Problem 1 **(21 Points)**

1. **(9 Points)** Consider the model

$$y = X\beta + e, \qquad e \sim (0, \sigma^2 I_T), \tag{1}$$

where X is a deterministic and nonsingular matrix.

   (a) Show that the OLS estimator $b = (X'X)^{-1}X'y$ is unbiased.

   (b) Derive the covariance matrix of b.

   (c) Explain what consistency of an estimator means. Show that b is consistent under the assumption that

$$\lim_{T\to\infty} \frac{X'X}{T} = Q \qquad \text{and} \qquad \operatorname*{plim}_{T\to\infty} \frac{X'e}{T} = 0,$$

   where Q is a finite and positive definite matrix.

2. **(12 Points)** Consider the linear regression model

$$y_t = \beta_1 + \beta_2 x_{t2} + e_t, \qquad e_t \sim N(0, \sigma^2)\ i.i.d., \qquad t = 1, ..., 50. \tag{2}$$

A structural break in the slope coefficient (β_2) is assumed in period $t = 30$, but the intercept (β_1) is not affected by the break.

   (a) Define a dummy variable that can be used to capture the structural break in β_2. Using this variable formulate the model in dummy variable representation. In addition to this, give the model in matrix notation by defining appropriate vectors and matrices.

   (b) Give the sets of equation representation of the model defined in 2(a). As above, give also the matrix version of the model.

   (c) Formulate the appropriate null hypothesis for testing that there is no structural break using the dummy variable representation. Which type of test statistic (no formula) would you use and how is it distributed under the null hypothesis?

<u>Problem 2</u> (28 Points)

1. **(2 Points)** Describe shortly what exact and near exact multicollinearity mean.

2. **(2 Points)** Give two symptoms of near exact multicollinearity.

3. Consider the following model

$$y = \mathbb{1}_T \alpha + x\beta + e, \qquad e \sim (0, \sigma^2 I_T), \tag{3}$$

where $\mathbb{1}_T$ is a $(T \times 1)$ vector of ones and x is a $(T \times 1)$ vector.

(a) **(4 Points)** Let $Q = I_T - \frac{\mathbb{1}_T \mathbb{1}_T'}{T}$. Write down the expressions for $y^* = Qy$ and $x^* = Qx$. How does a typical element for y^* and x^*, respectively, look like?

(b) **(4 Points)** Show that $y^* = Qy$ and $x^* = Qx$ are the vectors of OLS residuals $(\hat{u}, \hat{v})$ from two different regressions, i.e.

$$\begin{aligned} y &= \mathbb{1}_T \gamma + u, \\ x &= \mathbb{1}_T \delta + v. \end{aligned} \tag{4}$$

(c) **(4 Points)** Apply OLS to

$$Qy = Qx\beta + Qe$$

to get an estimator b for β (expressed in terms of y and x).

(d) **(4 Points)** Show that a direct application of the Frisch-Waugh Theorem to (3) provides the same estimator as in 3(c).

(e) **(5 Points)** Compute the variance of the estimator b derived in 3(c) and show that it can be represented in terms of the vector of OLS residuals $\hat{v}$ of the auxiliary regression (4).

(f) **(3 Points)** Explain why $Var(b)$ will be large, in case of mulicollinearity, i.e. when the elements in x are ‘nearly constant’ (so that model (4) will provide a very good fit).

<u>Problem 3</u> (23 Points)

Consider the model:

$$y_t = f(x_t, \beta) + e_t, \qquad f(x_t, \beta) = \beta_1 x_t^{\beta_2}, \qquad e_t \sim N(0,1)\ i.i.d., \qquad t = 1, \dots, T, \tag{5}$$

where $\beta = (\beta_1,\ \beta_2)' \in \mathbb{R}^2$ and $x_t \in \mathbb{R}$ are non–stochastic.

1. **(3 Points)** Show that in model (5) estimating β by minimizing the sum of squared errors is equivalent to maximizing the likelihood function. Hint: Look at the likelihood function $l\,(\beta) = (\beta|y, X)$.

2. **(9 Points)** Derive the Fisher information matrix $I[\beta] = -\mathbb{E}\left[\frac{\partial^2 \ln l(\beta)}{\partial\beta\partial\beta'}\right]$.

   (a) Express the log-likelihood function $\ln l(\beta)$ explicitly as a function of $f(x_t, \beta) = \beta_1 x_t^{\beta_2}$.

   (b) Calculate all 1st and 2nd derivatives of $\ln l(\beta)$ with respect to β_1 and β_2.
   Hint: $\frac{\partial a^{bx}}{\partial x} = ba^{bx} ln(a)$.

   (c) Calculate the expected values of all the second derivatives and give $I[\beta]$.

3. **(9 Points)** Consider the following null hypothesis $H_0 : \beta_2 = 1$. Test this hypothesis by means of a Wald test using the significance level $\alpha = 0.05$. Assume that the maximum likelihood procedure provided the following estimates:

$$\hat{\beta} = (\hat{\beta}_1,\ \hat{\beta}_2)' = (2,\ 1.5)',\ \hat{\Sigma}_{\hat{\beta}} = \widehat{Cov(\hat{\beta})} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}.$$

   (a) Write the null hypothesis in the form $g(\beta) = 0$. Calculate $F(\beta) = \frac{\partial g(\beta)}{\partial \beta'}$.

   (b) Why can $\hat{\Sigma}_{\hat{\beta}}$ be used to calculate the Wald test statistic?

   (c) Calculate the value of the Wald test statistic.

   (d) What is the test decision?

4. **(2 Points)** Which other test would you recommend here? Give reasons for your choice.

<u>Problem 4</u> (28 Points)

Consider the following model:

$$\begin{aligned} y_1 &= \gamma_{21} y_2 + e_1, & (6) \\ y_2 &= \gamma_{12} y_1 + \beta_{12} x_1 + \beta_{22} x_2 + e_2, & (7) \end{aligned}$$

where y_1, y_2 are endogenous and x_1, x_2 are exogenous $(T \times 1)$ vectors. In addition to this the following "sample moments" are given:

$$x_1' y_1 = 2, \quad x_2' y_1 = 3, \quad x_1' y_2 = 4, \quad x_2' y_2 = 8,$$

$$x_1' x_1 = 5, \quad x_1' x_2 = 5, \quad x_2' x_2 = 10.$$

1. **(3 Points)** Write the model in the form $Y\Gamma + XB + E = 0$.
2. **(1 Point)** Which condition must the parameters γ_{21} and γ_{12} satisfy, for Γ to be invertible?
3. **(4 Points)** Assume that the condition above holds for the parameters γ_{21} and γ_{12}. Give the reduced form for equation (6). Express γ_{21} as a function of the reduced form parameters. What do you conclude?
4. **(4 Points)** Check the identification of equation (6) using the order and the rank conditions.
5. **(2 Points)** Explain why equation (7) cannot be identified.
6. **(14 Points)** Estimate equation (6) (i.e. $\delta_1 = \gamma_{21}$) by the 2SLS procedure.
   (a) Which variables appear in Y_1, Y_1^*, X_1, X_1^*, Z_1, X, respectively?
   (b) Rewrite the expression for $\hat{\delta}_{1(2SLS)} = \left[\hat{Z}_1' \hat{Z}_1\right]^{-1} \hat{Z}_1' y_1$ in dependence of X, y_1, y_2 and simplify, so that you can insert the "sample moments".
   (c) Give the numerical value of $\hat{\delta}_{1(2SLS)}$.

<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 the linear model the coefficient of determination will never decrease after the inclusion of an additional regressor.
[ ] [ ]	2. In case of heteroscedasticity the FGLS estimator is always preferable to the OLS estimator.
[ ] [ ]	3. The vector of the fitted values $\hat{y} = (\hat{y}_1, ..., \hat{y}_T)' = Xb$ is orthogonal to the vector of OLS residuals $\hat{e} = (\hat{e}_1, ..., \hat{e}_T) = y - \hat{y}$, i.e. $\hat{y}'\hat{e} = 0$.
[ ] [ ]	4. If there are contemporaneous correlations between the errors of the different equations of a SUR model, SUR-GLS estimator is always more efficient than the OLS estimator.
[ ] [ ]	5. In case of stochastic regressors the OLS estimator is always an unbiased.