1. Consider the model

$$y = S\delta + X\beta + u$$

where X is a $T \times K$ matrix of regressors and S is a matrix of seasonal dummy variables of the form

$$S = (s_1, s_2, s_3, s_4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Specifically, assume 4 seasons and N years of data so that $T = 4N$.

(a) Show that the OLS estimator of β can be written

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

ANS: Frisch-Waugh Theorem

(b) Paying particular attention to the definitions of the relevant variables, describe a two-stage procedure for estimating β.
ANS: $M_S X$ and $M_S y$ removes seasonal means prior to regression.

(c) Why would it be inadvisable to include an intercept in this model?
ANS: The model would be, say

$$y = \gamma i + S\delta + X\beta + u$$

and the regressor matrix would be muticollinear since $i = \sum_{j=1}^{4} s_j$.

(d) Suppose you estimated the model

$$y = Z\gamma + X\beta + u$$

where

$$Z = [i, s_2, s_3, s_4]$$

and i is the $T \times 1$ sum vector. How, if at all, would this change the estimation results? (Hint: Find the matrix A, such that $Z = SA$.).
ANS: Since

$$Z = [s_1 + s_2 + s_3 + s_4, s_2, s_3, s_4]$$

It follows that

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

and

$$
\begin{aligned}
P_Z &= Z\left(Z'Z\right)^{-1} Z' \\
&= SA\left(A'SS'A\right)^{-1} A'S' \\
&= SA\left(A^{-1}\left(SS'\right)^{-1}\left(A'\right)^{-1}\right) A'S' \\
&= P_S
\end{aligned}
$$

Hence $M_Z = M_S$. Together these imply the same fitted values and resifuals

2. Consider the model

$$
\begin{aligned}
y &= X\beta + u \\
&= X_1\beta + \gamma z + u
\end{aligned}
$$

where X_1 is $T \times (K-1)$, z is $T \times 1$ and $u \sim N\left(0, \sigma^2 I\right)$.

(a) Show that

$$
\frac{z'M_1 y}{\sigma\sqrt{z'M_1 z}} \sim N\left(0,1\right)
$$

ANS:

$$
\begin{aligned}
\hat{\gamma} &= \left(z'M_1 z\right)^{-1} z'M_1 y \\
&= \frac{z'M_1}{\left(z'M_1 z\right)}\left(X\beta + \gamma z + u\right) \\
&\quad \gamma + \frac{z'M_1 u}{z'M_1 z}
\end{aligned}
$$

Hence $E\left(\hat{\gamma}\right) = \gamma$ and

$$
\begin{aligned}
V\left(\hat{\gamma}\right) &= E\left[\left(z'M_X z\right)^{-1} z'M_x uu'M_X z\left(z'M_X z\right)^{-1}\right] \\
&= \sigma^2\left(z'M_X z\right)^{-1}
\end{aligned}
$$

Under the null

$$
\frac{\hat{\gamma}}{\sqrt{V\left(\hat{\gamma}\right)}} = \frac{z'M_1 y}{\sigma\sqrt{z'M_1 z}}
$$

(b) Show that $y'M_X y/\sigma^2$ has a Chi-squared distribution with $T-K$ degrees of freedom.
ANS:

$$
y'M_X y = u'M_X u
$$

Define

$$
w = B'u
$$

where B is the orthogonal matrix which diagonalises M_X such that

$$B'M_XB = \Lambda = \begin{bmatrix} I_{T-K} & 0 \\ 0 & 0 \end{bmatrix}$$

We have

$$E(w) = B'E(u) = 0$$

and

$$Var(w) = E\left[(B'u)(B'u)'\right] = \sigma^2 B'B = \sigma^2 I$$

Therefore, $w \backsim N\left(0, \sigma^2 I\right)$. However,

$$Bw = BB'u = u$$

and so

$$\begin{aligned} u'M_Xu &= w'B'M_XBw \\ &= w'\Lambda w \\ &= \sum_{i=1}^{T-K} w_i^2 \end{aligned}$$

Hence

$$\sum_{i=1}^{T-K} \left(\frac{w_i}{\sigma}\right)^2 \sim \chi^2_{T-K}$$

(c) Show that $y'M_Xy$ is independent of $z'M_1y$.
ANS: Given normality. independence follows if the covariance is zero.

$$\begin{aligned} COV(M_Xy, z'M_1y) &= COV(M_Xu, z'M_1u) \\ &= E(M_Xuu'M_1z) \\ &= \sigma E(M_XP_Xz) \\ &= 0 \end{aligned}$$

(d) Hence, show that

$$\frac{z'M_1y}{s\sqrt{z'M_1z}} \sim \chi^2_{T-K}$$

where

$$s^2 = \frac{e'e}{T-K}$$

and e is the residual vector

ANS:

$$\begin{aligned}
\frac{z'M_1y}{s\sqrt{z'M_1z}} &= \frac{z'M_1y}{\sigma\sqrt{z'M_1z}}\frac{\sigma}{s} \\
&= \frac{z'M_1y}{\sigma\sqrt{z'M_1z}}\sqrt{\frac{(T-K)\,\sigma^2}{y'M_Xy}} \\
&= \frac{\dfrac{z'M_1y}{\sigma\sqrt{z'M_1z}}}{\sqrt{\dfrac{(y/\sigma)'\,M_X\,(y/\sigma)}{(T-K)}}} \\
&= \frac{N\,(0,1)}{\sqrt{\chi^2_{T-K}/T-K}} \\
&\sim t_{T-K}
\end{aligned}$$

3. Describe the Likelihood Ratio, Wald and Lagrange Multiplier approaches to hypothesis testing. What do each of these imply for testing the null hypothesis that $\gamma = 0$ in the model

$$y = X\beta + Z\gamma + u$$

4. Consider the model

$$y = X\beta + u$$

where y is a $T \times 1$ vector, X is a $T \times K$ matrix, β is a $K \times 1$ vector. If the data is divided into two subsamples of size T_1 and T_2 with $T_1 + T_2 = T$, we can allows for the possibility that the coefficients are different in the two subsamples by writing the model as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

(a) Describe a procedure for testing the hypothesis that $\beta_1 = \beta_2$ by running a restricted regression and an unrestricted regression. Carefully state the test statistic and the distribution of the test statistic.

(b) Show that the unrestricted residual sum of squares can be obtained by running individual regressions for the two subsamples and adding together the two sums of squared residuals.

(c) The general form of the test statistic for a set of linear restrictions is

$$\frac{\left(R\hat{\beta} - r\right)'\left[s^2R\,(X'X)^{-1}\,R'\right]^{-1}\left(R\hat{\beta} - r\right)}{q}$$

Use this expression to derive a test statistic for the hypothesis that $\beta_1 = \beta_2$.

(d) Show that the model can be written

$$y = X\beta_1 + Z\gamma + u$$

where

$$Z = \begin{bmatrix} 0 \\ X_2 \end{bmatrix}$$

is a $T \times K$ matrix with zeros in the first T_1 rows and $\gamma = \beta_1 - \beta_2$. How is the test statistic for the hypothesis $\gamma = 0$ related to the test statistic for the hypothesis that $\beta_1 = \beta_2$?

5. Examine the main statistical properties of the OLS estimator of β in the linear model

$$y = X\beta + u$$

making clear which assumptions are required for each property to hold. Briefly compare the consequences of including irrelevant variables with those of excluding relevant variables.

6. Examine the role of the F Distribution in testing linear restrictions. Comment on the difference between the small sample case and the large sample case.

7. Consider the model

$$\begin{aligned} y_t &= \alpha + \beta x_t + u_t \\ u_t &\sim IN(0, \gamma) \quad \gamma > 0 \end{aligned}$$

(a) Find the Maximum Likelihood estimators of α, β and σ^2.
ANS:

$$\mathcal{L}\left(a, \beta, \sigma^2 | x_1, x_2, ..., x_T, y_1, ..., y_T\right) = \left(\frac{1}{2\pi\gamma}\right)^{T/2} \exp\left[-\frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t - \alpha - \beta x_t\right)^2}{\gamma}\right]$$

so

$$L = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln(\gamma) - \frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t - \alpha - \beta x_t\right)^2}{\gamma}$$

Differentiating

$$\begin{aligned} \frac{\partial L}{\partial \alpha} &= \sum_{t=1}^{T}\left(\frac{y_t - \alpha - \beta x_t}{\gamma}\right) \\ \frac{\partial L}{\partial \beta} &= \sum_{t=1}^{T}\left(\frac{y_t - \alpha - \beta x_t}{\gamma}\right) x_t \\ \frac{\partial L}{\partial \sigma^2} &= -\frac{T}{2\gamma} + \frac{1}{2\gamma^2}\sum_{t=1}^{T}\left(y_t - \alpha - \beta x_t\right)^2 \end{aligned}$$

Setting these to zero, the first two equations give

$$\begin{aligned}\sum_{t=1}^{T}\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right) &= 0\\ \sum_{t=1}^{T}\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)x_t &= 0\end{aligned}$$

From the first

$$\hat{\alpha}=\frac{1}{T}\sum_{t=1}^{T}y_t-\hat{\beta}\frac{1}{T}\sum_{t=1}^{T}x_t$$

Substituting into the second

$$\begin{aligned}\sum_{t=1}^{T}y_tx_t &= \left[\frac{1}{T}\sum_{t=1}^{T}y_t-\hat{\beta}\frac{1}{T}\sum_{t=1}^{T}x_t\right]\sum_{t=1}^{T}x_t+\hat{\beta}\sum_{t=1}^{T}x_t^2\\ &= \frac{1}{T}\sum_{t=1}^{T}y_t\sum_{t=1}^{T}x_t+\hat{\beta}\left[\sum_{t=1}^{T}x_t^2-\frac{1}{T}\left(\sum_{t=1}^{T}x_t\right)^2\right]\end{aligned}$$

or

$$\hat{\beta}=\frac{\frac{1}{T}\sum_{t=1}^{T}y_tx_t-\frac{1}{T}\sum_{t=1}^{T}y_t\frac{1}{T}\sum_{t=1}^{T}x_t}{\frac{1}{T}\sum_{t=1}^{T}x_t^2-\left(\frac{\sum_{t=1}^{T}x_t}{T}\right)^2}$$

Rearranging the third equation

$$\hat{\gamma}=\frac{1}{T}\sum_{t=1}^{T}\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)^2$$

(b) Now consider the model

$$\begin{aligned}y_t &= \alpha+\beta x_t+u_t\\ u_t &\sim IN\left(0,\gamma+\phi x_t^2\right) \quad \gamma>0,\phi\geq 0\end{aligned}$$

Use the first order conditions for a maximum to show that the ML estimators for α and β are the same as those in (a).
ANS:

$$f\left(\alpha,\beta,\gamma,\phi;y_t,x_t\right)=\frac{1}{\sqrt{2\pi\left(\gamma+\phi x_t^2\right)}}\exp\left[-\frac{1}{2}\frac{\left(y_t-\alpha-\beta x_t\right)^2}{\gamma+\phi x_t^2}\right]$$

$$\ln f\left(\alpha,\beta,\gamma,\phi;y_t,x_t\right)=-\frac{1}{2}\ln\left[2\pi\left(\gamma+\phi x_t^2\right)\right]-\frac{1}{2}\frac{\left(y_t-\alpha-\beta x_t\right)^2}{\gamma+\phi x_t^2}$$

$$L=-\frac{T}{2}\ln\left(2\pi\right)-\frac{1}{2}\sum_{t=1}^{T}\ln\left(\gamma+\phi x_t^2\right)-\frac{1}{2}\sum_{t=1}^{T}\frac{\left(y_t-\alpha-\beta x_t\right)^2}{\left(\gamma+\phi x_t^2\right)}$$

$$\begin{aligned}
\frac{\partial L}{\partial \alpha} &= -\frac{1}{2}\sum_{t=1}^{T}\frac{2\left(y_t-\alpha-\beta x_t\right)}{\left(\gamma+\phi x_t^2\right)}(-1)=\sum_{t=1}^{T}\frac{\left(y_t-\alpha-\beta x_t\right)}{\left(\gamma+\phi x_t^2\right)}\\
\frac{\partial L}{\partial \beta} &= -\frac{1}{2}\sum_{t=1}^{T}\frac{2\left(y_t-\alpha-\beta x_t\right)}{\left(\gamma+\phi x_t^2\right)}(-x_t)=\sum_{t=1}^{T}\frac{\left(y_t-\alpha-\beta x_t\right)x_2}{\left(\gamma+\phi x_t^2\right)}
\end{aligned}$$

$$\frac{\partial L}{\partial \alpha}=0\Longrightarrow\sum_{t=1}^{T}\frac{\left(y_t-\alpha\hat{}-\hat{\beta}x_t\right)}{\left(\hat{\gamma}+\hat{\phi}x_t^2\right)}=0\Longrightarrow\sum_{t=1}^{T}\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)=0\Longrightarrow\hat{\alpha}=\bar{y}-\hat{\beta}\bar{x}$$

$$\frac{\partial L}{\partial \beta}=0\Longrightarrow\sum_{t=1}^{T}\frac{\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)x_2}{\left(\hat{\gamma}+\hat{\phi}x_t^2\right)}=0\Longrightarrow\sum y_t x_t=\hat{\alpha}\sum x_t+\hat{\beta}\sum x_2^2$$

(c) Find the ML destimators of γ and ϕ.
ANS:

$$\begin{aligned}
\frac{\partial L}{\partial \gamma} &= -\frac{1}{2}\sum_{t=1}^{T}\left(\frac{1}{\gamma+\phi x_t^2}\right)-\frac{1}{2}\sum_{t=1}^{T}\left[-\frac{\left(y_t-\alpha-\beta x_t\right)^2}{\left(\gamma+\phi x_t^2\right)^2}\right]\\
\frac{\partial L}{\partial \phi} &= -\frac{1}{2}\sum_{t=1}^{T}\left(\frac{2\phi x_t}{\gamma+\phi x_t^2}\right)-\frac{1}{2}\sum_{t=1}^{T}\left[-\frac{\left(y_t-\alpha-\beta x_t\right)^2}{\left(\gamma+\phi x_t^2\right)^2}2\phi x_t\right]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial L}{\partial \gamma} &= 0\Longrightarrow\sum_{t=1}^{T}\left(\frac{1}{\hat{\gamma}+\hat{\phi}x_t^2}\right)=\sum_{t=1}^{T}\left[\frac{\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)^2}{\left(\hat{\gamma}+\hat{\phi}x_t^2\right)^2}\right]\\
&\Longrightarrow \sum_{t=1}^{T}\left[\frac{\left(\hat{\gamma}+\hat{\phi}x_t^2\right)-\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)^2}{\left(\hat{\gamma}+\hat{\phi}x_t^2\right)^2}\right]=0\\
&\Longrightarrow T\hat{\gamma}+\hat{\phi}\sum x_t^2=\sum\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)^2=\sum e_t^2\\
\frac{\partial L}{\partial \phi} &= 0\Longrightarrow\sum_{t=1}^{T}\left(\frac{x_t}{\hat{\gamma}+\hat{\phi}x_t^2}\right)=\sum_{t=1}^{T}\left[\frac{\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)^2}{\left(\hat{\gamma}+\hat{\phi}x_t^2\right)^2}x_t\right]\\
&\Longrightarrow \sum_{t=1}^{T}\left[\frac{x_t\left(\hat{\gamma}+\hat{\phi}x_t^2\right)-x_t\left(y_t-\hat{\alpha}-\hat{\beta}x_t\right)^2}{\left(\hat{\gamma}+\hat{\phi}x_t^2\right)^2}\right]=0\\
&\Longrightarrow \sum_{t=1}^{T}x_t\left(\hat{\gamma}+\hat{\phi}x_t^2\right)=\sum x_t e_t^2\\
&\Longrightarrow \hat{\gamma}\sum_{t=1}^{T}x_t+\hat{\phi}\sum_{t=1}^{T}x_t^2=\sum x_t e_t^2
\end{aligned}$$

$$\begin{aligned}\sum e_t^2 &= T\hat{\gamma} + \hat{\phi}\sum x_t^2 \\ \sum x_t e_t^2 &= \hat{\gamma}\sum_{t=1}^{T} x_t + \hat{\phi}\sum_{t=1}^{T} x_t^2\end{aligned}$$

$$\begin{aligned}\hat{\phi} &= \frac{\left[\sum x_t e_t^2 - \frac{1}{T}\sum_{t=1}^{T} x_t \sum e_t^2\right]}{\left[\sum_{t=1}^{T} x_t^2 - \frac{1}{T}\left(\sum x_t^2\right)^2\right]} \\ &= \frac{\frac{1}{T}\sum x_t e_t^2 - \left(\frac{1}{T}\sum_{t=1}^{T} x_t\right)\left(\frac{1}{T}\sum e_t^2\right)}{\frac{1}{T}\sum_{t=1}^{T} x_t^2 - \left(\frac{\sum x_t^2}{T}\right)^2}\end{aligned}$$

(d) Use your results to suggest a test for heteroscedasticity using an auxilliary regression based on estimating the model with $\phi = 0$.
ANS: Regress e_t^2 on a constant and x_t and carry out a t test.

8. The following Table shows the results from estimating a number of regressions explaining the size of the government in a sample of countries. The basic hypotheses is that there are increasing returns to government spending so that countries with larger populations can afford to have a smaller government. The basic hypothesis is testsed in Regressin (1). The other regressions include various control variables, in particular, regional dummy variables.

   (a) Test the hypothesis that the four dummy variables are jointly zero.

   (b) Explain what is meant by a 'heteroscdastic-consistent standard error' and discuss the circumstances in which these will be used.

   (c) Assess the regression results.