University of Pennsylvania **Economics 705, Spring 2017**

Prelim Examination

Friday June 9, 2017. Time limit: 150 minutes

Instructions:

(i) The exam consists of two parts. The total number of points for each part is 35. The number of points for each question is given below.

(ii) The exam is closed book and closed notes.

(iii) To receive full credit for your answers you have to explain your calculations. **You may state additional assumptions.**

Part I

Question 1: TRUE or FALSE Questions (12 Points)

You need to explain your answers to get full credit.

(i) (2 Points) TRUE or FALSE? Suppose that random variables X and Y are jointly distributed with density $p(x, y)$. Then $\mathbb{E}[X] = 0$ implies $\mathbb{E}[X|Y = y] = 0$ for each y.

(ii) (3 Points) TRUE or FALSE? To test the joint null hypothesis $H_0 : \beta_1 = \beta_2 = 0$ against the alternative $H_1 : \beta_1 \neq 0$ or $\beta_2 \neq 0$, an econometrician decides to conduct two separate t tests for $H_{0i} : \beta_i = 0$ and to reject the joint hypothesis H_0 if at least one of the two t-tests rejects. Assuming each t-test has a type 1 error of 5%, the type 1 error of the resulting joint test is exactly 10%.

(iii) (3 Points) TRUE or FALSE? Consider the location-shift experiment $Y \sim N(\theta, 1)$, $\theta \sim N(0, 1/\lambda)$. In this experiment a 95% Bayesian credible interval is shorter than a 95% frequentist confidence interval.

(iv) (4 Points) TRUE or FALSE? The presence of heteroskedasticity in an error term of a regression renders the OLS estimator inconsistent.

Question 2: Likelihood Estimation of a Mixture Model (23 Points)

Consider the following location-shift model:

$$Y_i = \theta_0 + U_i, \quad U_i \sim \begin{cases} N(0,1) & \text{with prob. } 1/2 \\ N(0,4) & \text{with prob. } 1/2 \end{cases} . \tag{1}$$

(i) (3 Points) Derive the likelihood function $\mathcal{L}(\theta|Y)$ for an *iid* sample of n observations.

(ii) (4 Points) Denote the log likelihood function as $l(\theta|Y)$. Derive the maximum likelihood estimator defined as $\hat{\theta} = \text{argmax}_{\theta \in \mathbb{R}} \, l(\theta|Y)$. Provide an explicit formula for $n = 1$ and an implicit formula for $n > 1$.

(iii) (4 Points) Show that the score $s(\theta) = l'(\theta|Y)$ evaluated at $\theta = \theta_0$ has mean zero.

(iv) (2 Points) For $n = 1$ compute the mean and the variance of $\hat{\theta}$.

(v) (2 Points) What is the Cramer-Rao lower bound?

(vi) (8 Points) For $n = 1$ verify whether $\hat{\theta}$ attains the Cramer-Rao lower bound in model (1).

Part II

Question 3: Omitted Variable Problem (24 Points)

(i) (2 Points) Consider the linear model

$$Y_i = X_i\theta_0 + U_i, \text{ where } E(X_iU_i) = 0 \text{ and } \theta_0 \in R.$$

Suppose we observe i.i.d. data $\{(Y_i, X_i) : i = 1, ..., n\}$. Provide an estimator of θ_0 and show its asymptotic distribution.

(ii) (2 Points) Now suppose the true model is

$$Y_i = X_i\theta_0 + D_i\gamma_0 + U_i, \text{ where } E(X_iU_i) = E(D_iU_i) = 0$$

and $\theta_0, \gamma_0 \in R$. Suppose you observe i.i.d. data $\{(Y_i, X_i, D_i) : i = 1, ..., n\}$. Provide an estimator of θ_0 and show its asymptotic distribution.

(iii) (2 Points) Suppose the true model is as in part (ii) with $\gamma_0 \neq 0$. However, D_i is unobserved. With i.i.d. data $\{(Y_i, X_i) : i = 1, ..., n\}$, if you estimate θ_0 with the estimator in part (i), what is the asymptotic limit of this estimator given that D_i is omitted? Discuss conditions under which this estimator is consistent.

(iv) (2 Points) Suppose the condition under which the estimator in part (iii) is consistent holds. Compare the estimators in part (ii) and part (iii). Which one has a smaller asymptotic variance?

(v) (2 Points) Again, assume the true model is as in part (ii) and D_i is unobserved. Now you want to obtain a consistent estimator of θ_0 by using instruments Z_i. Provide conditions on Z_i such that they are exogenous and relevant.

(vi) (6 Points) Provide an efficient GMM estimator of θ_0 with the instrument Z_i and show its asymptotic distribution.

(vii) (4 Points) How to test whether the instruments are exogenous. Be specific about the test statistic and how to obtain the critical value.

(viii) (4 Points) Provide a test for $H_0 : \theta = \theta_0$ vs $H_0 : \theta \neq \theta_0$ that is robust to weak instruments. Be specific about the test statistic and how to obtain the critical value.

Question 4: Limited Dependent Variable Model (11 Points)

Consider the model

$$Y_i^* = g(X_i, \theta_0) + U_i,$$

where X_i and U_i are independent, $U_i \sim N(0, \sigma^2)$, and $g(x, \theta_0)$ is a nonlinear function with unknown parameter θ_0.

(i) (4 points) Suppose we observe i.i.d. data $\{(X_i, Y_i\}, i = 1, ..., n\}$ with

$$Y_i = \begin{cases} 1 & Y_i^* \geq 0 \\ 0 & Y_i^* < 0. \end{cases}$$

Write down the log likelihood function for the estimation of θ_0.

(ii) (3 points) Suppose $g(X_i, \theta_0) = X_i' \theta_0$, can we identify θ_0 in part (i)? Explain.

(iii) (4 points) Suppose we observe i.i.d. data $\{(X_i, Y_i\}, i = 1, ..., n\}$ with

$$Y_i = \begin{cases} Y_i^* & Y_i^* \geq 0 \\ 0 & Y_i^* < 0. \end{cases}$$

Write down the log likelihood function for the estimation of θ_0.