Market and Information Economics
Preliminary Examination

Department of Agricultural Economics
Texas A&M University

May 2021

Instructions: This examination consists of six questions. You must answer the first question and you must answer four of the remaining five questions (i.e. answer four of the questions numbered 2-6). Each question answered (five in total) has a weight of 20% in the final examination score. Please read through the entire examination before making a decision on the particular set of five questions you actually answer. The examination proctor will review the content of the exam at the beginning of the time period (9:00 am). He or she will answer general questions for the entire set of students writing this prelim. You have until 1:15 pm to complete the exam. Good Luck!

You Must Answer this Question

1. A restaurant has three types of customers. A third of its customers, Type A, are willing to spend \$5 on an appetizer but only \$2 on a dessert. Another third, Type B, are willing to spend \$3.50 on an appetizer and \$3.50 on a dessert. The remaining third, Type C, are willing to spend only \$2 on an appetizer but \$5 on a dessert. All three types are willing to spend \$10 on the main course. It costs the restaurant a constant \$2 to prepare an appetizer or a dessert, and \$7 to prepare the main course. Which is optimal for the restaurant, to offer appetizers and desserts à la carte (with separate prices on the menu), or to offer them only as a complete meal, tied in with the main course?

Answer four of the following five questions

2. Consider the following panel data model

$$\text{Income}_{i,t} = \beta_0 + \beta_1 \log(\text{age}_{i,t}) + \beta_2 \text{education}_i + \beta_3 \text{gender}_i + c_i + u_{i,t},$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$, where $\text{gender}_i = 1$ if the ith individual is male, and zero otherwise, and c_i is a time invariant individual effect.

(a) Suppose c_i is correlated with some of the covariates. What are the plausible sources of the correlation? Shall we use the fixed effect or random effect estimator in this case? Explain your choice.

(b) If the fixed effect estimator is used, indicate which coefficient(s) among $\beta_0, \beta_1, \beta_2, \beta_3$ are identified.

(c) Suppose the fixed effect estimator is employed. How does one estimate possibly time-varying effect of education. Explain how to test if the effect of education is indeed time varying.

(d) Suppose the variance of $u_{i,t}$ increases with the level of education. Propose an estimator that can mitigate this heteroscedasticity.

3. An analyst estimates the following specification over the period 2011.1 to 2020.4 (quarterly data from 2011 to 2020) using OLS:

$$ln(SALES) = 4 - 150P + 25PS + 0.02Q_1 - 0.06Q_2 - 0.08Q_3$$

Q_1, Q_2, and Q_3 are quarterly dummy variables to account for seasonality; P denotes own-price, and PS denotes the price of another product.

In addition, suppose that the sample means of the respective variables over this time period are as follows:

$SALES$=\$30; P=\$0.005/unit; and PS=\$0.003/unit.

(a) On the basis of this information, what is the own-price elasticity of demand?

(b) Is the demand for this product elastic? What would you recommend as pricing strategy to maximize sales?

(c) Obtain the cross-price elasticity. Interpret your answer.

(d) Characterize the level of sales in quarter 3 relative to quarter 4.

(e) Forecast the level of sales for quarter 2 when P=\$0.005/unit and when PS=\$0.003/unit.

(f) What additional explanatory variable(s) would you add, if any, to the specification of this demand model? Defend your answer.

(g) Suppose that the regression sum of squares is equal to 90, and that the error or residual sum of squares is equal to 30. What is the adjusted R^2 for this model?

(h) Suppose that the Jarque-Bera test statistic associated with the residuals has a p-value of 0.45. The partial autocorrelation function for the residuals at lags 1 and 2 are statistically significant. All other partial correlation coefficients are not statistically different from zero. What can be said about the behavior of the residuals?

(i) From your answer in (h), what would you recommend econometrically?

(j) Let h_t represent the variance of the error or disturbance term. Suppose that $h_t = 3+0.25(e_{t-1})^2+0.45h_{t-1}$, where e_{t-1} represents a 1-period lag of the residuals. What can be concluded from the information about the residual variance of this model? Assume that each of the estimated coefficients associated with the residual variance are statistically different from zero.

4. A marketing research consulting firm hired you to estimate a model for consumer willingness-to-pay (WTP) for apples produced in the state of Washington. They are specifically interested in a model that estimates $WTP = f(X_s, X_m, X_l)$, where X are apple attributes. X_s is the size and it is equal to 1 if the size is large, and 0 otherwise; X_m is the production method, and it is equal to 1 if it is organic, and 0 otherwise; and X_l is the local designation and it is equal to 1 if produced in Washington, and 0 otherwise. The marketing firm launched a survey where each participant stated their WTP for four apple products: 1,2,3,4. See Table 1 below. They are interested in estimating the following model:

$$Y_i = \beta_0 + \beta_1 X_s + \beta_2 X_m + \beta_3 X_l + \beta_4 X_s X_m + \beta_5 X_s X_l + \beta_6 X_m X_l + \beta_7 X_s X_m X_l + \epsilon_i$$

Table 1: List of products in the survey

Product	X_s	X_m	X_l
1	0	0	1
2	0	1	0
3	1	0	0
4	0	0	0

In order to get credit, please make sure you provide an explanation to support your answers.

(a) Which parameters in the model are identified (i.e. which parameters can be estimated)?

(b) If the WTP is elicited without market transactions (i.e. no real purchases) the measurement of the dependent variable, WTP may be biased. Explain what type of measurement error bias results in this setup? In which direction would WTP be biased?

(c) Propose two different ways to overcome the problems of the bias above.

(d) In addition to the apple attributes, which other variables would you include in the model?

5. Consider a Bayesian linear regression model

$$y = X\beta + u$$

where y is an $n \times 1$ vector, X is an $n \times k$ data matrix, β is a $k \times 1$ vector, and u are independent and identically normally distributed random variables: $u \sim N(0, \sigma^2)$. For simplicity, let $\theta = (\beta, \sigma^2)$. Conventionally, the posterior distribution of θ is expressed as the product of likelihood function and the prior distribution:

$$p(\theta|y, X) = \frac{p(y|\theta, X)p(\theta)}{p(y|X)} \propto p(y|\theta, X)p(\theta).$$

(a) What are the major differences between the classical OLS and Bayesian estimation?

(b) How the choice of prior would impact the estimation results when the sample size is small? How about when sample size is sufficiently large?

(c) Assuming non-informative priors for β and σ^2, derive the posterior joint density function of β and σ^2 for estimation purpose. Note that you only need to derive an expression that is proportional to the exact posterior density.

(d) Propose two methods to obtain the posteriors joint distribution of β and σ^2. Briefly explain how you are going to implement these proposed methods in R.

6. Consider an i.i.d. sample $\{Y_i, X_i\}, i = 1, \ldots, N$, where Y_i takes the value of either zero and one, and $X_i = (X_{i,1}, X_{i,2}, X_{i,3})$.

   (a) Describe how to formulate a Gaussian Process Classification estimator for this problem.

   (b) Since this estimator does not have a tractable solution, the Laplace approximation is typically used. Explain how to implement this approximation.

   (c) Some of the covariates may not contribute to the classification of Y. Explain how to determine the relevance of the covariates.