Q1. (30 pts.) Consider the following panel data model:

\[ y_i = X_i \beta + (e_i + a_i), \]

where \( i = 1, ..., N \), and all of the symbols are defined in the class notes. \( X \) contains \( k \) variables. Note that this model does not have time-invariant regressors. For the entire data, this model can be expressed as \( y = X \beta + (v + \epsilon) \), where \( \alpha = (a_1, ..., a_N)' \). Answer the following questions.

1. (10 pts.) Assuming that \( \sigma^2_e \) is known, derive \( \text{Cov}(\hat{\beta}_W) \).
2. (10 pts.) Assume that the \( a_i \) are random with \( N(0, \sigma^2_a) \) and uncorrelated with \( X_i \) (the random effects assumption). Assume that both \( \sigma^2_a \) and \( \sigma^2_e \) are known. Under these assumptions, derive \( \text{Cov}(\hat{\beta}_{GLS}) \).
3. (10 pts.) Under the random effects assumptions, show that \( \hat{\beta}_{GLS} \) is more efficient than \( \hat{\beta}_W \).

Q2. (10 pts.) Show that the estimator \( \hat{\theta} = \sqrt{s^2 / S^2_\hat{\theta}} \) given in p. Basic Panel – 21 of the class note is consistent.

Q3. (20 pts.) Consider the model given in Q1. Suppose you have unbalanced data: that is, the \( T_i \) are different across \( i \).

1. (10 pts.) Derive \( \hat{\beta}_{GLS} \) and \( \text{Cov}(\hat{\beta}_{GLS}) \) for the data.
2. (10 pts.) Explain the GLS procedure for the data. Be specific as much as you can.

Q3. (40 pts.) Use panel_1.txt (This data set contains 4 variables, \( y, x_1, x_2, \) and \( z \); and \( N = 30 \) and \( T = 5 \)). You wish to estimate:

\[ y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \gamma_1 + \gamma_2 z_i + (a_i + \epsilon_{it}), \]

where the \( a_i \) are iid \( (0, \sigma^2_a) \) over different \( i \) and the \( \epsilon_{it} \) are iid \( (0, \sigma^2_e) \) over different \( i \) and \( t \). Use Gauss to answer the following questions. Make your own programs and report your programs.

1. (10 pts.) Estimate the above model by within. Report variable names, estimates, standard errors, and t-statistics. Also, report \( R^2 \).
2. (10 pts.) Assuming that the effects \( a_i \) are random and uncorrelated with the regressors, estimate the parameters by GLS. Report variable names, estimates, standard errors, and t-statistics. Also, report \( R^2 \).
(3) (10 pts.) Using the Hausman test, decide whether the random effects assumption is appropriate for the above model. Use 95% of confidence level.

(4) (10 pts.) Now assume that the errors \( \varepsilon_{it} \) are cross-sectional heteroskedastic over \( i \), and/or are heteroskedastic and autocorrelated over \( t \). Compute the appropriate standard errors for the within estimates of \( \beta \)'s.