<u>Problem 1</u> (22 Points)

Consider the following linear regression model

$$y_t = \beta_1 + \beta_2 x_{t2} + e_t, \quad e_t \sim N(0, \sigma_t^2) \text{ independent for } t = 1, ..., T, \tag{1}$$

where x_{t2} is a non-stochastic regressor. Let $\Phi = diag[\sigma_1^2, ..., \sigma_T^2]$ be a positive definite matrix.

1. **(7 Points)** Assume that Φ is known.

   (a) Derive the covariance matrix of the OLS estimator b for $\beta = (\beta_1, \beta_2)'$ in model (1).

   (b) Propose an efficient estimator for β and derive its expectation.

   (c) What does efficiency of an estimator mean?

2. **(9 Points)** The following data are available:

$$X'X = \begin{bmatrix} 50.000 & 5.410 \\ 5.410 & 43.315 \end{bmatrix}, \quad (X'X)^{-1} = \begin{bmatrix} 0.020 & -0.003 \\ -0.003 & 0.023 \end{bmatrix},$$

$$X'y = \begin{bmatrix} 19.690 \\ 46.213 \end{bmatrix}, \quad y'y = 1995.$$

   Let Φ be unknown and proceed with the wrong assumption that $\Phi = \sigma^2 I_T$.

   (a) Calculate the OLS estimate b of the vector of coefficients $\beta = (\beta_1, \beta_2)'$.

   (b) Estimate the variance of the error term. (Hint: $\hat{e}'\hat{e} = y'y - b'X'y$.)

   (c) Estimate the covariance matrix of the OLS estimator b.

   (d) Compute the test statistic for testing the null hypothesis

$$H : \beta_2 = 0. \tag{2}$$

   Which critical value would you need at a 5% significance level?

3. **(6 Points)** In addition to the data above, let

$$X'\hat{\Phi}X = \sum_{t=1}^{T} \hat{e}_t^2 x_{(t)} x_{(t)}' = \begin{bmatrix} 1941.775 & 175.990 \\ 175.990 & 737.029 \end{bmatrix}$$

   be given, where the $\hat{e}_t$'s are the OLS residuals from part 2 and $\hat{\Phi} = diag[\hat{e}_1^2, ..., \hat{e}_T^2]$.

   (a) Compute the White heteroscedasticity consistent covariance matrix estimator.

   (b) For testing the null hypothesis (2) compute the test statistic based on the OLS estimate of β_2 and its White standard error. What can you conclude when comparing your results with that of part 2(c)?

<u>Problem 2</u> (23 Points)

Consider the following model with two equations

$$\begin{aligned} y_{t1} &= \beta_1 x_{t1} + e_{t1}, \\ y_{t2} &= \beta_2 x_{t2} + e_{t2}, \end{aligned}$$

where $x_{t1} = 1$ for all $t = 1, ..., T$ and x_{t2} is non-stochastic. For $T = 50$ observations the following sample moments are given:

$$x_2'x_2 = 100, \quad x_2'y_2 = 50, \quad x_2'y_1 = 60, \quad y_2'y_2 = 90, \quad y_1'y_1 = 500, \quad y_1'y_2 = 40,$$

$$\sum_{t=1}^{50} x_{t2} = 100, \quad \sum_{t=1}^{50} y_{t1} = 150, \quad \sum_{t=1}^{50} y_{t2} = 50.$$

1. **(2 Points)** Compute the OLS estimates of β_1 and β_2.
2. **(6 Points)** Assume that the error terms have the following structure:

   $$e = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \sim (0, \Sigma \otimes I_T) \quad \text{with} \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}.$$

   Estimate Σ using the OLS residuals $(\hat{e}_1, \hat{e}_2)$ from part 1 and

   $$\hat{\sigma}_{ij} = \frac{\hat{e}_i'\hat{e}_j}{T} \quad \text{for} \quad i, j = 1, 2.$$

3. **(8 Points)** Compute the feasible GLS estimator for β_1 and β_2 under the assumption of part 2:
   - Let $\hat{\Sigma}^{-1} = \begin{pmatrix} \hat{\sigma}^{11} & \hat{\sigma}^{12} \\ \hat{\sigma}^{21} & \hat{\sigma}^{22} \end{pmatrix}$. Rewrite the formula for $\hat{\hat{\beta}}_G$ in terms of $\hat{\sigma}^{ij}$ and the sample moments.
   - Calculate $\hat{\Sigma}^{-1}$.
   - Give the estimates for β_1 and β_2.
4. **(2 Points)** What is the motivation behind using the feasible GLS estimator?
5. **(5 Points)** Show that the GLS estimator of $\beta = (\beta_1, \beta_2)'$ is equivalent to the OLS estimator if Σ is a diagonal matrix.

<u>Problem 3</u> (29 Points)

The following model is given:

$$y \;=\; Z\beta + e, \qquad e \sim N(0, \sigma^2 I_T), \tag{3}$$

where Z is a stochastic $(T \times K)$ matrix. Furthermore it holds, for $T \to \infty$, that $\frac{\sum_{t=1}^{T} \mathbb{E}\left[z_{(t)} z'_{(t)}\right]}{T} \to A$, where A is a constant $(K \times K)$ matrix.

1. **(12 Points)** Assume that e and Z are stochastically independent. Show that then it holds that $\text{plim}_{T\to\infty} \frac{Z'e}{T} = 0$:

   - Write $\frac{Z'e}{T}$ as a sum.
   - Show that $\mathbb{E}\left[z_{(t)} e_t\right] = 0$.
   - How does Cov $\left[z_{(t)} e_t\right]$ look like?
   - Compute Cov $\left[z_{(t)} e_t, z_{(s)} e_s\right]$, $s \neq t$.
   - Show that, for $T \to \infty$, $\mathbb{E}\left[\frac{Z'e}{T}\right] \to 0$.
   - Show that, for $T \to \infty$, Cov $\left[\frac{Z'e}{T}\right] \to 0$.
   - Why can you conclude now that $\text{plim}_{T\to\infty} \frac{Z'e}{T} = 0$?

2. **(15 Points)** Consider now the following model:

$$y_t = \beta y_{t-1} + e_t, \qquad e_t = \nu_t + \rho \nu_{t-1}, \quad t = 1, \ldots, T, \tag{4}$$

$$|\beta| < 1,\ |\rho| < 1,\ \nu_t \sim N(0, \sigma_\nu^2)\ i.i.d.$$

   In addition to this, y_{-1} and y_0 are observable. From the assumptions above it follows that for all t, $\mathbb{E}[y_t] = 0$, Var $[y_t] = \sigma_y^2$ and $\mathbb{E}\left[\nu_t y_{t-s}\right] = 0$ for $s > 0$.

   (a) Write the matrix Z for model (4) in terms of the observations of which it is composed. Write $\frac{Z'e}{T}$ as a sum which depends on the regressor variable y_{t-1}.

   (b) Show that $\mathbb{E}\left[y_{t-1} e_t\right] \neq 0$, if $\rho \neq 0$. What do you conclude?

   (c) Show that $\mathbb{E}\left[y_{t-2} e_t\right] = 0$ and in general $\mathbb{E}\left[y_{t-2} y_{t-1}\right] \neq 0$.

   (d) Give the elements of the instrument matrix X for model (4), if y_{t-2} will be used as an instrument for y_{t-1}. Represent $\frac{X'Z}{T}$ and $\frac{X'e}{T}$ as sums depending on the instrumental variable y_{t-2}.

   (e) How would you proceed to show that $\text{plim}_{T\to\infty} \frac{X'e}{T} = 0$ and $\text{plim}_{T\to\infty} \frac{X'Z}{T} \neq 0$?

3. **(2 Points)** What is wrong with the following text? Explain shortly.

   The instrumental variable estimator is a special case of the OLS estimator, and it is consistent if the instruments are uncorrelated with the stochastic regressors of the model.

<u>Problem 4</u> (26 Points)

The following model is given:

$$
\begin{aligned}
y_1 &= \gamma_1 y_2 + \beta_{11} x_1 + e_1 \qquad (5)\\
y_2 &= \gamma_2 y_1 + \beta_{22} x_2 + \beta_{32} x_3 + e_2 \qquad (6)
\end{aligned}
$$

where y_1, y_2 are endogenous and x_1, x_2, x_3 are exogenous $(T \times 1)$ vectors.

1. **(3 Points)** Write the model in the form $Y\Gamma + XB + E = 0$.

2. **(8 Points)** Check the identification of both equations using the order and the rank conditions.

3. **(6 Points)** Assume that an estimation $\hat{\Pi}$ of the reduced form parameters is known. Use the relation $\hat{\Pi}\Gamma_2 = -B_2$ (see the formulary) to estimate the structural parameters $\gamma_2, \beta_{22}, \beta_{32}$.

4. **(7 Points)** Show for equation (6) that $\hat{\delta}_{2(2SLS)} = \hat{\delta}_{2(ILS)}$:

   - Rewrite $\hat{Z}_2$ in dependence on Y_2 and X_2.
   - How many rows and columns does $X'Z_2$ have? Hint: Think about the variables, which appear in X, X_2, Y_2 and Z_2.
   - Assume that $X'Z_2$ has full column rank. Starting with $\hat{\delta}_{2(2SLS)}$ show that $\hat{\delta}_{2(2SLS)} = \hat{\delta}_{2(ILS)}$, i.e. $(\hat{Z}_2'\hat{Z}_2)^{-1}\hat{Z}_2' y_2 = (X'Z_2)^{-1}X'y_2$.

5. **(2 Points)** How does the identifiability of the model change, when you assume for economic reasons that $\beta_{22} = \beta_{32}$?

<u>Problem 5</u> (10 Points) Multiple Choice

For the following statements indicate whether they are True (T) or False (F). For each true answer you will get 2 points, for each false answer you will lose 1 point and for unanswered questions you will get 0 points. For the whole problem you will not get less than zero points.

Statements

T F	
[ ] [ ]	1. In a linear regression model the sum of the OLS residuals $(\sum_{t=1}^{T} \hat{e}_t)$ is always zero.
[ ] [ ]	2. The OLS residuals $\hat{e}_t$ are uncorrelated, if this is the case for the errors e_t.
[ ] [ ]	3. A break in the intercept of a linear regression model can be modelled by an appropriately defined dummy variable.
[ ] [ ]	4. If the regressor matrix X in the linear model $y = X\beta + e$ is not a full rank, then an unbiased estimator for β does not exist.
[ ] [ ]	5. The Gauss-Newton method for estimating the parameters in a nonlinear regression model is based on a second order Taylor series approximation of the least squares criterion.