COLORADO STATE UNIVERSITY
DEPARTMENT OF AGRICULTURAL & RESOURCE ECONOMICS

Exam 2 **Agricultural & Resource Economics / Economics 535** *S.R. Koontz*
Fall 2021 **Applied Econometrics**

This exam is worth 100 points. The points allocated to each question are listed with the question number. Use your time appropriately. Show the work necessary to communicate clear answers each question. Good answers always communicate.

1. (30 pts.) You are interested in estimating a model with cross sectional data. The general model is

$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_k x_{ki} + e_i \qquad i = 1,\ldots,N.$$

Heteroskedasticity is the most likely violation of OLS assumptions with this data.

a. What is heteroskedasticity and how does it impact OLS estimates?

b. Outline the Breusch-Pagan test for heteroskedasticity. State the null hypothesis and the distribution of the test statistic under the null.

c. You determine your software will report White's heteroskedasticity consistent standard errors. Outline what this is.

d. You determine your software will perform weighted least squares. Discuss how this can be used to address the problem.

e. You determine your software will estimate following along with the model at the beginning of this question

$$\sigma^2_i = \exp(\alpha_0 + \alpha_1 x_{1i} + \ldots + \alpha_p x_{pi})$$

Outline what this alternative model is.

f. Between c., d. and e. above, how do you choose?

2. (20 pts.) A colleague of yours is interested in estimating a model of an expenditure function. Your colleague has surveyed a sample of consumers, asked about expenditures, and asked about characteristics that describe their socio-economic background and values. The model is specified as follows

$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_k x_{ki} + e_i \qquad i = 1,\ldots,N.$$

However, 40% of the sample had zero expenditures. The remainder had positive expenditures. Furthermore, the response rate for the survey was 50% – 50 out of every 100 surveys were returned with usable data.

First, outline statistical issues with this model. Second, outline your estimation options.

3. (25 pts.) Another of your colleagues is interested in estimating a model of how economic output across developing countries is impacted by foreign direct investment (FDI). The basic model is specified as follows

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \ldots + \beta_k x_{kit} + e_{it}$$

where $i = 1, \ldots, N$ and $t = 1, \ldots, T$.

The sample are a combined cross-section and time-series of data. The dependent variable is the GDP growth rate. The first independent variable is FDI growth rate and the last variable is, for example, country population growth rate. Two alternative models are estimated

(1) $y_{it} = \beta_0 + \beta_1 x_{1it} + \ldots + \beta_k x_{kit} + (\Sigma \alpha_i D_i) + e_{it}$

(2) $y_{it} = \beta_0 + \beta_1 x_{1it} + \ldots + \beta_k x_{kit} + (\Sigma \theta_t D_t) + e_{it}$

where D is a dummy variable and the subscript denotes the dimension of the panel with which the variable is associated.

a. Describe what the authors are allowing the conditional mean (the regression) to do in these alternative models. Why might this be important?

b. How would you interpret β_k across the two alternative models? Your colleague is surprised that the two results are rather different. Can you offer a reason why the results might be rather different?

c. Your colleague settles on (1) alternative model but recognizes there still might be a concern with time impacting observed economic output. You suggest using Generalize Least Squares and the weighting matrix as follows

$$\Omega = \begin{bmatrix} \sigma^2 I_T & \sigma_v^2 I_T & \cdots & \sigma_v^2 I_T \\ \sigma_v^2 I_T & \sigma^2 I_T & \cdots & \sigma_v^2 I_T \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_v^2 I_T & \sigma_v^2 I_T & \cdots & \sigma^2 I_T \end{bmatrix}$$

where $e_{it} = v_t + w_{it}$ and $V(e_{it}) = \sigma^2 = \sigma^2_v + \sigma^2_w$. Describe the issue that this weighting scheme addresses.

d. Your colleague next wants to know if you are sure and if there a way to test if the approach in c. above is appropriate. If yes, name the test and state the null.

4. (15 pts.) Probit and Logit are qualitative choice models. These are models where the dependent variable is a dummy variable. Below is an example model

$$z_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + e_i \qquad i = 1,\ldots,N$$

where z_i is an index that captures when the dependent variable equals 0 or 1.

a. How is the change in probability of observing the choice given a change in x_1 measured? How is the marginal interpreted?

b. How is the probability of observing the choice given a specific set of x's calculated?

c. How is the goodness-of-fit measured in Probit and Logit models?

5. (10 pts.) While conducting a literature review for your research, you read an online article in which the authors report the following OLS regression results using a sample of data over time. The article concludes the work is important because a variable (x_3) that has not been examined before – and is suggested by a new emerging theory and literature – is important in explaining the economic behavior in the dependent variable.

$$y_t = \underset{(11.50)}{95.05} + \underset{(0.102)}{1.596}\, x_{1t} + \underset{(0.142)}{0.982}\, x_{2t} - \underset{(1.475)}{2.985}\, x_{3t} + \underset{(0.015)}{0.818}\, y_{t-1} + e_t$$

$R^2 = 0.749$ $F_{4,44} = 56.62$ $\sigma = 3.724$ $T = 50$

Adjusted $R^2 = 0.811$ DW-d = 0.989 Durbin-h = 5.83

Standard errors are in parentheses below the coefficients. Do you agree with the conclusion of the authors? Be precise.