PS3 Solutions

Jingle Fu

Solution (a).

- **Unit (Firm) Fixed Effects:** Each firm i has intercept α_i. When the number of firms N is large but the time period T is fixed, as N increases, the number of these nuisance parameters increases too. The IPP refers to potential bias in the estimation of the structural parameters when the number of incidental parameters increases with the sample size. Such problem can lead to inconsistency in the estimation of the common parameters, because the estimation error in α_i doesn't vanish as $N \to \infty$.

- **Time Fixed Effects:** In contrast, time fixed effects add only T dummies, the number of parameters associated with time dummies is fixed. If we perform the within transformation as above, the time dummies remain in the model in the form

$$\mathbf{1}\{t = \tau\} - \bar{d}_{it}$$

  and if we use first-differences, for any firm i, we have:

$$\Delta f_t = f_t - f_{t-1}$$

  Thus, the time fixed effects enter the differenced equation as differences, preserving their role in controlling for period-specific shocks. Hence, their estimation does not create an incidental parameters problem.

In short, while unit fixed effects can cause IPP when T is relative small to N, the addition of time fixed effects does not because the number of time dummies remains fixed and is asymptotically negligible.

```
use GMdata.dta, clear
xtset index yr
xtreg ldsal lemp ldnpt ldrst i.yr, fe robust
```

Solution (b).
For consecutive time periods t and $t-1$, we take the difference and have:

$$\Delta ldsal_{it} = \beta_1 \Delta lemp_{it} + \beta_2 \Delta ldnpt_{it} + \beta_3 \Delta ldrst_{it} + (f_t - f_{t-1}) + \Delta u_{it}$$

Table 1: Fixed Eects Model

	(1) FE b/se
log of employment	0.751***
	(0.059)
log of deflated capital	0.053
	(0.043)
log of deflated R&D capital	0.264***
	(0.051)
Observations	2971

Since α_i does not vary over time, it drops out in the differencing. For the dummies, we have:

$$\mathbf{1}\{t=\tau\} - \mathbf{1}\{t-1=\tau\} = \begin{cases} +1 & \text{if } t=\tau, \\ -1 & \text{if } t-1=\tau, \\ 0 & \text{otherwise.} \end{cases}$$

The time effects appear as differences $f_t - f_{t-1}$. Thus, in the first-differenced equation the levels of the year dummies disappear, but their differences remain.

Solution (c).
Adding a time-invariant dummy variable to the regression $d357_i$, we get:

$$ldsal_{it} = \alpha_i + \beta_1 lemp_{it} + \beta_2 ldnpt_{it} + \beta_3 ldrst_{it} + \delta d357_i + \sum_{\tau=1}^{T} f_t \mathbf{1}\{t=\tau\} + u_{it}$$

We take the first-differences of the equation:

$$\Delta \delta d357_i = 0.$$

Thus, in the first-difference model, the term $\delta d357_i$ drops out, and the time dummies are differenced out. The IPP arises from having many firm-specific parameters with a limited number of observations. So,including a time-invariant dummy does not worsen the IPP since it adds only one parameter that is either absorbed (or eliminated in first differences). If we consider Time-Varying Effects of the industry dummy $\delta_t d357_{it}$: taking the first-difference, we have:

$$\Delta(\delta_t d357_{it}) = \delta_t d357_{it} - \delta_{t-1} d357_{i,t-1} = \delta_t - \delta_{t-1}$$

as $\delta d357_{it} = 1$ for all t if the firm is in the computer industry. For firms not in industry

357, the term remains zero. Thus, in the first-difference equation, the industry effect appears as the difference $\delta_t - \delta_{t-1}$, which is time-varying.
Although allowing the industry effect to vary over time introduces several parameters, the number of these parameters is tied to the fixed. And the key issue with IPP is the growth of parameters with N, which does not occur here.

Solution (d).
As we define $\ddot{y}_{it} = y_{it} - \overline{y}_{it}$ and $\ddot{X}_{it} - \overline{X}_{it}$, we have:

$$\hat{\beta}_{FE-W} = \left(\sum_{i,t} \ddot{X}'_{it}\ddot{X}_{it}\right) \sum_{i,t} \ddot{X}_{it}\ddot{y}_{it}$$

$$\hat{\beta}_{RE} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}y$$

Table 2: Fixed Effects Model

	(1) log of deflated sales b/se
log of employment	0.685***
	(0.030)
log of deflated capital	0.180***
	(0.027)
log of deflated R&D capital	0.099***
	(0.027)
Observations	856

Table 3: Random Effects Model

	(1) log of deflated sales b/se
log of employment	0.598***
	(0.025)
log of deflated capital	0.335***
	(0.019)
log of deflated R&D capital	0.065***
	(0.019)
Observations	856

```
use "GMdata_balanced.dta", clear
xtset index yr
gen d357 = (sic3 == 357)
```

```

xtreg ldsal lemp ldnpt ldrst i.yr##i.d357, fe robust
eststo fe_w
esttab fe_w using d1.tex, replace label booktabs title("FE-W Estimator")
    ///
    cells("b(fmt(3)) se(fmt(3))")

xtreg ldsal lemp ldnpt ldrst i.yr##i.d357, re robust
eststo RE
esttab RE using d2.tex, replace label booktabs title("RE Estimator") ///
    cells("b(fmt(3)) se(fmt(3))")
```

Solution (e).
The Hausman test statistics is:

$$H = \left(\hat{\beta}_{FE} - \hat{\beta}_{RE}\right)' \left[A\mathbb{V}[\hat{\beta}_{FE}] - A\mathbb{V}[\hat{\beta}_{RE}]\right]^{-1} \left(\hat{\beta}_{FE} - \hat{\beta}_{RE}\right)$$

```
hausman fe_w re, sigmamore
```

```
Test of H0: Difference in coefficients not systematic

    chi2(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
            =  61.64
Prob > chi2 = 0.0000
(V_b-V_B is not positive definite)
```

Hausman test output yields a chi-square statistic of 61.64 (with p-value= 0.0000). The null hypothesis is strongly rejected, indicating that the RE estimator is inconsistent and that the FE estimator is preferred.

Solution (f).
Let$\theta = \beta_1 + \beta_2$. Under the null hypothesis, we have

$$H_0 : \theta = 1.$$

An estimator for θ is

$$\hat{\theta} = \hat{\beta}_1 + \hat{\beta}_2.$$

Suppose the estimated coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ have the following variance-covariance matrix:

$$V = \begin{pmatrix} \text{Var}(\hat{\beta}_1) & \text{Cov}(\hat{\beta}_1, \hat{\beta}_2) \\ \text{Cov}(\hat{\beta}_1, \hat{\beta}_2) & \text{Var}(\hat{\beta}_2) \end{pmatrix}.$$

Because $\hat{\theta}$ is a sum of $\hat{\beta}_1$ and $\hat{\beta}_2$, by the properties of variance we have:

$$\text{Var}(\hat{\theta}) = \text{Var}(\hat{\beta}_1 + \hat{\beta}_2) = \text{Var}(\hat{\beta}_1) + \text{Var}(\hat{\beta}_2) + 2\ \text{Cov}(\hat{\beta}_1, \hat{\beta}_2).$$

Under $\mathcal{H}_0$, the deviation of $\hat{\theta}$ from its hypothesized value is:

$$\hat{\theta} - 1 = \hat{\beta}_1 + \hat{\beta}_2 - 1.$$

Since, by the Central Limit Theorem, $\hat{\theta}$ is approximately normally distributed in large samples, we can standardize the difference:

$$Z = \frac{\hat{\theta} - 1}{\sqrt{\mathrm{Var}(\hat{\theta})}}.$$

Under H_0, Z is asymptotically standard normal. Squaring this Z statistic gives a chi-square statistic with 1 degree of freedom:

$$\chi^2 = \frac{(\hat{\beta}_1 + \hat{\beta}_2 - 1)^2}{\mathrm{Var}(\hat{\beta}_1) + \mathrm{Var}(\hat{\beta}_2) + 2\,\mathrm{Cov}(\hat{\beta}_1, \hat{\beta}_2)}.$$

This test statistic is compared to the χ^2_1 distribution.

- If χ^2 is large (and the corresponding p-value is small), we reject H_0 and conclude that the sum of β_1 and β_2 is statistically different from 1.
- If χ^2 is not large (and the p-value is large), we do not reject H_0 and there is no evidence against constant returns to scale.

At the conventional 5% level, the p-value (0.0000) is much smaller than 0.05, meaning we reject $\mathcal{H}_0$.

```
. test lemp + ldnpt = 1

 ( 1)  lemp + ldnpt = 1

       F(  1,   633) =   26.74
            Prob > F =    0.0000
```