1_i}$$

1.4 Part iv

GLS and OLS are the same estimators when the panel is balanced. Therefore, GLS is not more efficient.

$$\Omega_i^{-1} = \frac{1}{\sigma^2} I_{T_i} - \frac{1}{\sigma 4}\frac{\sigma_a^2 I_{T_i}1_i1_i'I_{T_i}}{1+\sigma_a^2 1_i'I_{T_i}1_i/\sigma^2} = \frac{1}{\sigma^2}\left(I_{T_i} - \frac{T_i\sigma_a^2}{T_i\sigma_a^2+\sigma^2}\frac{1_i1_i'}{T_i}\right)$$

$$\tilde{Z}_i = \Omega_i^{-1}1_i = \frac{1_i}{T_i\sigma_a^2+\sigma^2}.$$

Now we impose $T_i = T$:

$$\hat{\mu}_{gls} = \frac{\sum \tilde{Z}_i'Y_i}{\tilde{Z}_i'1_i} = \frac{\sum \frac{1}{T\sigma_a^2+\sigma^2}1_i'Y_i}{\sum \frac{1}{T\sigma_a^2+\sigma^2}1_i'1_i} = \frac{\sum 1_i'Y_i}{\sum 1_i'1_i} = \hat{\mu}_{ols}.$$

1.5 Part v

Let $\bar{\epsilon}_i = \frac{1}{T_i}\sum_{t=1}^{T_i}\epsilon_{it}$.

$$\begin{aligned}
\hat{\sigma}_i^2 &= \frac{1}{T_i-1}\sum_{t=1}^{T_i}(Y_{it}-\bar{Y}_i)^2 \\
&= \frac{1}{T_i-1}\sum_{t=1}^{T_i}(\epsilon_{it}-\bar{\epsilon}_i)^2 \\
&= \frac{1}{T_i-1}\sum_{t=1}^{T_i}\epsilon_{it}(\epsilon_{it}-\bar{\epsilon}_i) \\
&= \frac{1}{T_i-1}\sum_{t=1}^{T_i}\epsilon_{it}^2 - \frac{1}{T_i(T_i-1)}\sum_{t=1}^{T_i}\sum_{s=1}^{T_i}\epsilon_{is}\epsilon_{it} \\
&= \frac{1}{T_i}\sum_{t=1}^{T_i}\epsilon_{it}^2 - \frac{1}{T_i(T_i-1)}\sum_{t=1}^{T_i}\sum_{s=1,s\neq t}^{T_i}\epsilon_{is}\epsilon_{it}.
\end{aligned}$$

Therefore,

$$\begin{aligned} E[\hat{\sigma}_i^2] &= \frac{1}{T_i}\sum_{t=1}^{T_i} E[\epsilon_{it}^2] \\ &= \sigma^2. \end{aligned}$$

The estimator $\hat{\sigma}^2$ is an average of independent random variables, so under very mild assumptions (existence of fourth moment), consistency holds.

1.6 Part vi

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

By the same logic as in part (v), we have the average of uncorrelated random variables as our estimator, so the estimator is consistent under very mild assumptions.

1.7 Part vii

Extending (iv),

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

We can use a plug-in estimator,

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

This will be consistent by the continuous mapping theorem.

2 Question 2

2.1 Part i

$$\begin{aligned}
\hat{\beta}_{FE} &\to_p \beta_0 + \frac{E[\sum_{t=1}^{T}(X_{it} - \bar{X}_i)\epsilon_{it}]}{E[\sum_{t=1}^{T}(X_{it} - \bar{X}_i)^2]} \\
&= \beta_0 + \frac{E[\sum_{t=1}^{T} X_{it}\epsilon_{it}] - E[\sum_{t=1}^{T} \bar{X}_i\epsilon_{it}]}{(T-1)\sigma_x^2} \\
&= \beta_0 + \frac{-E[\sum_{t=1}^{T} \bar{X}_i\epsilon_{it}]}{(T-1)\sigma_x^2} \\
&= \beta_0 + \frac{-E[\sum_{t=1}^{T}\sum_{s=1}^{T} X_{is}\epsilon_{it}]}{T(T-1)\sigma_x^2} \\
&= \beta_0 + \frac{-(T-1)\delta\sigma_x^2}{T(T-1)\sigma_x^2} \\
&= \beta_0 - \frac{\delta}{T}
\end{aligned}$$

Therefore, the asymptotic bias of $\hat{\beta}_{FE}$ is $-\frac{\delta}{T}$.

$$\begin{aligned}
\hat{\beta}_{FD} &\to_p \beta_0 + \frac{E[\sum_{t=2}^{T}(X_{it} - X_{i,t-1})(\epsilon_{it} - \epsilon_{i,t-1})]}{E[\sum_{t=1}^{T}(X_{it} - X_{i,t-1})^2]} \\
&= \beta_0 + \frac{E[\sum_{t=2}^{T} X_{it}\epsilon_{it} - X_{i,t-1}\epsilon_{it} - X_{it}\epsilon_{i,t-1} + X_{i,t-1}\epsilon_{i,t-1}]}{E[\sum_{t=2}^{T} X_{it}^2 - 2X_{it}X_{i,t-1} + X_{i,t-1}^2]} \\
&= \beta_0 + \frac{-(T-1)\delta\sigma_x^2}{2(T-1)\sigma_x^2} \\
&= \beta_0 + \frac{-\delta}{2}
\end{aligned}$$

Therefore, the asymptotic bias of $\hat{\beta}_{FD}$ is $-\frac{\delta}{2}$.

2.2 Part ii

For $T = 2$, the asymptotic biases are the same.

3 Question 3

3.1 Part i

We run the simulation a single time, output is below:

	FE	OLS	δ_2	δ_3	δ_4	EW LB	EW UB	C LB	C UB
$\phi = 0, n = 40$	1.28	2.79	1.1	0.815	0.943	0.685	1.87	0.56	2
$\phi = 0.8, n = 40$	0.775	2.4	1.22	1.14	1.2	0.227	1.32	-0.103	1.65
$\phi = 0, n = 70$	0.527	2.25	1.02	1.31	1.19	0.0732	0.982	0.0256	1.03
$\phi = 0.8, n = 70$	0.441	2.24	1.01	1.08	1.04	0.0546	0.828	-0.15	1.03
$\phi = 0, n = 100$	0.846	2.69	1.13	1.22	1.13	0.508	1.18	0.489	1.2
$\phi = 0.8, n = 100$	1.15	2.94	1.01	0.939	0.879	0.836	1.47	0.663	1.64

3.2 Part ii

We run the simulation 10000 times. The output is below, and code follows at the end of the problem set:

	FE	OLS	EW	Cluster
$\phi = 0, n = 40$	0.999	2.67	0.894	0.92
$\phi = 0.8, n = 40$	1.01	2.68	0.79	0.919
$\phi = 0, n = 70$	1	2.67	0.902	0.938
$\phi = 0.8, n = 70$	1	2.67	0.807	0.935
$\phi = 0, n = 100$	1	2.68	0.9	0.939
$\phi = 0.8, n = 100$	0.996	2.67	0.811	0.938

3.3 Part iii

Our Monte-Carlo results demonstrate that the fixed effects estimation is unbiased, whereas the (incorrectly specified) OLS estimated coefficient is badly upwardly biased. It also demonstrates that the cluster standard errors have a coverage rate much closer to the correct value than the heteroskedasticity robust standard errors. The cluster standard errors are just as robust in the case without autocorrelation in the error term as they are with autocorrelation, whereas the heteroskedasticity robust standard errors lose a lot of coverage in that case.

```
clear; close all; clc
beta0 = 1;
pdelta = [0 1 1 1]';
philist = [0 0.8];
nlist = [40 70 100];
T = 4;
b_ols = zeros(length(nlist),length(philist),5);
b_fe = zeros(length(nlist),length(philist),4);
CIr = zeros(length(nlist),length(philist),2);
CIc = CIr;
scl = norminv(0.975);
for in = 1:length(nlist)
    n = nlist(in);
    for iphi = 1:length(philist)
        phi = philist(iphi);
        [X,Y] = sim_panel(beta0,pdelta,phi,n,T);
        %t = repmat((1:T)',1,n);
        t = repmat(diag([0 1 1 1]),n,1);
        t = t(:,2:4);
        b_ols(in,iphi,:) = ([ones(T*n,1) X(:) t])\Y(:);
        dX = X - mean(X);
        dY = Y - mean(Y);
        dt = t - mean(t);
        XX = ([dX(:) dt]);
        b_fe(in,iphi,:) = XX\dY(:);
        r = dY(:) - XX*reshape(b_fe(in,iphi,:),4,1);
        Vr = inv(XX'*XX)*(XX'*diag(r.^2)*XX)*inv(XX'*XX);
        clustsum = zeros(4,4);
        for i = 1:n
            clustsum = clustsum + XX(T*(i-1)+1:T*i,:)' *r(T*(i-1)+1:T*i,:) * r(T*(i-1)+1:T*i,:)' * XX(T*(i-1)+1:T*i,:)
        end
        Vc = inv(XX'*XX)*clustsum*inv(XX'*XX);
        CIr(in,iphi,1) = b_fe(in,iphi,1) - scl*sqrt(Vr(1));
        CIr(in,iphi,2) = b_fe(in,iphi,1) + scl*sqrt(Vr(1));
        CIc(in,iphi,1) = b_fe(in,iphi,1) - scl*sqrt(Vc(1));
        CIc(in,iphi,2) = b_fe(in,iphi,1) + scl*sqrt(Vc(1));
    end
end
```

```
nsim = 10000;
b_ols_sim = zeros(length(nlist),length(philist),5,nsim);
b_fe_sim = zeros(length(nlist),length(philist),4,nsim);
CRr = zeros(length(nlist),length(philist),nsim);
CRc = CRr;

for ns = 1:nsim
for in = 1:length(nlist)
    n = nlist(in);
    for iphi = 1:length(philist)
        phi = philist(iphi);
        [X,Y] = sim_panel(beta0,pdelta,phi,n,T);
        %t = repmat((1:T)',1,n);
        t = repmat(diag([0 1 1 1]),n,1);
        t = t(:,2:4);
        b_ols_sim(in,iphi,:,ns) = ([ones(T*n,1) X(:) t])\Y(:);
        dX = X - mean(X);
        dY = Y - mean(Y);
        dt = t - mean(t);
        XX = ([dX(:) dt]);
        b_fe_sim(in,iphi,:,ns) = XX\dY(:);
        r = dY(:) - XX*reshape(b_fe_sim(in,iphi,:,ns),4,1);
        Vr = inv(XX'*XX)*(XX'*diag(r.^2)*XX)*inv(XX'*XX);
        clustsum = zeros(4,4);
        for i = 1:n
            clustsum = clustsum + XX(T*(i-1)+1:T*i,:)' *r(T*(i-1)+1:T*i,:) * r(T*(i-1)+1:T*i,:)' * XX(T*(i-1)+1:T*i,:
        end
        Vc = inv(XX'*XX)*clustsum*inv(XX'*XX);
        CIr_sim(1) = b_fe_sim(in,iphi,1,ns) - scl*sqrt(Vr(1));
        CIr_sim(2) = b_fe_sim(in,iphi,1,ns) + scl*sqrt(Vr(1));
        CIc_sim(1) = b_fe_sim(in,iphi,1,ns) - scl*sqrt(Vc(1));
        CIc_sim(2) = b_fe_sim(in,iphi,1,ns) + scl*sqrt(Vc(1));
        CRr(in,iphi,ns) = (beta0>CIr_sim(1))*(beta0<CIr_sim(2));
        CRc(in,iphi,ns) = (beta0>CIc_sim(1))*(beta0<CIc_sim(2));
    end
end
end
CRr = mean(CRr,3);
CRc = mean(CRc,3);
b_ols_sim = median(b_ols_sim,4);
b_fe_sim = median(b_fe_sim,4);
function [X,Y] = sim_panel(beta0,pdelta,phi,n,T)
% generates panel data
X = zeros(T,n);
alpha = normrnd(0,1,1,n);
ind = alpha>0.6;
X(3:4,ind) = 1;
epsilon = zeros(T+1,n);
epsilon(1,:) = normrnd(0,1,1,n);
u = normrnd(0,1,T,n);
for tt = 1:T
    epsilon(tt+1,:) = phi*epsilon(tt,:) + u(tt,:);
end
Y = X*beta0 + repmat(pdelta,1,n) + repmat(alpha,T,1) + epsilon(2:end,:);
```