1. **Question 1**: Show that the OLS estimator $\boldsymbol{b}$ can be written as a weighted sum of the within and between (i.e. fixed and random effects) estimators.

2. **Question 2**: Consider the random effects model:

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + (\alpha + u_i) + \varepsilon_{it}$$

where
- $E[\varepsilon_{it} \mid \mathbf{X}] = E[u_i \mid X] = 0$
- $E[\varepsilon_{it}^2 \mid \mathbf{X}] = \sigma_\varepsilon^2$, $E[u_i^2 \mid X] = \sigma_u^2$
- $E[\varepsilon_{it} u_i \mid \mathbf{X}] = 0 \ \forall \ i, t, j$
- $E[\varepsilon_{it}\varepsilon_{js} \mid \mathbf{X}] = 0 \ \forall \ t \neq s, i \neq j$
- $E[u_i u_j \mid \mathbf{X}] = 0 \ \forall \ i \neq j$

Show whether or not the OLS classical assumption of spherical disturbances (Assumption 4) holds.

3. **Question 3**: Download the data from the class website (file "HW10a data"). It is a panel data on investment (y) and profit (x) for 3 firms over T=10 periods.

   (a) Pool the data and compute the least squares regression coefficients of the model:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it}$$

   (b) Estimate the fixed effects model of Greene (11-13) (9-12 in the 6th edition) and then test the hypothesis that the constant term is the same for all three firms.

   (c) Estimate the random effects model of Greene (11-29) (9-26 in the 6th edition) and then carry out the Lagrange multiplier test of hypothesis that the classical model without the common effect applies.

   (d) Carry out Hausman's specification test for the random versus the fixed effects model.

4. **Question 4**: The data in Greene Appendix Table F6.1 (and found on the class website, file "HW10b data") are a balanced panel on six U.S. airlines between 1970-1984, before the industry was deregulated. The data includes output, total cost, and fuel price, as well as data on the load factor, which is the average capacity utilization of the airline's fleet.

   Use these data to build a cost model for airline service. Allow for cross-airline heterogeneity in the constants in the model. Use both random and fixed effects specifications and statistical tests to determine which is the preferred model. An appropriate cost model to begin the analysis would be:

$$\ln cost_{it} = \alpha + \beta \ln price_{it} + \gamma \ln output_{it} + \varepsilon_{it}$$

   You might want to generalize the cost function by including a quadratic term of the output in the function. A translog model would include the square of the log of price and products of the log of output and the log of price (as in the previous homework problem). The load factor variable might also influence costs. Consider including it in your model and use the appropriate test statistic to test whether they are, indeed, relevant to the determination of (log) total cost.