IB9BMO ECONOMETRICS I

Deadline: 11 January 2016

Warwick ID: 1257635
WBS PhD Finance

Part A

(1) Maximum Likelihood (ML) Estimation

$$Y_t = \alpha + u_t,\ \ for\ t = 1, 2, ...T$$

where the observations are independent and the distribution of Y_t is given by:

$$f(Y_t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\frac{Y^t - \alpha}{\sigma})^2}.$$

(i) Log-likelihood Function, Score Vector and ML Estimator

Likelihood function is:

$$f(Y_1, ..., Y_T|\alpha, \sigma^2) = \prod_{t=1}^{T} f(Y_t|\alpha, \sigma^2) = L(\alpha, \sigma^2)$$

Then the log-likelihood function is:

$$lnL(\alpha, \sigma^2) = \sum_{t=1}^{T} ln[f(Y_t|\alpha, \sigma^2)]$$

Firstly,

$$\begin{aligned} lnL(\alpha, \sigma^2) &= ln(\frac{1}{\sqrt{2\pi\sigma^2}}) + lne^{-\frac{1}{2}(\frac{Y^t - \alpha}{\sigma})^2} \\ &= -\frac{1}{2} ln 2\pi\sigma^2 - \frac{1}{2}(\frac{Y^t - \alpha}{\sigma})^2 \\ &= -\frac{1}{2} ln 2\pi\sigma^2 - \frac{1}{2}(\frac{\alpha + u_t - \alpha}{\sigma})^2 \\ &= -\frac{1}{2} ln 2\pi\sigma^2 - \frac{1}{2}\frac{u_t^2}{\sigma^2} \end{aligned}$$

Therefore,

$$\begin{aligned} lnL(\alpha, \sigma^2) = \sum_{t=1}^{T} ln[f(Y_t|\alpha, \sigma^2)] &= -\frac{T}{2} ln 2\pi\sigma^2 - \frac{1}{2}\sum_{t=1}^{T}\frac{1}{\sigma^2}(Y_t - \alpha)^2 \\ &= -\frac{T}{2} ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\underbrace{u_t^2}_{RSS} \end{aligned}$$

As a result, maximising $lnL(\alpha, \sigma^2)$ w.r.t α corresponds to minimising the RSS (residual sum of squares) $\Rightarrow$ ML estimation is **identical** to the OLS.

Next, the score vector in this case is a 2 by 1 vector with the first order derivatives of log-likelihood function w.r.t. each parameter.

$$
score\,vector:\ g(\theta) = \begin{bmatrix} \dfrac{\partial lnL(\alpha,\sigma^2)}{\partial \alpha} \\ \\ \dfrac{\partial lnL(\alpha,\sigma^2)}{\partial \sigma^2} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma^2}\sum_{t=1}^{T}(Y_t-\alpha) \\ \\ -\dfrac{T}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum_{t=1}^{T}(Y_t-\alpha)^2 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma^2}\sum_{t=1}^{T}u_t \\ \\ -\dfrac{T}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum_{t=1}^{T}u_t^2 \end{bmatrix}
$$

Finally, to derive the ML estimators we need to solve the following maximisation problem:

$$
\max_{\hat{\alpha},\hat{\sigma}^2} \quad -\frac{T}{2}ln2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(Y_t-\alpha)^2
$$

$$
\begin{aligned}
FOC: \quad \frac{\partial lnL(\alpha,\sigma^2)}{\partial \alpha}\Big|_{\alpha=\hat{\alpha}} &= -\frac{1}{2\sigma^2}2\sum_{t=1}^{T}(Y_t-\hat{\alpha})(-1) = 0 \\
&\Rightarrow \sum_{t=1}^{T}(Y_t-\hat{\alpha}) = 0 \\
&\Rightarrow \frac{1}{T}\sum_{t=1}^{T}(Y_t-\hat{\alpha}) = 0 \\
&\therefore\ \hat{\alpha}_{ML} = \frac{1}{T}\sum_{t=1}^{T}Y_t \equiv \bar{Y}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial lnL(\alpha,\sigma^2)}{\partial \sigma^2}\Big|_{\sigma^2=\hat{\sigma}^2} &= -\frac{T}{2}\frac{2\pi}{2\pi\hat{\sigma}^2} - \frac{1}{2}\frac{1}{(\hat{\sigma}^2)^2}(-1)\sum_{t=1}^{T}\hat{u}_t^2 = 0 \\
&\Rightarrow -\frac{T}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\sum_{t=1}^{T}\hat{u}_t^2 = 0 \\
&\Rightarrow \frac{T}{\cancel{2}}\frac{1}{\hat{\sigma}^2} = \frac{\sum_{t=1}^{T}\hat{u}_t^2}{\cancel{2}\hat{\sigma}^4} \\
&\therefore\ \hat{\sigma}^2_{ML} = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2
\end{aligned}
$$

Hence, ML estimator of variance $\hat{\sigma}^2_{ML}$ is a *consistent* estimator, but it is not *unbiased.*

(ii) Information Matrix and Asymptotic Covariance Matrix

Information matrix

$$I(\theta_0) = -E[H(\theta_0)],$$

where $H(\theta_0)$ is the Hessian matrix for log-likelihood function $lnL(\alpha, \sigma^2)$.

Calculate the Hessian:

$$\begin{aligned}
from\ (i): \qquad \frac{\partial lnL(\alpha,\sigma^2)}{\partial \alpha} &= \frac{1}{\sigma^2}\sum_{t=1}^{T}(Y_t-\alpha) = \frac{1}{\sigma^2}\sum_{t=1}^{T}Y_t - \frac{T}{\sigma^2}\alpha \\
\therefore\ \frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial \alpha^2} &= -\frac{T}{\sigma^2} \\
\frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial \alpha \partial \sigma^2} &= -\frac{1}{\sigma^4}\sum_{t=1}^{T}(Y_t-\alpha) = -\frac{1}{\sigma^4}\sum_{t=1}^{T}u_t
\end{aligned}$$

$$\begin{aligned}
and \qquad \frac{\partial lnL(\alpha,\sigma^2)}{\partial \sigma^2} &= -\frac{T}{2}\frac{1}{\sigma^2} + \frac{1}{2}\frac{1}{\sigma^4}\sum_{t=1}^{T}(Y_t-\alpha)^2 \\
\therefore\ \frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial (\sigma^2)^2} &= \frac{T}{2}\frac{1}{\sigma^4} + \frac{1}{2}(-2)\frac{1}{(\sigma^2)^3}\sum_{t=1}^{T}u_t^2 \\
&= \frac{T}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{t=1}^{T}u_t^2
\end{aligned}$$

$$\therefore I(\theta_0) = -E[H(\alpha,\sigma^2)] = -E\begin{bmatrix} \frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial \alpha^2} & \frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial \alpha \partial \sigma^2} \\ \frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial \alpha \partial \sigma^2} & \frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial (\sigma^2)^2} \end{bmatrix}$$

Taking expectation of each element in the matrix:

$$
\begin{aligned}
-E\Big[\frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial\alpha^2}\Big] &= -E\Big[-\frac{T}{\sigma^2}\Big] = \frac{T}{\sigma^2} \\
-E\Big[\frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial\alpha\partial\sigma^2}\Big] &= -E\Big[-\frac{1}{\sigma^4}\sum_{t=1}^{T} u_t\Big] = \frac{1}{\sigma^4}\sum_{t=1}^{T}\underbrace{E(u_t)}_{=0} = 0 \\
-E\Big[\frac{\partial^2 lnL(\alpha,\sigma^2)}{\partial(\sigma^2)^2}\Big] &= -E\Big[\frac{T}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{t=1}^{T} u_t^2\Big] \\
&= -\frac{T}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{t=1}^{T}\underbrace{E(u_t^2)}_{=\sigma^2} \\
&= -\frac{T}{2\sigma^4} + \frac{1}{\sigma^6}T\sigma^2 \\
&= -\frac{T}{2\sigma^4} + \frac{T}{\sigma^4} \\
&= \frac{T}{2\sigma^4}
\end{aligned}
$$

$\therefore$ Information matrix is:

$$
I(\alpha,\sigma^2) = \begin{bmatrix} \frac{T}{\sigma^2} & 0 \\ 0 & \frac{T}{2\sigma^4} \end{bmatrix}
$$

The asymptotic variance-covariance matrix is the *inverse* of information matrix:

$$
var(\theta_0) = var((\alpha,\sigma^2)) = [I(\alpha,\sigma^2)]^{-1} = \begin{bmatrix} \frac{\sigma^2}{T} & 0 \\ 0 & \frac{2\sigma^4}{T} \end{bmatrix}
$$

(iii) Estimate the Standard Error

The standard error of $\hat{\alpha}_{ML}$ is

$$
\sqrt{[I(\alpha,\sigma^2)]_{11}^{-1}} = \sqrt{\frac{\sigma^2}{T}}
$$

where σ^2 can be estimated by

$$
\hat{\sigma}^2_{ML} = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2
$$

where $\hat{u}_t = Y_t - \hat{\alpha}_{ML}$.

(2) Regression Model with Endogenous Regressors

Consider the linear regression model

$$y_t = \beta' x_t + u_t, \qquad t = 1, 2, ..., T$$

If the correlation between x_t and u_t is not zero, it means that the exogeneity assumption of standard OLS regression fails to hold:

$$E(x_t u_t) \neq 0$$

which is called the issue of *endogeneity*. This will make the OLS estimator **inconsistent**. As a result, the OLS estimator is invalid since it is biased/incorrect even if the sample size goes to infinity.

The modelling situations in which one might expect such a correlation to arise, or the issue of endogeneity to occur, are discussed in the next section.

Cases of Endogeneity

There are many possible situations in which one might encounter an endogeneity problem. According to Greene (2012), some of the most common reasons are:

1. *Measurement Error* :
   Consider

   $$Y = \beta_0 + \beta_1 X^* + \varepsilon$$

   where the classical assumptions are satisfied (and X^* is unobserved), hence:

   $$E[\varepsilon | X^*] = 0 \ \Rightarrow \ E[Y|X^*] = \beta_0 + \beta_1 X^*.$$

   Under this case,

   - $E[\varepsilon] = 0$ and $Corr(X^*, \varepsilon) = 0$
   - $\beta_0 = E[Y] - \beta_1 E[X^*]$ and $\beta_1 = \dfrac{Corr(Y, X^*)}{Var(X^*)}$

   However, X^* has measurement error. We observe X such that $X = X^* + v_1$, where v_1 is the measurement error.

   Substituting $X^* = X - v_1$, we get

   $$Y = \beta_0 + \beta_1 X + \underbrace{\varepsilon - \beta_1 v_1}_{u}$$

As a result, $Corr(X, u) \neq 0 \;\Rightarrow\; X$ is endogenous.

2. *Nonrandom Sampling and Attrition* :
Nonrandom sampling: In a model of the effect of a training program, an employment program, or the labour supply behaviour of a particular segment of the labour force, the sample may have voluntarily selected themselves into the observed sample. For instance, the Job Training Partnership Act (JTPA) was a job training program intended to provide employment assistance to disadvantaged youth. However, some researchers found that for a sample that they examined, the program appeared to be administered most frequently to the best qualified applicants. In an earnings equation estimated for such a nonrandom sample, the implication is that the disturbances are not truly random. This nonrandomness of the sample translates to a form of omitted variable bias known as **sample selection bias**.

   Attrition: There are two closely related important cases of nonrandom sampling. In panel data studies of firm performance, the firms still remain in the sample at the end of the observation period are probably a subset of those present at the start of the period - those firms that underperform may simply drop out of the sample. In these situations, least squares regression of the performance variable on the covariates (whatever they are), suffers from a form of selection bias known as **survivorship bias**. In this case, the distribution of outcomes, firm performances, for the survivors is systematically higher than that for the population of firms as a whole. This produces a phenomenon known as **truncation bias**.

3. *Omitted Variables* :
In a 'long regression' (1st line) setting with cross-section data,
$$\begin{aligned} y_t &= \beta' x_t + \gamma' w_t + u_t \\ &\equiv \beta' x_t + \varepsilon_t \end{aligned}$$
with $\varepsilon_t \equiv w_t \gamma + u_t$, we may only observe y_t and x_t, and only care about the estimation of β. However, if x_t and the "missing regressors" w_t are correlated with x_t in the population, x_t and ε_t will be correlated in the 'short regression' (2nd line) of y_t on x_t.

4. *Simultaneity* :
Suppose that we have two structural equations,
$$\begin{aligned} y_t &= \beta_1' x_t + \gamma_1' z_t + u_t \\ z_t &= \beta_2' x_t + \gamma_2' y_t + v_t \end{aligned}$$

   Regarding the first structural equation, it can be shown that $E(z_t u_t) \neq 0$. First, solving for z_t we can get (assuming that $1 - \gamma_1 \gamma_2 \neq 0$),

$$z_t = \frac{\beta_2 + \gamma_2\beta_1}{1-\gamma_1\gamma_2}x_t + \frac{1}{1-\gamma_1\gamma_2}v_t + \frac{\gamma_2}{1-\gamma_1\gamma_2}u_t.$$

Assuming that x_t and v_t are uncorrelated with u_t, we find that,

$$E(z_t u_t) = \frac{\gamma_2}{1-\gamma_1\gamma_2}E(u_t^2) \neq 0$$

Therefore, attempts to estimating either structural equation will be contaminated by endogeneity.

Alternative Estimators

In order to remedy the endogeneity problems, many alternative estimators with better properties have been proposed. Some of these estimators are discussed here.

1. *Two Stage Least Squares (2SLS) Estimator* :
   Consider the linear model with k explanatory variables x_t

   $$y_t = x_t\beta + u_t$$

   Now suppose some of the x_t variables are endogenous such that

   $$E(x_t u_t) \neq 0$$

   Let z_t denote r instrumental variables that are correlated with x_t but uncorrelated with u_t

   $$E(z_t x_t) \neq 0 \quad and \quad E(z_t u_t) = 0$$

   Apply the $2SLS$ method:

   - First Stage: apply OLS to estimate

     $$x_t = z_t\Gamma_k + \varepsilon_t$$

     We obtain

     $$\hat{\Gamma} = (Z'Z)^{-1}Z'X \quad and \quad \hat{X} = Z\hat{\Gamma}$$

   - Second Stage: apply OLS to the equation

     $$y_t = \hat{x}_t\beta + v_t$$

     we can obtain a consistent estimator for β.

2. *Generalised Method of Moments (GMM) Estimator* :
GMM moves away from parametric assumptions towards estimators that are robust to variations in the underlying DGP. The idea is to estimate parameters by matching sample moments to population moments.

Using the previous example,

$$y_t = x_t\beta + u_t \quad and \quad E(z_t u_t) = 0$$

We obtain r moment/orthogonality conditions

$$E(z_t(y_t - x_t\beta)) = 0$$

In this case the GMM estimator coincides with $2SLS$ estimator. In other words, $2SLS$ estimator is a special case of GMM estimator here, which generates consistent estimates.

(3) Kalman Filter

Consider the time-varying parameter regression model:

$$\underset{1\times 1}{y_t} = \underset{1\times k}{\beta_t'}\ \underset{k\times 1}{x_t} + \underset{1\times 1}{u_t}, \qquad t = 1, 2, ..., T,\ u_t \sim N(0, \sigma_u^2)$$
$$\underset{k\times 1}{\beta_t} = \beta_0 + \underset{k\times k}{\rho}\ \underset{k\times 1}{\beta_{t-1}} + \underset{k\times 1}{\varepsilon_t},\ \varepsilon_t \sim N(0, \underset{k\times k}{\Sigma})$$

(i) State-space Representation

The observation/measurement equations:

$$y_t = \beta_t' x_t + u_t, \quad t = 1, 2, ..., T,\ u_t \sim N(0, \sigma_u^2)$$

The state/transition equations:

$$\beta_t = \beta_0 + \rho\beta_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma)$$

(ii) Kalman Filter and Likelihood Function

Under the assumption that β_0 is Gaussian and both u_t and ε_t are Gaussian, then

$$y_t|\Upsilon^{t-1} \sim N(y_{t|t-1}, f_{t|t-1})$$

and the log-likelihood (over all t) is:

$$-\frac{1}{2}\sum_{t=1}^{T} ln(2\pi f_{t|t-1}) - \frac{1}{2}\sum_{t=1}^{T} \eta_{t|t-1}' f_{t|t-1}^{-1} \eta_{t|t-1}$$

Hence, we need to specify three elements: $y_{t|t-1},\ \eta_{t|t-1},\ f_{t|t-1}$.

(1) Start with the best prediction for β_t with information available up to $t-1$:

$$\beta_{t|t-1} = \beta_0 + \rho\beta_{t-1|t-1}$$

and the variance of $\beta_{t|t-1}$ is:

$$P_{t|t-1} = \rho P_{t-1|t-1}\rho' + \Sigma$$

(2) At t,, y_t is observed so we can compute the prediction error $\eta_{t|t-1}$:

$$\begin{aligned} y_{t|t-1} &= \beta'_{t|t-1}x_t \\ \eta_{t|t-1} &= y_t - y_{t|t-1} = y_t - \beta'_{t|t-1}x_t \end{aligned}$$

and the conditional variance of the prediction error is:

$$f_{t|t-1} = x'_t P_{t|t-1} x_t + \sigma_u^2$$

(3) The Kalman filter updating/filtering equations are:

$$\begin{aligned} \beta_{t|t} &= \beta_{t|t-1} + \underbrace{P_{t|t-1}x_t f_{t|t-1}^{-1}}_{\equiv K_t} \eta_{t|t-1} \\ &= \beta_{t|t-1} + K_t \eta_{t|t-1} \end{aligned}$$

$$\begin{aligned} P_{t|t} &= P_{t|t-1} - \underbrace{P_{t|t-1}x_t f_{t|t-1}^{-1}}_{\equiv K_t} x'_t P_{t|t-1} \\ &= P_{t|t-1} - K_t x'_t P_{t|t-1} \end{aligned}$$

where K_t is the *Kalman gain* and guides how much new information is included in the prediction error about the state variable β.

(4) Initial values:

$$\begin{aligned} \beta_{0|0} &= \beta_0 + \rho\beta_{0|0} \\ \therefore\ \beta_{0|0} &= \frac{\beta_0}{1-\rho} \end{aligned}$$

$$\begin{aligned} P_{0|0} &= \rho P_{0|0}\rho' + \Sigma \\ \therefore vec(P_{0|0}) &= (I - \rho \otimes \rho)^{-1} vec(\Sigma) \end{aligned}$$

(iii) Kalman Smoother

The smoothing equations are:

$$\begin{aligned} \beta_{t|T} &= \beta_{t|t} + P_{t|t}\rho' P_{t+1|t}^{-1}(\beta_{t+1|T} - \rho\beta_{t|t} - \beta_0) \\ P_{t|T} &= P_{t|t} + P_{t|t}\rho' P_{t+1|t}^{-1}(P_{t+1|T} - P_{t+1|t}) P_{t+1|t}^{-1'} \rho P'_{t|t} \ , \end{aligned}$$

the initial values are the last filtered values, $\beta_{T|T}$ and $P_{T|T}$.

(4) Structural Break

(a) Chow Breakpoint Test

Apply the Chow's breakpoint test:

$$\frac{[RSS_{full} - (RSS_{sub1} + RSS_{sub2})]/k}{(RSS_{sub1} + RSS_{sub2})/(T-2k)} \sim F(k, T-2k)$$

where k is the number of regressors in the model,
RSS_{full} is the residual sum of squares of the full sample (79Q4-03Q4),
RSS_{sub1} is the residual sum of squares of the sub-sample 1 (79Q4-92Q3), and
RSS_{sub2} is the residual sum of squares of the sub-sample 2 (92Q3-03Q4).

H_0 : regression parameters are the *same* in both periods (assuming *constant variance* across two sub-periods).

- $k = 4, \quad T = 97$
- $RSS_{full} = 0.288$
- $RSS_{sub1} = 0.226, RSS_{sub2} = 0.03253$

$$\begin{aligned} \therefore \ test\ statistics &= \frac{[0.288 - (0.226 + 0.03253)]/4}{(0.226 + 0.03253)/(97 - 2 \times 4)} \\ &= \frac{0.0073675}{0.0029048} \\ &= 2.54 \quad > \quad F(4, 89) = 2.47 \end{aligned}$$

$\Rightarrow$ **Reject** the null under 5 % significance level.
$\rightarrow$ parameters are NOT the same in two sub-periods.

(b) Variance Ratio/Goldfeld-Quandt Test

Apply the Variance Ratio/Goldfeld-Quandt test:

$$\frac{RSS_{sub1}/(T_1 - k)}{RSS_{sub2}/(T_2 - k)} \sim F(T_1 - k, T_2 - k)$$

where k is the number of regressors in the model,
T_1 is the sample size of sub-sample 1 (79Q4-92Q3), and
T_2 is the sample size of sub-sample 2 (92Q3-03Q4).

H_0 : variances across the two sub-periods are the *same*.

- $k = 4$
- $T_1 = 52,\ T_2 = 45$
- $RSS_{sub1} = 0.226, RSS_{sub2} = 0.03253$

$$\begin{aligned} test\ statistics &= \frac{0.226/(52-4)}{0.03253/(45-4)} \\ &= 5.93 \quad > \quad F(48, 41) = 1.64 \end{aligned}$$

⇒ **Reject** the null that variances in two sub-periods are the same under 5 % significance level.

(c) From (b) to (a)

Chow's breakpoint test assumes that the variances across two sub-periods are the same, which has been **rejected** by results from (b):

→ the result from (a) could be incorrect and misleading.

(d) Evidence for Misspecification

The test for serial correlation (*SC*, with a p-value of 0.017) and test for heteroskedasticity (*Het*, with a p-value of 0.025) for the full sample regression are both **rejected** under 5 % significance level. This indicates that the full sample model may be subject to the problem of misspecification.

However, the SC and Het tests for two sub-samples are all **insignificant**, suggesting no misspecification issue for each sub-sample model.

Overall, it is concluded that the problem of misspecification for the full sample model may result from the structural break in regression parameters across the two sub-samples.

(e) Comparison betweem Estimates for the Two Subperiods

Firstly, the t tests of statistical significance from zero are conducted for parameter estimates of two sub-sample models and summarised in the following table:

	constant	Δy_t^*	r_{t-1}	Δy_{t-1}
sub1 79Q4-92Q3	**0.020**	**0.459**	**-0.670**	0.071
sub2 92Q3-03Q4	0.000	**0.585**	0.239	-0.006

Note: **boldface** indicates statistical significance from 0

It is worth noting from the table that:

- Lagged growth, Δy_{t-1}, has **no** explanatory power for both sub-periods;
- The constant over sub-period 1 is **significant** at 0.02 or 2 %, while that over the sub-period 2 is insignificantly different from 0;
- Lagged long term interest rate, r_{t-1}, has a negative and **significant** effect on the growth rate of GDP, Δy_t, over sub-period 1, whereas it has a positive but insignificant effect on Δy_t over sub-period 2;
- Last but not the least, the growth rate of trade weighted foreign GDP, Δy_t^*, has a **significant** impact on Δy_t in both sub-period 1 and 2.

In summary, there is a change in s of the growth rate of GDP, Δy_t, from sub-period 1 to sub-period 2. Although the growth rate of trade weighted foreign GDP, Δy_t^*, remain a significant explanatory power, the lagged long term interest rate, r_{t-1}, ceases explaining the dynamics of Δy_t. Moreover, there is also a change in long term average level of Δy_t across two sub-periods, indicating by the magnitude and (in)significance of the constant term.

Part B

Consider the following dynamic panel data model:

$$\Delta s_{i,t+1} = \alpha_i + \beta_i z_{it} + \epsilon_{it+1},\ \epsilon_{it+1} \sim iid\ Normal$$

where

- $\Delta s_{i,t+1} = s_{i,t+1} - s_{it}$;
- s_{it} = natural log of the nominal exchange rate measured in terms of domestic currency (the US dollar) per unit of foreign currency;
- $z_{it} \equiv (p_{0t} - p_{it}) - s_{it}$ = deviation from an equilibrium determined accoring to PPP, and measure of real exchange rate;
- p_{0t} = the log of US price levels;
- p_{it} = the log of foreign price levels;
- i = country subscript, $i = 1, 2, ..., N$; $N = 9$;
- t = time subscript, $t = 1, 2, ..., T - 1$; $T = 156$.

In this section, this particular panel data model regarding Purchasing Power Parity (PPP) is examined. The model has been estimated using different estimators that lead to different estimation results. Results from different estimators are compared, contrasted, commented and analysed from both an econometric and economic perspective.

Section (1) provides regression results from Pooled OLS (POLS), separate-country OLS and mean group (MG) estimators. Secion (2) shows results from fixed effects (FE) and random effects models, as well as results of a likelihood ratio (LR) test for equality of parameters across countries and a Hausman test for RE (null) v.s. FE (alternative) model. Section (3) compares the regression results from section (1) and those from the FE model in section (2).

1. Pooled OLS, Separate OLS and Mean Group Estimators

(a) Pooled OLS (POLS) Estimator

Country	Coefficient	Estimate	StdErr	SER	Adj.R^2	ADF	ARCH	DW	NT
Across All	Constant	0.0014	(0.0010)						1395
	β_P	-0.0003	(0.0003)	0.0285	0.0007	0.1792	**0.0000**	**0.0047**	

(b) Separate OLS Estimator for Each Country

Country	Coefficient	Estimate	StdErr	SER	Adj.R^2	ADF	ARCH	DW	T
Canada	Constant	0.0052	(0.0045)						155
	β_{OLS}	-0.0098	(0.0142)	0.0192	0.0107	0.3069	**0.0206**	0.6001	
Denmark	Constant	-0.0192	(0.0293)						155
	β_{OLS}	0.0109	(0.0157)	0.0263	-0.0019	0.5200	0.4039	**0.0473**	
Euro	Constant	0.0020	(0.0026)						155
	β_{OLS}	0.0067	(0.0128)	0.0221	-0.0026	0.5877	0.5183	**0.0000**	
Japan	Constant	-0.1489	(0.0797)						155
	β_{OLS}	0.0320	(0.0172)	0.0317	0.0165	0.9239	0.1001	0.3240	
Korea	Constant	-0.2577	(0.1562)						155
	β_{OLS}	0.0367	(0.0223)	0.0449	0.0116	0.6723	**0.0011**	0.1350	
Norway	Constant	-0.0277	(0.0355)						155
	β_{OLS}	0.0149	(0.0183)	0.0276	0.0000	0.5137	0.5678	0.9252	
Sweden	Constant	-0.0292	(0.0305)						155
	β_{OLS}	0.0150	(0.0151)	0.0278	0.0011	0.6315	0.3667	0.5263	
Switz	Constant	-0.0060	(0.0049)						155
	β_{OLS}	0.0271	(0.0172)	0.0286	0.0102	0.4557	0.5961	0.7738	
UK	Constant	0.0159	(0.0116)						155
	β_{OLS}	0.0262	(0.0209)	0.0206	0.0089	0.7980	0.3654	0.4784	

(c) Mean Group (MG) Estimator

Country	Coefficient	Estimate	StdErr
Across All	Constant	-0.0517	(0.0305)
	β_{MG}	**0.0177***	(0.0048)

Note: 1. * denotes statistical significance under 95% confidence level;
2. Standard errors (StdErr) are in parenthese:
3. SER is the the standard error of regression;
4. ADF, ARCH, and DW are *p-values* for Augmented Dickey-Fuller test, ARCH type heteroskedasticity test, and Durbin-Watson autocorrelation test, respectively.

Figure 1: table for results of POLS, OLS and MG Estimators (Balanced Panel)

POLS Estimator

Figure 3 panel (a) shows the results using Pooled OLS estimator, which is the estimator in the POLS regression

$$\Delta s_{i,t+1} = \alpha + \beta_P z_{it} + \varepsilon_{it+1}.$$

This least squares regression stacks the cross-section or individuals with the time series such that the regression y_{it} on x_{it} using $NT \times 1$ observations for balanced panel, or $(\sum_{i=1}^{N} T_i) \times 1$ observations for unbalanced panel. By stacking all the time series together, the POLS estimator leads to a general β_P that is applied to all cross sectional countries.

A few points are worth emphasising:

- Both $\hat{\alpha}$ and $\hat{\beta}_P$ are **insignificantly** different from zero;
- *Adjusted* R^2 is almost zero (0.0007), implying that z_{it} plays **little role** in explaining the variations in $\Delta s_{i,t+1}$;
- Diagnostic test on z_{it}: ADF reports p-value of the ADF test on z_{it} (not on residuals). The p-value of 0.1792 indicates failure of rejection to the null of a unit root, suggesting that the model's explanatory variable z_{it} is **nonstationary** (while in the question it is *assumed* to be stationary). This implies that the model results are strongly biased and probably invalid;
- Diagnostic test on residuals: ARCH and DW report p-values of the Engle's ARCH test for residual heteroskedasticity and Durbin-Watson test for 1st-order serial correlation, respectively. The null of ARCH test (of homoskedasticity) and that of DW test (of no serial correlation) are both **rejected under 1 % significance level**, indicating existence of serial correlation and heteroskedasticity in regression residuals. This also implies that the standard errors for parameter estimates are biased/incorrect and hypothesis testing results are misleading.

Separate OLS Estimators

Figure 1 panel (b) displays results applying standard OLS estimator to each of the 9 countries. As a result, regression outcomes from each country are obtained.

Results are summarised as follows:

- $\hat{\alpha}_{OLS,i}$ and $\hat{\beta}_{OLS,i}$ $(i = 1, 2, ..., N)$ for each country are **insignificantly** different from zero;

- *Adjusted* R_i^2 for each country varies. *Adjusted* R_i^2 for Canada, Japan, Korea, Sweden, Switzerland and UK are higher than that for POLS regression. This suggests that, although the z_{it} still plays **little role** in explaining the variations in $\Delta s_{i,t+1}$ due to the **nonstationarity** of z_{it} (which is to be re-emphasised below), estimating each country separately is a better practice in general than estimating the pooled regression;

- Diagnostic test on z_{it}: the p-value of ADF test for each country indicates the failure of rejection to the null of a unit root, suggesting that the model's explanatory variable z_{it} is **nonstationary**, which explains why the estimators and models behave poorly;

- Diagnostic test on residuals: the p-value of ARCH and DW tests for each country generate better results, on average, than those for the pooled regression. However, these results do not alter the bad performance of the models due to the built-in problem of **nonstationarity**.

MG Estimator

The Mean Group (MG) estimates are computed as follows:

$$\hat{\alpha}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_{OLS,i}$$
$$\hat{\beta}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_{OLS,i}$$

The MG estimator's covariance matrix is estimated as

$$V\Big(\begin{bmatrix}\hat{\alpha}_{MG}\\ \hat{\beta}_{MG}\end{bmatrix}\Big) = \frac{1}{N(N-1)}\sum_{i=1}^{N}\Big(\begin{bmatrix}\hat{\alpha}_{OLS,i}\\ \hat{\beta}_{OLS,i}\end{bmatrix} - \begin{bmatrix}\hat{\alpha}_{MG}\\ \hat{\beta}_{MG}\end{bmatrix}\Big)\Big(\begin{bmatrix}\hat{\alpha}_{OLS,i}\\ \hat{\beta}_{OLS,i}\end{bmatrix} - \begin{bmatrix}\hat{\alpha}_{MG}\\ \hat{\beta}_{MG}\end{bmatrix}\Big)',$$

and the standard errors are

$$stderr\ \hat{\alpha}_{MG} = \sqrt{V_{11}}$$
$$stderr\ \hat{\beta}_{MG} = \sqrt{V_{22}}$$

Figure 1 panel (c) reports coefficient estimates and associated standard errors for MG estimator. It should be noted that the standard error of $\hat{\beta}_{MG}$ is **substantially**

lower (i.e. more reliable estimate) than its separate OLS counterparts. This leads to **significance** of the slope coefficient, although this significance may carry little meaning because of (1)misspecification (i.e. omitted dynamics) of model, (2) non-stationarity of explanatory variable and (3) serial correlation and heteroskedasticity of residuals.

Summary

A few points are worth recapitulating to highlight the differences among estimators:

- The POLS estimator for the dynamic panel data model in question has the worst overall performance compared to OLS and MG estimators. This is probably because there is a heterogeneity in slope coefficients across different countries (a hypothesis test of (in)equality of slope coefficient can be performed, which is not offered here);

- When we have large enough sample sizes (i.e. large T panel), it is usually preferred to performing separate OLS regression for each cross sectional unit so that the heterogeneity problem can be mitigated;

- The MG estimator tends to outperform among the 3 estimators. Moreover, the covariance matrix of MG estimator is a non-parameteric estimator of the standard errors of $\hat{\alpha}_{MG}$ and $\hat{\beta}_{MG}$, which does not depend on estimation of $V([\hat{\alpha}_{OLS,i}, \hat{\beta}_{OLS,i}]')$ that may not be robust to autocorrelation and/or heteroskedasticity.

2. FE Models, RE Models and Hypothesis Testing

(2.1) Fixed Effects (FE) Model

Country	FE	β_{FE}	StdErr (β_{FE})	SER	Adj.R^2	ADF	ARCH	DW	NT
Canada	-0.0034								
Denmark	-0.0349								
Euro	0.0040								
Japan	-0.0903								
Korea	-0.1364	**0.0193***	(0.0058)	0.0285	0.0036	0.1792	**0.0000**	**0.0016**	1395
Norway	-0.0362								
Sweden	-0.0379								
Switz	-0.0041								
UK	0.0121								

Variance Across Countries

Canada	Denmark	Euro	Japan	Korea	Norway	Sweden	Switz	UK
0.0004	0.0007	0.0005	0.0010	0.0020	0.0008	0.0008	0.0008	0.0004

LR Test

Test Statistic	Critical Value	Test Result
15.3703	15.5073	Fail to Reject the Null of Constant Beta Across Units Under 95% CL

(2.2) Random Effects (RE) Model

Country	Coefficient	Estimate	StdErr	SER	Adj.R^2	ADF	ARCH	DW
Across All	Constant	0.0011	(0.0010)					
	β_{RE}	-0.0003	(0.0004)	0.0285	0.0002	0.1797	**0.0000**	**0.0048**

(2.3) Hausman Test

Test Statistic	Critical Value	Test Result
11.5931	3.8415	Reject the Null Under 95% CL, FE Model Is Appropriate

Note: 1. * denotes statistical significance under 95% confidence level;
2. Standard errors (StdErr) are in parenthese:
3. SER is the the standard error of regression;
4. ADF, ARCH, and DW are *p-values* for Augmented Dickey-Fuller test, ARCH type heteroskedasticity test, and Durbin-Watson autocorrelation test, respectively.

Figure 2: table for results of FE and RE Estimators, LR and Hausman Tests (Balanced Panel)

Fixed Effects (FE) Model

Figure 2 panel (2.1) provides estimation outcomes for the one-way FE model, which is estimated by OLS using *within group* estimator:

$$(\Delta s_{i,t+1} - \Delta \bar{s}_i) = \beta_{FE,W}(z_{it} - \bar{z}_i) + u_{it+1}$$

where

$$\Delta \bar{s}_i = \frac{1}{T-1}\sum_{t=1}^{T-1} \Delta s_{i,t+1}, \quad \bar{z}_i = \frac{1}{T-1}\sum_{t=1}^{T-1} z_{it}.$$

The N intercepts (i.e. fixed effects) are then estimated as

$$\hat{\alpha}_i = \Delta \bar{s}_i - \hat{\beta}_{FE,W}\bar{z}_i.$$

It is shown in the table that the fixed effect (i.e. intercept term) varies across countries. Moreover, a general slope coefficient $\beta_{FE,W}$ is applied to every country. We now turn to test on the null hypothesis of a constant slope coefficient across countries in the next section.

Likelihood Ratio (LR) Test

The hypothesis of equality of slopes is tested, however, conditional on equality of residual variances (i.e. $\sigma_i^2 = \sigma^2$). Although the assumption of equal variances is not tested, it is likely to be reasonable. This can be justified in the section 'Variance Across Countries' in figure 2, where 'variance' is the variance of residuals of separate OLS regression for each country. It turns out that residual variances do not vary a lot across countries.

The LR test statistic is

$$LR = 2(MLL_{OLS} - MLL_{FE}) \ \sim \ \chi^2(N-1)k$$

where $k = 1$ and MLL_{OLS} of the homogeneous variance model is

$$\begin{aligned} MLL_{OLS} &= -\frac{N(T-1)}{2}[ln(2\pi)+1] - \frac{N(T-1)}{2}ln\hat{\sigma}^2 \\ \hat{\sigma}^2 &= \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\hat{\epsilon}_{it+1}^2 \\ \hat{\epsilon}_{it+1} &= \Delta s_{i,t+1} - \hat{\alpha}_i - \hat{\beta}_i z_{it}, \ t = 1,2,...,T-1, \ i = 1,2,...,N, \end{aligned}$$

and MLL_{FE} of the one-way FE model is

$$
\begin{aligned}
MLL_{FE} &= -\frac{N(T-1)}{2}[ln(2\pi)+1] - \frac{N(T-1)}{2} lns^2 \\
s^2 &= \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=1}^{T-1} \hat{u}_{it+1} \\
\hat{u}_{it+1} &= (\Delta s_{i,t+1} - \Delta \bar{s}_i) - \hat{\beta}_{FE,W}(z_{it} - \bar{z}_i),\ t = 1, ..., T-1,\ i = 1, ..., N.
\end{aligned}
$$

From Figure 2, the LR test statistic is 15.37 < 15.51 (critical value 5%). Hence, we **fail to reject** the null of slope equality (conditional on the homogeneity of variances) under 5 % significance level.

Random Effects (RE) Model

In order to perform Hausman test (which is discussed in the next section), a RE model is required to be estimated since RE estimate of slope coefficient and the associated variance of RE estimate are ingredients for Hausman test.

Estimation of RE model is performed by Generalised Least Squares (GLS), which is equal to OLS on

$$(\Delta s_{i,t+1} - \theta \Delta \bar{s}_i) = \alpha + \beta_{RE}(z_{it} - \theta \bar{x}_i) + e_{it+1}$$

where

$$
\begin{aligned}
\theta &= 1 - \sqrt{\frac{\sigma^2}{(T-1)\sigma_\eta^2 + \sigma^2}} \\
\sigma^2 &= \frac{RSS_{POLS}}{N(T-1)-1-k} \\
\sigma_\eta^2 &= \frac{RSS_{FE}}{N(T-1)-N-k} \\
RSS_{POLS} &= residual\ sum\ of\ squares\ from\ POLS\ regression, \\
RSS_{FE} &= residual\ sum\ of\ squares\ from\ the\ \text{one-way}\ FE\ model.
\end{aligned}
$$

Hausman Test

The null and alternative hypothesis of Hausman test are:

$$
\begin{aligned}
H_0 &= RE\ model\ is\ appropriate; \\
H_1 &= FE\ model\ is\ appropriate.
\end{aligned}
$$

The Hausman test statistic is:

$$\tau = (\hat{\beta}_{FE,W} - \hat{\beta}_{RE})^2[V(\hat{\beta}_{FE,W}) - V(\hat{\beta}_{RE})] \ \sim \ \chi^2(1)$$

where $V(.)$ represents variance of the coefficient estimates. Note that we can write the form $(\hat{\beta}_{FE,W} - \hat{\beta}_{RE})^2$ because $\hat{\beta}_{FE,W}$ and $\hat{\beta}^2_{RE}$ are scalars in this case.

From Figure 2 panel (2.3), the Hausman test statistic is $11.59 > 3.84$ (critical value 5%). Therefore, we **reject** the null under 5% significance level, meaning that the FE model is appropriate in present case.

3. Comparison between Different Estimates

Comparative analysis among different estimates in econometric and economic aspects is performed in this section, assuming that the real exchange rate is stationary. The comparison is made mainly among POLS, MG and FE estimates.

Econometric Analysis

- *Heterogeneity, Endogeneity and Consistency* :
  The main advantage of FE estimator over POLS estimator is the FE estimator's ability to remedy the potential issue of **endogeneity** caused by correlation between explanatory variables with (unobserved) **heterogeneity**. If this issue of endogeneity does exist, FE estimator will generate **consistent** estimates while POLS estimator will produce **inconsistent** regression results. Since MG estimator uses estimates from OLS regression, which may also suffer from endogeneity problem. In this case, it is clear graphically that explanatory variable z_{it} fluctuates around different means for different countries, which signals the potential of heterogeneity.

  As is shown in Figure 1 panel (a), $\hat{\beta}_P$ is **insignificant** whereas in Figure 2 panel (2.1) $\hat{\beta}_{FE,W}$ is **significant** (from zero).

- *Robustness* :
  One advantage of MG estimates over separate OLS estimator is that the covariance matrix of MG estimator is a non-parameteric estimator of the standard errors of $\hat{\alpha}_{MG}$ and $\hat{\beta}_{MG}$, which does not depend on estimation of $V([\hat{\alpha}_{OLS,i}, \hat{\beta}_{OLS,i}]')$ that may not be robust to autocorrelation and/or heteroskedasticity. From Figure 1 panel (c), it can be seen that the standard error of MG estimate for slope coefficient ($\hat{\beta}_{MG}$) is much lower than the OLS counterpart of any country.

Economic Analysis

As is mentioned in the last section, $\hat{\beta}_P$ is **insignificant** while $\hat{\beta}_{FE,W}$ is **significant**. This significance of $\hat{\beta}_{FE,W}$ carries both statistical and economic meaning. Since the explanatory variable z_{it} represents an error-correction term, the significance implies the existence of an error-correction behaviour of exchange rate in the long run, which is suggested by the PPP theory. By the same token, the insignificance of $\hat{\beta}_P$ indicates no error-correction behaviour and invalidity of PPP theory when estimating by POLS regression. Again, this is probably because POLS estimator fails to handle with heterogeneity across countries.

To conclude, when facing panel data with a large sample size (as is in this case), cross sectional heterogeneity is a main issue to consider. As a result, fixed effects estimator, which has better ability to deal with heterogeneity over POLS, separate OLS and MG estimators, is preferred for modelling practice.

References

Greene, W. H. (2012), *Econometric analysis, 7th editions*, Pearson Education.

Appendix: Matlab Code

Code for Balanced Panel

```
%% Assignment for Econometrics I, Part B: Purchasing Power Parity (PPP)
%  Module Code: IB9BM0;
%  Author: Dalong Sun (Warwick ID 1257635)

%% 1. Load Data
Canada  = csvread('Canada_M.csv');
Denmark = csvread('Denmark_M.csv');
Euro    = csvread('Euro_M.csv');
Japan   = csvread('Japan_M.csv');
Korea   = csvread('Korea_M.csv');
Norway  = csvread('Norway_M.csv');
Sweden  = csvread('Sweden_M.csv');
Switz   = csvread('Switz_M.csv');
UK      = csvread('UK_M.csv');
US      = csvread('US_M.csv');

%% 2. Truncate the Data to Have Equal Length
%     Switz Is the Base for Truncation: 156 Obs, 95m1-07m12
%     US Is the Base Country
Norminal_FX = zeros(156,9); % no norminal FX data for US
CPI         = zeros(156,10); % include US data for CPI

Norminal_FX(:,1)  = Canada(277:432,1);
Norminal_FX(:,2)  = Denmark(253:408,1);
Norminal_FX(:,3)  = Euro(61:216,1);
Norminal_FX(:,4)  = Japan(277:432,1);
Norminal_FX(:,5)  = Korea(181:336,1);
Norminal_FX(:,6)  = Norway(193:348,1);
Norminal_FX(:,7)  = Sweden(277:432,1);
Norminal_FX(:,8)  = Switz(:,1);
Norminal_FX(:,9)  = UK(277:432,1);

CPI(:,1)  = Canada(277:432,4);
CPI(:,2)  = Denmark(253:408,4);
CPI(:,3)  = Euro(61:216,4);
CPI(:,4)  = Japan(277:432,4);
CPI(:,5)  = Korea(181:336,4);
CPI(:,6)  = Norway(193:348,4);
CPI(:,7)  = Sweden(277:432,4);
CPI(:,8)  = Switz(:,4);
CPI(:,9)  = UK(277:432,4);
CPI(:,10)= US(277:432,3);

%% 3. Data Transformation
```

```
% 3.1 Convert the measurement of nominal FX rate in terms of
%     domestic currency price (the US dollar price) PER UNIT of foreign currency
S_it = 1./Norminal_FX;

% 3.2 Log transformation
s_it = log(S_it);

p_0t = log(CPI(:,10)); % price levels for US (the base country)
p_it = log(CPI(:,1:9)); % price levels for the other 9 countries

%% 4. Create Dependent and Independent Variables for Regressions
Delta_s_itplus1 = s_it(2:end,:) - s_it(1:end-1,:);

z_it_Canada  = (p_0t(1:end-1) - p_it(1:end-1,1)) - s_it(1:end-1,1);
z_it_Denmark = (p_0t(1:end-1) - p_it(1:end-1,2)) - s_it(1:end-1,2);
z_it_Euro    = (p_0t(1:end-1) - p_it(1:end-1,3)) - s_it(1:end-1,3);
z_it_Japan   = (p_0t(1:end-1) - p_it(1:end-1,4)) - s_it(1:end-1,4);
z_it_Korea   = (p_0t(1:end-1) - p_it(1:end-1,5)) - s_it(1:end-1,5);
z_it_Norway  = (p_0t(1:end-1) - p_it(1:end-1,6)) - s_it(1:end-1,6);
z_it_Sweden  = (p_0t(1:end-1) - p_it(1:end-1,7)) - s_it(1:end-1,7);
z_it_Switz   = (p_0t(1:end-1) - p_it(1:end-1,8)) - s_it(1:end-1,8);
z_it_UK      = (p_0t(1:end-1) - p_it(1:end-1,9)) - s_it(1:end-1,9);

z_it = [z_it_Canada, z_it_Denmark, z_it_Euro, z_it_Japan, z_it_Korea,...
    z_it_Norway, z_it_Sweden, z_it_Switz, z_it_UK];

% Plot time series
Name = {'Canada', 'Denmark', 'Euro', 'Japan', 'Korea','Norway',...
    'Sweden','Switz','UK'};
% for independent variables
figure
for i = 1:9
    subplot(3,3,i), plot(z_it(:,i),'r')
    title(Name{i})
    axis tight
end
print -depsc PartB_Q1_IndepVar.eps

% for dependent variables
figure
for i = 1:9
    subplot(3,3,i), plot(Delta_s_itplus1(:,i))
    title(Name{i})
    axis tight
end
print -depsc PartB_Q1_DepVar.eps

%% 5. Question 1
[T, N] = size(z_it);
k      = 1; % k is the number of exogenous variables in the model, which is 1 is this cas
```

```matlab
% Part (a): Pooled OLS (POLS)
z_it_Pool = z_it(:); % stack z_it matrix into an NT by 1 vector
% alternatively, use z_it_P = reshape(z_it,[1395,1]);

X = [ones(length(z_it_Pool),1),z_it_Pool];
Y = Delta_s_itplus1(:); % alternatively, use Y = reshape(Delta_s_itplus1,[1395,1]);

% POLS estimator
Beta_Pool = lscov(X,Y);

% Covariance matrix and standard error (StdErr)
Y_hat_Pool        = X*Beta_Pool;
Resid_Pool        = Y - Y_hat_Pool;
RSS_Pool          = sum(Resid_Pool.^2);
Cov_Pool          = (RSS_Pool/(N*T-(k+1)))*inv(X'*X); % assume homoskedasticity
StdErr_Pool(1,1) = sqrt(Cov_Pool(1,1)); % StdErr for constant term
StdErr_Pool(2,1) = sqrt(Cov_Pool(2,2)); % StdErr for Beta_Pool
SER_Pool          = std(Y); % standard error of regression

% Diagnostic tests
Adj_R2_Pool          = 1 - (RSS_Pool/(N*T-(k+1))) / (Y'*Y/(N*T-1)); % Adjusted R square
[~,ADF_Pval_Pool]  = adftest(z_it_Pool); % Augmented Dickey-Fuller test
[~,ARCH_Pval_Pool] = archtest(Resid_Pool); % ARCH type heteroskedasticity
DW_Pval_Pool         = dwtest(Resid_Pool,X); % Durbin-Watson test

% Combine regression results into one matrix
Estimator_POLS(1,1)  = Beta_Pool(1);
Estimator_POLS(1,2)  = StdErr_Pool(1);
Estimator_POLS(2,1)  = Beta_Pool(2);
Estimator_POLS(2,2)  = StdErr_Pool(2);
Estimator_POLS(2,3)  = SER_Pool;
Estimator_POLS(2,4)  = Adj_R2_Pool;
Estimator_POLS(2,5)  = ADF_Pval_Pool;
Estimator_POLS(2,6)  = ARCH_Pval_Pool;
Estimator_POLS(2,7)  = DW_Pval_Pool;
Estimator_POLS       = RoundToDecimalPlace(Estimator_POLS,4); % round to 4 decimal places

% Convert double to cell format
Estimator_POLS       = num2cell(Estimator_POLS);

% T test
for i = 1:2
    if abs(Estimator_POLS{i,1}/Estimator_POLS{i,2}) >= 1.96
        Estimator_POLS{i,1} = num2str(Estimator_POLS{i,1});
        Estimator_POLS{i,1} = strcat(Estimator_POLS{i,1},'*');
    end
end

% Put StdErr in parenthese
for i = 1:2
    Estimator_POLS{i,2} = strcat('(',num2str(Estimator_POLS{i,2}),')');
```

```
end

% Write in Excel (doesn't work for Mac)
% xlswrite('Econometrics I Part B.xlsx',Estimator_POLS,...
%     'PartB Q1','D8:J9');

%-----------------------------------------------------------------------------
% Part (b): Seperate Country
Beta_Sep    = zeros(2,N); % 1st row is for alpha, 2nd row is for beta
Resid_Sep   = zeros(T,N);

StdErr_Sep = zeros(2,N); % 1st row is for alpha, 2nd row is for beta
SER_Sep    = zeros(1,N);

Adj_R2_Sep    = zeros(1,N); % Adjusted R square
ADF_Pval_Sep  = zeros(1,N); % Augmented Dickey-Fuller test
ARCH_Pval_Sep = zeros(1,N); % ARCH type heteroskedasticity
DW_Pval_Sep   = zeros(1,N); % Durbin-Watson test

for i = 1:9
    X    = [ones(T,1),z_it(:,i)];
    Y    = Delta_s_itplus1(:,i);
    Beta = lscov(X,Y);
    Beta_Sep(:,i)  = Beta;
    Resid_Sep(:,i) = Y - X*Beta;

    % Covariance matrix and standard error (StdErr)
    RSS_Sep         = sum(Resid_Sep(:,i).^2);
    Cov_Sep         = (RSS_Sep/(T-(k+1)))*inv(X'*X); % assume homoskedasticity
    StdErr_Sep(1,i) = sqrt(Cov_Sep(1,1));
    StdErr_Sep(2,i) = sqrt(Cov_Sep(2,2));
    SER_Sep(i)      = std(Y); % standard error of regression

    % Diagnostic tests
    Adj_R2_Sep(i)          = 1 - (RSS_Sep/(T-(k+1))) / (Y'*Y/(T-1)); % Adjusted R square
    [~,ADF_Pval_Sep(i)]  = adftest(z_it(:,i)); % Augmented Dickey-Fuller test
    [~,ARCH_Pval_Sep(i)] = archtest(Resid_Sep(:,i)); % ARCH type heteroskedasticity
    DW_Pval_Sep(i)         = dwtest(Resid_Sep(:,i),X); % Durbin-Watson test
end

% Combine regression results into one matrix
Estimator_OLS(1:2:17,1) = Beta_Sep(1,:);
Estimator_OLS(1:2:17,2) = StdErr_Sep(1,:);
Estimator_OLS(2:2:18,1) = Beta_Sep(2,:);
Estimator_OLS(2:2:18,2) = StdErr_Sep(2,:);
Estimator_OLS(2:2:18,3) = SER_Sep;
Estimator_OLS(2:2:18,4) = Adj_R2_Sep;
Estimator_OLS(2:2:18,5) = ADF_Pval_Sep;
Estimator_OLS(2:2:18,6) = ARCH_Pval_Sep;
Estimator_OLS(2:2:18,7) = DW_Pval_Sep;
Estimator_OLS           = RoundToDecimalPlace(Estimator_OLS,4); % round to 4 decimal plac
```

```
% Convert double to cell format
Estimator_OLS            = num2cell(Estimator_OLS);

% T test
for i = 1:18
    if abs(Estimator_OLS{i,1}/Estimator_OLS{i,2}) >= 1.96
        Estimator_OLS{i,1} = num2str(Estimator_OLS{i,1});
        Estimator_OLS{i,1} = strcat(Estimator_OLS{i,1},'*');
    end
end

% Put StdErr in parenthese
for i = 1:18
    Estimator_OLS{i,2} = strcat('(',num2str(Estimator_OLS{i,2}),')');
end

%-----------------------------------------------------------------------------
% Part (c): Mean Group (MG) Estimator (for RC Model)
Beta_MG = sum(Beta_Sep,2)/N;

% Covariance matrix and standard error (StdErr)
Summation = 0;
for i = 1:9
    Summation = Summation + ...
        (Beta_Sep(:,i) - Beta_MG)*(Beta_Sep(:,i) - Beta_MG)';
end
Cov_MG = Summation/(N*(N-1));
StdErr_MG(1,1) = sqrt(Cov_MG(1,1));
StdErr_MG(2,1) = sqrt(Cov_MG(2,2));

% Combine regression results into one matrix
Estimator_MG(1,1) = Beta_MG(1);
Estimator_MG(1,2) = StdErr_MG(1);
Estimator_MG(2,1) = Beta_MG(2);
Estimator_MG(2,2) = StdErr_MG(2);
Estimator_MG      = RoundToDecimalPlace(Estimator_MG,4); % round to 4 decimal places

% T test
Estimator_MG = num2cell(Estimator_MG);

for i = 1:2
    if abs(Estimator_MG{i,1}/Estimator_MG{i,2}) >= 1.96
        Estimator_MG{i,1} = num2str(Estimator_MG{i,1});
        Estimator_MG{i,1} = strcat(Estimator_MG{i,1},'*');
    end
end

% Put StdErr in parenthese
for i = 1:2
    Estimator_MG{i,2} = strcat('(',num2str(Estimator_MG{i,2}),')');
```

```
end

%% 6. Question 2 FE Model

% 6.1 FE Model Estimation
X = z_it(:); % stack z_it matrix (by cross-section) into an NT by 1 vector
Y = Delta_s_itplus1(:); % stack Y by cross-section

[Coeff_FE,Cov_Beta_FE,Resid_FE] = panFE(Y,X,155); % T = 155 obs
% function panFE() can be found on my.wbs lecture 4

Beta_FE  = RoundToDecimalPlace(Coeff_FE.slope,4); % Beta_FE uses 'Within' estimator
Alpha_FE = RoundToDecimalPlace(Coeff_FE.fe,4);

% Standard error (StdErr)
StdErr_FE(1,1) = RoundToDecimalPlace(std(Alpha_FE),4);
StdErr_FE(2,1) = RoundToDecimalPlace(sqrt(Cov_Beta_FE),4);

SER_FE         = RoundToDecimalPlace(std(Y),4); % standard error of regression

% Diagnostic tests
RSS_FE           = sum(Resid_FE.^2);
Adj_R2_FE        = 1 - (RSS_FE/(N*T-N-k)) / (Y'*Y/(N*T-1)); % Adjusted R square
[~,ADF_Pval_FE]  = adftest(X); % Augmented Dickey-Fuller test
[~,ARCH_Pval_FE] = archtest(Resid_FE); % ARCH type heteroskedasticity
DW_Pval_FE       = dwtest(Resid_FE,X); % Durbin-Watson test

% T test for Alpha_FE
Alpha_FE = num2cell(Alpha_FE);

for i = 1:9
    if abs(Alpha_FE{i}/StdErr_FE(1)) >= 1.96
        Alpha_FE{i} = num2str(Alpha_FE{i});
        Alpha_FE{i} = strcat(Alpha_FE{i},'*');
    end
end

% T test for Beta_FE
if abs(Beta_FE/StdErr_FE(2)) >= 1.96
    beta_FE = strcat(num2str(Beta_FE),'*');
end

% Put StdErr in parenthese
StdErr_FE = num2cell(StdErr_FE);

for i = 1:2
    StdErr_FE{i} = strcat('(',num2str(StdErr_FE{i}),')');
end

% Combine regression results into one matrix
alpha_FE(1:2:17,1)    = Alpha_FE;
```

```
Estimator_FE(1:17,1) = alpha_FE;
Estimator_FE(1,2)    = StdErr_FE(1);
Estimator_FE(1,3)    = {beta_FE};
Estimator_FE(1,4)    = StdErr_FE(2);
Estimator_FE(1,5)    = {SER_FE};
Estimator_FE(1,6)    = {RoundToDecimalPlace(Adj_R2_FE,4)};
Estimator_FE(1,7)    = {RoundToDecimalPlace(ADF_Pval_FE,4)};
Estimator_FE(1,8)    = {RoundToDecimalPlace(ARCH_Pval_FE,4)};
Estimator_FE(1,9)    = {RoundToDecimalPlace(DW_Pval_FE,4)};

%---------------------------------------------------------------------------
% 6.2 Test the Hypothesis: the constant term is the same for all cross sectional units
%     we may assume constant/homogeneous variance across all units here
Var_Resid_Sep = RoundToDecimalPlace(var(Resid_Sep),4); % check for homoskedasticity

var_Sep_hat = sum(sum(Resid_Sep.^2))/(N*T); % assume homoskedasticity
MLL_Sep = -0.5*N*T*(log(2*pi)+1) - 0.5*N*T*log(var_Sep_hat); % This is MLL_B in lecture 4

var_FE_hat = sum(Resid_FE.^2)/(N*T-N-k);
MLL_FE = -0.5*N*T*(log(2*pi)+1) - 0.5*N*T*log(var_FE_hat); % This is MLL_C in lecture 4 s

Test_LR    = 2*(MLL_Sep - MLL_FE); % This is the LR test 'tao_2'  in lecture 4 slide P53
C_value_LR = chi2inv(0.95,(N-1)*k); % Chi Square inverse CDF with DoF of 8 under 95% CL

% Result for Hypothesis Testing
if Test_LR >= C_value_LR
    HT_LR = 'Reject the Null of Constant Beta Across Units Under 95% CL';
else
    HT_LR = 'Fail to Reject the Null of Constant Beta Across Units Under 95% CL';
end

% Combine results into one vector
Result_LR(1,[1,3,5]) = {'Test Stats','Critival Value','Result'};
Result_LR(2,1) = {RoundToDecimalPlace(Test_LR,4)};
Result_LR(2,3) = {RoundToDecimalPlace(C_value_LR,4)};
Result_LR(2,5) = {HT_LR};

%---------------------------------------------------------------------------
% 6.3 Should We Use a RE Model? (Use the Hausman Test)

% (1) Feasible GLS for RE Model

% The feasible GLS requires an estimate of theta, which in the current case
% is based on the FE model
var_eta_hat = (RSS_Pool/(N*T-(k+1))) - var_FE_hat;
theta_hat   = 1 - sqrt(var_FE_hat/(T*var_eta_hat + var_FE_hat));

X_i_bar = mean(z_it);
Y_i_bar = mean(Delta_s_itplus1);
```

```
count = 1;
for i = 1:9
    X(count:T*i) = X(count:T*i) - theta_hat*X_i_bar(i);
    Y(count:T*i) = Y(count:T*i) - theta_hat*Y_i_bar(i);
    count = T*i + 1;
end

X = [ones(length(X),1), X];

Beta_RE      = lscov(X,Y);
Resid_RE     = Y - X*Beta_RE;
RSS_RE       = sum(Resid_RE.^2);
Cov_RE       = (RSS_RE/(N*T-(k+1)))*inv(X'*X); % assume homoskedasticity
Cov_Beta_RE = Cov_RE(2,2);

% Diagnostic tests
SER_RE            = std(Y); % standard error of regression
Adj_R2_RE         = 1 - (RSS_RE/(N*T-(k+1))) / (Y'*Y/(N*T-1)); % Adjusted R square
[~,ADF_Pval_RE]  = adftest(X(:,2)); % Augmented Dickey-Fuller test
[~,ARCH_Pval_RE] = archtest(Resid_RE); % ARCH type heteroskedasticity
DW_Pval_RE        = dwtest(Resid_RE,X); % Durbin-Watson test

% Combine regression results into one matrix
Estimator_RE(1,1)  = {RoundToDecimalPlace(Beta_RE(1),4)};
Estimator_RE(1,2)  = {RoundToDecimalPlace(sqrt(Cov_RE(1,1)),4)};
Estimator_RE(2,1)  = {RoundToDecimalPlace(Beta_RE(2),4)};
Estimator_RE(2,2)  = {RoundToDecimalPlace(sqrt(Cov_Beta_RE),4)};
Estimator_RE(2,3)  = {RoundToDecimalPlace(SER_RE,4)};
Estimator_RE(2,4)  = {RoundToDecimalPlace(Adj_R2_RE,4)};
Estimator_RE(2,5)  = {RoundToDecimalPlace(ADF_Pval_RE,4)};
Estimator_RE(2,6)  = {RoundToDecimalPlace(ARCH_Pval_RE,4)};
Estimator_RE(2,7)  = {RoundToDecimalPlace(DW_Pval_RE,4)};

% T test for Beta_RE
for i = 1:2
    if abs(Estimator_RE{i,1}/Estimator_RE{i,2}) >= 1.96
        Estimator_RE{i,1} = strcat(num2str(Estimator_RE{i,1}),'*');
    end
end

% Put StdErr in parenthese
for i = 1:2
    Estimator_RE{i,2} = strcat('(',num2str(Estimator_RE{i,2}),')');
end

% (2) Hausman Test
q   = Beta_FE - Beta_RE(2);
v_q = Cov_Beta_FE - Cov_Beta_RE;

Test_Hausman     = q'*inv(v_q)*q;
C_value_Hausman = chi2inv(0.95,k);
```

```
% Result for Hypothesis Testing
if Test_Hausman >= C_value_Hausman
    HT_Hausman = 'Reject the Null Under 95% CL, FE Model Is Appropriate';
else
    HT_Hausman = 'Fail to Reject the Null Under 95% CL, RE Model Is Appropriate';
end

% Combine results into one vector
Result_Hausman(1,[1,3,5]) = {'Test Stats','Critival Value','Result'};
Result_Hausman(2,1) = {RoundToDecimalPlace(Test_Hausman,4)};
Result_Hausman(2,3) = {RoundToDecimalPlace(C_value_Hausman,4)};
Result_Hausman(2,5) = {HT_Hausman};

%% 7. Remove all the variables for intermediate steps
% Only keep the final resulting variables
clearvars -except...
    Estimator_POLS Estimator_OLS Estimator_MG Estimator_FE Estimator_RE...
    Result_LR Result_Hausman Var_Resid_Sep


%% Classical estimation of the fixed effects panel data model
%%
%%
function[coeff,COVb,res]=panFE(Y,X,T)
% Y and X stacked by cross-section; T is the time dimension
% Estimator for panel data with fixed effects (balanced panel)
% coeff contains the estimator of the slope (slope) and the fixed effects (fe)
% COVb contains the estimated covariance matrix of the slope estimator
[NT,m] = size(Y);
[S,K]=size(X);
N=NT/T;
%within estimator
%build the matrix D
D=zeros(NT,N);
c=1;
for i=1:N,
D(c:T*i,i)=ones(T,1);
c=T*i+1;
end;
M=eye(NT)-D*inv(D'*D)*D';
b=inv(X'*M*X)*X'*M*Y;
a=inv(D'*D)*D'*(Y-X*b);
coeff.slope=b;
coeff.fe=a;
%compute the covariance matrix for the estimated coefficients
Xm=M*X;
Ym=M*Y;
res=Ym-Xm*b;
varres=(1/(NT-N-K))*res'*res;
COVb=varres*inv(X'*M*X);
```