DEPARTMENT OF ECONOMICS
UNIVERSITY OF CALIFORNIA, BERKELEY

Problem Set #3

ECONOMICS 240B
SPRING 2006

Due May 3

PART I: "Theoretical" Questions

Turn in (correct) answers to the following exercises from Ruud's text:

Chapter 17: Exercises 17.1, 17.6, 17.9.

Extra Questions:

1. Suppose that the restricted ML estimator $\tilde{\theta}$ is such that $\tilde{\theta} \xrightarrow{p} \theta_*$ where $\theta_* \neq \theta_0$ and $E[L_\theta(\theta_*)] \neq 0$. Stating appropriate assumptions sketch a proof that the likelihood ratio and score (Lagrange multiplier) tests define consistent tests for the null hypothesis $H_0 : r(\theta_0) = 0$ against the alternative hypothesis $H_1 : r(\theta_0) \neq 0$.

2. Consider the conditionally heteroskedastic normal linear regression model:

$$y = x'\beta_0 + u, u|x \sim N(0, \sigma_0^2(x)).$$

A random sample of observations $\{y_i, x_i\}_{i=1}^n$ on (y, x) is available. Let $\sigma_0^2(x) = h(\delta_{00} + \delta_{10}' z(x))$ where $h(\cdot)$ is a non-negative valued function such that $h'(0) \neq 0$. It is desired to test the null hypothesis $H_0 : \delta_{10} = 0$ against the alternative hypothesis $H_1 : \delta_{10} \neq 0$.

(a) Show that the score vector under the null hypothesis $H_0 : \delta_{10} = 0$ is given by

$$\begin{aligned} E_n[L_\beta(\beta_0, \delta_{00}, 0)] &= \frac{1}{\sigma_0^2(\delta_{00})} E_n[xu], \\ E_n[L_\delta(\beta_0, \delta_{00}, 0)] &= \frac{h'(\delta_{00})}{2\sigma_0^4(\delta_{00})} E_n[\begin{pmatrix} 1 \\ z(x) \end{pmatrix} (u^2 - \sigma_0^2(\delta_{00}))]. \end{aligned}$$

(b) Show that the conditional information matrix under the null hypothesis $H_0 : \delta_{10} = 0$ is

$$\mathcal{I}(\beta_0, \delta_{00}, 0|x) = E_n[\begin{pmatrix} \frac{1}{\sigma_0^2(\delta_{00})} xx' & 0 \\ 0' & \frac{h'(\delta_{00})^2}{2\sigma_0^4(\delta_{00})} \begin{pmatrix} 1 \\ z(x) \end{pmatrix} \begin{pmatrix} 1 \\ z(x) \end{pmatrix}' \end{pmatrix}].$$

(c) Hence, or otherwise, show that the score statistic for testing the null hypothesis H_0 : $\delta_{10} = 0$ against the alternative hypothesis $H_1 : \delta_{10} \neq 0$ is given by

$$\mathcal{S} = \frac{1}{2} n E_n[\left(\frac{\hat{u}^2}{\hat{\sigma}^2} - 1\right) \begin{pmatrix} 1 \\ z(x) \end{pmatrix}'] E_n[\begin{pmatrix} 1 \\ z(x) \end{pmatrix} \begin{pmatrix} 1 \\ z(x) \end{pmatrix}']^{-1} E_n[\begin{pmatrix} 1 \\ z(x) \end{pmatrix} \left(\frac{\hat{u}^2}{\hat{\sigma}^2} - 1\right)]$$

where $\hat{u}_i = y_i - x_i'\hat{\beta}$, $(i = 1, ..., n)$, and $\hat{\beta}$ and $\hat{\sigma}^2$ denote the H_0-ML estimators for β_0 and $\sigma_0^2 = \sigma_0^2(\delta_{00})$.

(d) What is the relationship between $\mathcal{S}$ and the Breusch-Pagan statistic for testing for heteroskedasticity? Justify your answer.

3. Consider the binary choice model with latent variable y^* determined by the normal linear regression model:

$$y^* = x'\beta_0 + \sigma_0(x)\varepsilon, \varepsilon|x \sim N(0,1),$$

where $\sigma_0^2(x) = h(\delta_0' z(x))$ where $h(\cdot)$ is a non-negative valued function such that $h(0) = 1$ and $h'(0) \neq 0$, and observability rule

$$y = \mathbb{I}(y^* > 0),$$

where the indicator function $\mathbb{I}(A) = 1$ if A and 0 otherwise. A random sample of observations $\{y_i, x_i\}_{i=1}^n$ on (y, x) is available. It is desired to test the null hypothesis $H_0 : \delta_0 = 0$ against the alternative hypothesis $H_1 : \delta_0 \neq 0$.

(a) Show that the score vector under the null hypothesis $H_0 : \delta_0 = 0$ is given by

$$\begin{aligned} E_n[L_\beta(\beta_0, 0)] &= E_n[x\varepsilon^{(1)}], \\ E_n[L_\delta(\beta_0, 0)] &= -\frac{h'(0)}{2} E_n[z(x)(x'\beta_0)\varepsilon^{(1)}], \end{aligned}$$

where

$$\varepsilon^{(1)} = y\frac{\phi(x'\beta_0)}{\Phi(x'\beta_0)} - (1-y)\frac{\phi(x'\beta_0)}{1-\Phi(x'\beta_0)}.$$

(b) Suggest a score statistic for testing the null hypothesis $H_0 : \delta_0 = 0$ against the alternative hypothesis $H_1 : \delta_0 \neq 0$.

PART II: "Empirical" Question

Reconsider the Probit model of Problem Set #1 relating *arr*86 to *pcnv*, *avgsen*, *tottime*, *ptime*86, *inc*86, *black*, *hispan*, and *born*60 based on the data file `crimerv.txt`.

(a) Test the joint significance of *avgsen* and *tottime* using a likelihood ratio statistic.

(b) Add the squares of *pcnv*, *ptime*86 and *inc*86 to the Probit model. Test whether they are individually or jointly significant using a Wald statistic.

(c) Construct a score test for the omission of the squares of *pcnv*, *ptime*86 and *inc*86 from the Probit model. Compare your result to those in (b).

(d) Construct a score test against heteroskedasticity with respect to the included covariates *pcnv*, *avgsen*, *tottime*, *ptime*86, *inc*86, *black*, *hispan*, and *born*60 for the Probit model. What conclusions do you draw?