Econometrics II, ECO351, Semester I, 2025-26
Homework IV (75 points)

Instructor: M.A. Rahman

Deadline: 6:00 pm, November 11, 2025 (in my office: Room 409, ESB-2).

Please read the instructions carefully and follow them while writing answers.

- *Solutions to homework should be written in A4 size loose sheets. If you are not comfortable writing on white sheets, please ask for biology paper in Tarun Book Store.*
- *Questions should be answered in order as they appear in the homework. Every new question should begin in a new page. Please number all the pages of your homework solution. Please leave a margin of one inch from top and one inch from left. Staple the sheets on the top-left.*
- *Please submit computational assignments (if any) and written answers together and in the correct order. Your answer script should directly address the questions, with all code included in the appendix. All questions in this assignment must be completed using MATLAB. Work done in any other software will NOT be accepted.*

---

1. **(5+5+5 = 15 points)** For the sample selection model presented in the class, show that the marginal effects (evaluated at the sample mean $\bar{x}$) have the following expressions.

(a) For a continuous covariate that appears only on the participation (or hurdle) equation, the covariate effect on outcome variable is,

$$\frac{\partial E[y_2|y_1^* > 0, \bar{x}, \beta, \Sigma]}{\partial(x_{1,j})} = -\beta_{1,j}\,\sigma_{12}\,\frac{\phi(\bar{\omega}_1)}{\Phi(\bar{\omega}_1)}\left[\frac{\phi(\bar{\omega}_1)}{\Phi(\bar{\omega}_1)} + \bar{\omega}_1\right]. \tag{1}$$

(b) For a continuous covariate that appears only on the outcome equation, the covariate effect on outcome variable is,

$$\frac{\partial E[y_2|y_1^* > 0, \bar{x}, \beta, \Sigma]}{\partial(x_{2,j})} = \beta_{2,j}. \tag{2}$$

(c) For a continuous covariate that appears in both participation (or hurdle) and outcome equations, the covariate effect on outcome variable is,

$$\frac{\partial E[y_2|y_1^* > 0, \bar{x}, \beta, \Sigma]}{\partial(x_{1,j}, x_{2,j})} = \beta_{2,j} - \beta_{1,j}\,\sigma_{12}\,\frac{\phi(\bar{\omega}_1)}{\Phi(\bar{\omega}_1)}\left[\frac{\phi(\bar{\omega}_1)}{\Phi(\bar{\omega}_1)} + \bar{\omega}_1\right], \tag{3}$$

where $\bar{\omega}_1 = \bar{x}_1'\beta_1$, and $\phi(\cdot)$ and $\Phi(\cdot)$ denotes the *pdf* and *cdf* of a standard normal distribution, respectively.

2. (4+2+4+5+5+5 = 25 points) Consider the panel data model,

$$y_{it} = x_{it}'\beta + c_i + \varepsilon_{it}, \qquad t = 1(1)T,\ i = 1(1)n, \tag{4}$$

where y_{it} is the dependent variable, x_{it}' is a covariate vector of dimension $(1 \times k)$, β is the parameter vector of dimension $(k \times 1)$, c_i is the individual specific effect, and ε_{it} is the error term. Moreover, define the following: $\tilde{y}_{it} = y_{it} - \bar{y}_i$, $\tilde{x}_{it} = x_{it} - \bar{x}_i$, and $\tilde{\varepsilon}_{it} = \varepsilon_{it} - \bar{\varepsilon}_i$.

Based on the above information and assuming we are interested in estimating a fixed effect model, answer the following.

(a) What are the two assumption on the error term for a fixed effects model? What is the assumption about the relationship between observed covariate vector and individual specific effect?

(b) Write the expression for equation (4), when the model is averaged over t.

(c) Write the panel data model in deviation form by subtracting the expression obtained in Part (b) from equation (4). Utilize the notation $\tilde{y}_{it}$, $\tilde{x}_{it}$, and $\tilde{\varepsilon}_{it}$. Stack the model over t, and express the model in a compact form clearly stating the dimension of all terms.

(d) Write the expression for fixed effects model in deviation form, when stacked over i and t. Clearly specify each term.

(e) The within-group FE estimator is pooled OLS on the transformed regression from Part (d). Derive the fixed effects estimator for β.

(f) Find the estimator for c_i and σ_ε^2.

2. (3+4+5+4+3+8+8 = 35 points) `Panel Data Model`:Consider the data present in the file "`Murder.xlsx`", where the sheet "`Data`" contains data and the sheet "`Desc`" contains the definition of variables. Using this state-level data (from the US) on murder rates and executions answer the following questions.

(a) Consider the panel data model,

$$\texttt{mrdrte}_{it} = \eta_t + \beta_1 \texttt{exec}_{it} + \beta_2 \texttt{unemp}_{it} + c_i + \varepsilon_{it}, \tag{5}$$

where η_t simply denotes different year intercepts and c_i is the unobserved state effect. If past executions of convicted murderers have a deterrent effect, what should be the sign of β_1? What sign do you think β_2 should have? Explain.

(b) Using just the years 1990 and 1993, estimate the equation from Part (a) by pooled OLS. Ignore the serial correlation problem in the composite errors. Do you find any evidence for a deterrent effect?

(c) Now using 1990 and 1993, estimate the equation via fixed effects. You may use first differencing because you are only using two years of data. Is there evidence of a deterrent effect? How strong?

(d) Compute the heteroskedastic robust standard error for the estimation in Part (b). Compare this with the non-robust standard error and comment on the significance.

(e) Find the state that has the largest number for the execution variable in 1993 (the variable `exec` is total executions in 1991, 1992, and 1993). How much bigger is this value than then next highest value?

(f) Estimate the equation using first differencing, dropping Texas from the analysis. Compute the usual and heteroskedastic robust standard errors, Now, what do you find? What is going on?

(g) Use all three years of data and estimate the model by fixed effects (within estimation). Include Texas in the analysis. Discuss the size and statistical significance of the deterrent effect compared with only using 1990 and 1993.