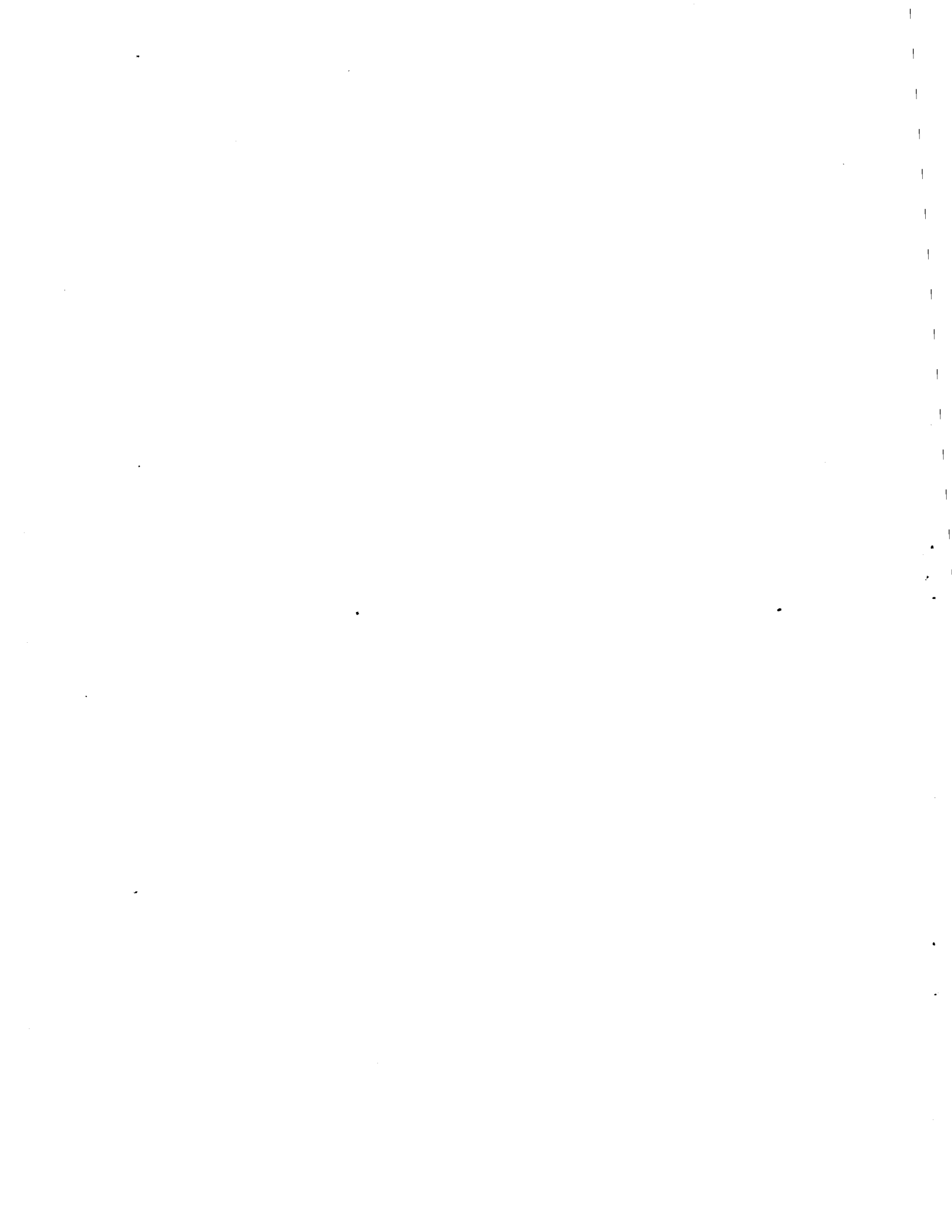


A SPECIFICATION TEST FOR NORMALITY  
IN THE GENERALIZED CENSORED REGRESSION MODELS

by  
Lung-Fei Lee

Discussion Paper No. 81 - 146, May 1981

Center for Economic Research  
Department of Economics  
University of Minnesota  
Minneapolis, Minnesota 55455



A SPECIFICATION TEST FOR NORMALITY  
IN THE GENERALIZED CENSORED REGRESSION MODELS

By

Lung-fei Lee

Department of Economics

University of Minnesota and University of Florida

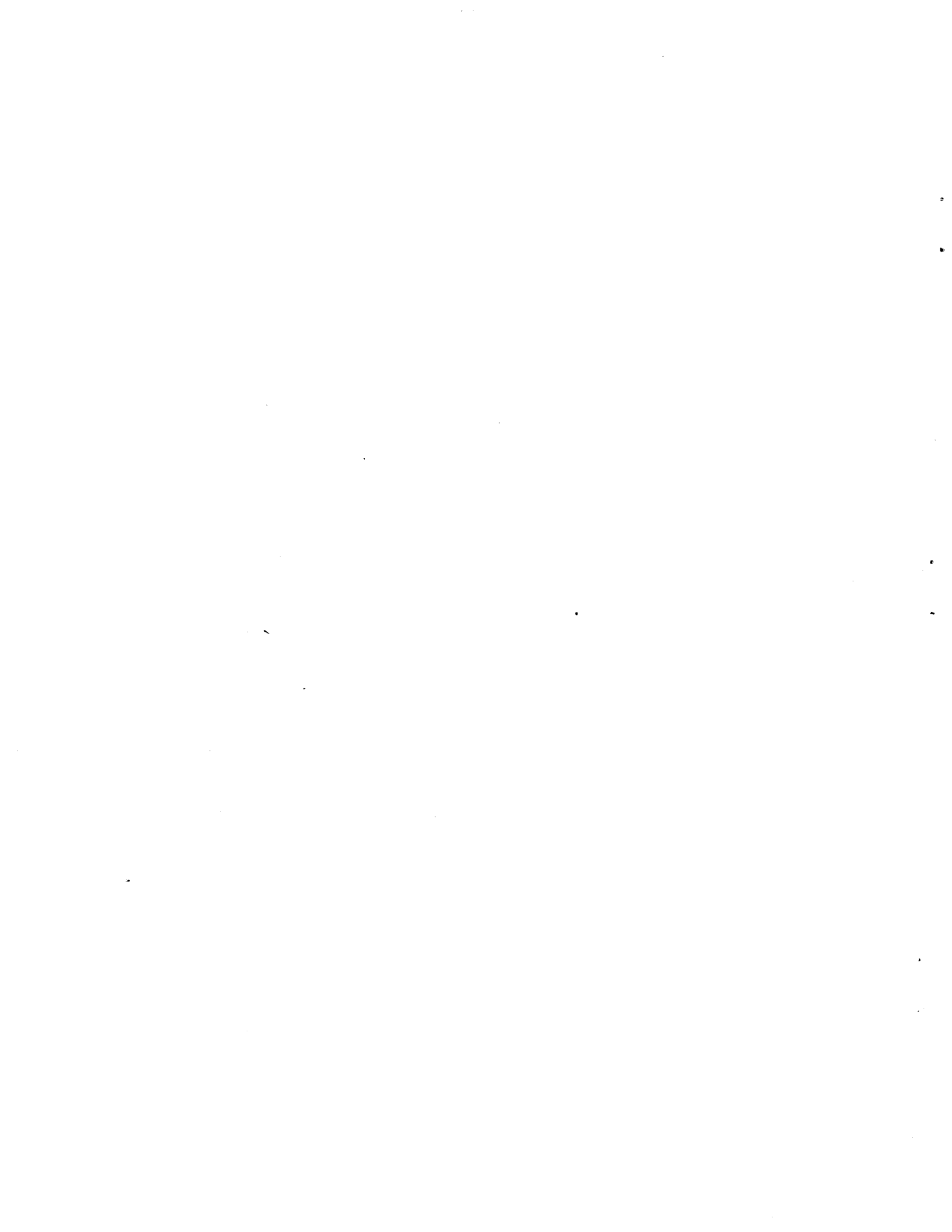
ABSTRACT

Based on the Pearson family of distributions, we have derived some Lagrangean multiplier tests for the normality and homoscedasticity assumptions in the censored regression models. The Lagrangean multiplier test statistic for the joint test of selectivity bias, homoscedasticity and normality is the sum of three components. Each component is shown to be a conditional Lagrangean multiplier test statistic. It has been shown that they can also be interpreted as tests of significance of coefficients in some linear models based on instrumental variable estimations. We have pointed out that for some very special cases, the Lagrange multiplier tests for selectivity have had no power, and are not equivalent, for large samples, to the likelihood ratio tests. This situation occurs as the likelihood evaluated at the constrained MLE is a stationary value of the unconstrained likelihood function. These examples provide, probably, the first set of examples to confirm a conjecture in Silvey [1959, p. 399].

---

Address: (valid until June 30, 1981)  
Department of Economics  
University of Florida  
Gainesville, Florida 32611, U.S.A.

(valid after June 30, 1981)  
Department of Economics  
University of Minnesota  
271 19th Avenue S.  
Minneapolis, Minnesota 55455, U.S.A.



A SPECIFICATION TEST FOR NORMALITY  
IN THE GENERALIZED CENSORED REGRESSION MODELS

By

Lung-fei Lee\*

1. INTRODUCTION

The recent developments on the econometric models with censored dependent variables attempt to deal with the problems of systematic missing data on the dependent variables for census survey data. The most common cases are the existence of some selection processes which determine the observed samples on the dependent variables. Conditional on the appropriate set of exogenous variables, if the dependent variable in a regression model is correlated with the selection processes, conventional estimation techniques which ignore the censoring will not provide consistent estimates of the parameters in the model. The common solution in the literature is to specify a joint probability distribution on the random elements in the selection processes, which are modeled as a probabilistic discrete choice model, and the regression model. Multivariate normal distribution is the most commonly specified assumption in those models. If the distributional assumption were correct, the maximum likelihood method would be consistent and asymptotically efficient under very general conditions. A rigorous proof of these large sample properties can be found in Amemiya [1973] for the Tobit model. Under the normality assumption, computationally simple

limited information methods have been developed in Amemiya [1974], Heckman [1976], Lee [1979], among others.

The normal distributional assumption is a crucial assumption in the model specification and in the development of the limited information estimation methods cited above. As contrary to the standard linear models, the misspecification of normality of the disturbances in the censored regression model and the Tobit model will, in general, provide inconsistent estimates of the parameters under both the maximum likelihood and the limited information methods. The main reason underlying the inconsistency of the estimators is that the probability of a sample being noncensored and its conditional mean depend on the functional form of the specified distribution. The misspecification of the distribution of the disturbances is similar, in a loose sense, to the misspecification of functional form in a nonlinear regression model. Theoretical and numerical evidences on the consequences of misspecification of normal distribution have been provided in the simple Tobit-type models in Goldberger [1980]. For the censored regression models, some investigations are in Olsen [1979]. The misspecification of the distributions is not the only source of misspecification in those models. Heteroscedasticity disturbances will also create inconsistent estimates when they are misspecified to be homoscedastic, as demonstrated in Hurd [1979], Maddala and Nelson [1975], and Nelson [1981] for the Tobit-type models. The consequences of these two types of misspecifications are serious for those models.

Since the misspecifications of normal distribution and homoscedasticity can create misleading estimates, these assumptions need to

be tested rigorously so that investigators can have more confidence in the fitted models. One can easily suggest some tests of homoscedasticity or incorporate heteroscedastic errors into the estimation procedure if the heteroscedastic variances can be parameterized as linear or non-linear regression function with finite number of unknown parameters. The test of normality seems to be a more difficult problem and needs more attention. It is the purpose of this article to provide a large sample test for the normality assumption in the censored regression model. Our test is based on the Lagrange Multiplier test principles as developed in Aitchison and Silvey [1958] and Silvey [1959]. The Lagrange Multiplier method is attractive since the basic model that we will consider is too general to be computationally tractible under the maximum likelihood method. Our test, however, can be easily implemented. The test of homoscedasticity can also be incorporated into the testing procedure in addition to the test of normality. Our approach can be modified to test normality and heteroscedasticity for related econometric models such as the probit and Tobit models and the stochastic frontier production function model in Aigner, Lovell and Schmidt [1979].

Our paper is organized as follows. In Section 2, we will specify a basic model which is quite general and contains the model with normal distributions as special case. In section 3, we derive the Lagrangean multiplier tests for the hypotheses of normality and homoscedasticity. In section 4, we analyze the Lagrangean multiplier tests for the selectivity bias, heteroscedasticity and normality. Some interpretations of the statistic and its relations with some conditional

tests are provided. We will point out that in an unusual case, the derived Lagrange multiplier test has no power and should not be used. This case may be the first example available to confirm to the conjecture in Silvey [1959] that there may be cases for which the maximum likelihood ratio principle is applicable but not the Lagrangian multiplier principle. In an appendix, we discuss briefly the modification of our test to the probit and Tobit models and provide the tests for those models. Finally, we draw our conclusions.

## 2. A BASIC CENSORED REGRESSION MODEL

Consider the simple two equation censored model

$$y_{1i} = x_i\beta + u_i \quad i=1, \dots, N \quad (2.1)$$

$$y_i^* = z_i\gamma - \varepsilon_i \quad i=1, \dots, N \quad (2.2)$$

where  $x$  and  $z$  are exogenous variable vectors of dimensions  $k_1$  and  $k_2$ , respectively,  $E(u|x,z) = 0$ ,  $E(\varepsilon|x,z) = 0$  and  $\text{var}(\varepsilon|x,z) = 1$ . The disturbances  $(u_i, \varepsilon_i)$  are independent for different  $i$ . The dependent variable  $y^*$  is unobservable but has a dichotomous observable indicator  $I$  which is related to  $y^*$  as follows:

$$I = 1 \quad \text{if and only if} \quad y^* \geq 0,$$

$$I = 0 \quad \text{if and only if} \quad y^* < 0.$$

The dependent variable  $y_1$  conditional on  $x$  and  $z$  has well-defined marginal distribution but  $y_1$  is not observed unless  $y^* \geq 0$ . The observed



samples  $y_i$ 's of  $y_1$  are thus censored. The exogenous variables are observable for each  $i$ . We assume further that the exogenous variables are uniformly bounded, its empirical distribution converges to a limiting distribution, the parameter space is compact and the true parameter vector is an interior point. These regular conditions are general enough to justify the asymptotic properties of the estimators and the tests that will be considered, see, e.g., Amemiya [1973] and Silvey [1959].

The popular distributional hypothesis for this model is that the disturbances are homoscedastic and the  $u_i$  and  $\epsilon_i$  are bivariate normally distributed. Some tests for normality have been suggested in Huang and Bolch [1974], White and MacDonald [1980] and Bera and Jarque [1980] in the linear regression models. These tests are, however, not applicable for our model since the observed samples of  $y_1$  are censored and the corresponding residuals  $u_i$  will not be normally distributed even though the marginal distributions of  $u_i$  are normal.

To provide a test of normality and homoscedasticity of the disturbances, we assume that the true distribution of  $u$  is a member of the general Pearson family of distributions. The general probability density function  $g(u)$  of this system satisfies a differential equation of form

$$\frac{1}{g(u)} \frac{dg(u)}{du} = \frac{a + u}{b_0 + b_1 u + b_2 u^2} \quad (2.3)$$

The density function  $g(u)$  can be solved from the above differential equation as

$$g(u) = \exp \left[ \int \frac{a+u}{b_0 + b_1 u + b_2 u^2} du \right] / \int_{-\infty}^{\infty} \exp \left[ \int \frac{a+t}{b_0 + b_1 t + b_2 t^2} dt \right] dt, \\ -\infty < u < \infty. \quad (2.4)$$

The denominator in (2.4) guarantees that the density function is proper.<sup>1</sup> This general distribution contains distributions with various shapes, in particular, it contains the normal, t, beta and gamma distributions as special cases (see, e.g., Elderton and Johnson [1969], Chapter 4, or Johnson and Kotz [1970], pp. 9-15). Since the disturbance  $u$  is specified to have zero mean, it implies that  $a = -b_1$  and the distribution is shifted to have mean at zero. After some reparameterizations, the density function with zero mean can be rewritten as

$$g(u) = \exp \left[ \int \frac{c_1 - u}{c_0 - c_1 u + c_2 u^2} du \right] / \int_{-\infty}^{\infty} \exp \left[ \int \frac{c_1 - t}{c_0 - c_1 t + c_2 t^2} dt \right] dt \\ (2.5)$$

This density function becomes the normal density function when  $c_1 = 0$  and  $c_2 = 0$ . To test normality for our model, we need to specify not only the marginal distribution of  $u$  but also the joint distribution of  $u$  and  $\epsilon$ . Theoretically, it seems desirable to consider the bivariate generalization of the Pearson system which contains the bivariate normal distribution as special case. Unfortunately, this generalized system is too complicated to be useful because the parameters in the system must satisfy a highly nonlinear relationship and cannot be chosen completely arbitrarily. For the detail discussions, one can consult Elderton and Johnson [1969, pp. 137-139] or Johnson and Kotz [1972, pp. 6-8].

As an alternative, we suggest the following joint distribution which is derived from the translation method in Mardia [1970] with given marginal distributions of  $u$  and  $\epsilon$ . Suppose the marginal distributions of  $u$  and  $\epsilon$  are specified as  $G(u)$  and  $F(\epsilon)$ , respectively. Let  $\Phi(\cdot)$  denote the standard normal distribution function. The random variables  $\epsilon$  and  $u$  can be transformed to the standard normal variates by the transformations  $J_1 \equiv \Phi_0^{-1}F$  and  $J_2 \equiv \Phi_0^{-1}G$ , respectively. The transformed random variables are then assumed to be bivariate normally distributed with correlation coefficient  $\rho$ . The corresponding joint density function is

$$h(\epsilon, u; \rho) = (1 - \rho^2)^{-1/2} f(\epsilon) g(u) \exp\{-\rho(2(1 - \rho^2))^{-1} [\rho(J_1^2(\epsilon) + J_2^2(u)) - 2J_1(\epsilon)J_2(u)]\} \quad (2.6)$$

where  $f$  and  $g$  are the marginal density functions of  $\epsilon$  and  $u$ , respectively. This approach has been proposed in Lee [1980] to construct censored regression models, when the marginal distributions are given, for the binary as well as polychotomous discrete choice cases. For the purpose of testing normality, our maintained hypothesis is that the marginal distribution of  $\epsilon$  is standard normal, the marginal distribution of  $u$  is the general Pearson distribution as in (2.5) and the joint density of  $\epsilon$  and  $u$  is the one in (2.6). This basic model is of interest since it does contain the bivariate normal distribution as a special case; the marginal distribution of  $u$  can have various shapes and the model is tractible. The maintained hypothesis that the choice

equation is a probit model deserves some comments. The distributional assumption in the choice model can apparently be tested based only on the observed dichotomous indicators  $I_i$ . A flexible distribution for testing the probit model can be chosen as the general Pearson distribution. Such test will be derived briefly in Appendix B. With only complications in notations, one can also derive, as a straightforward extension, a joint test of normality for both  $\varepsilon$  and  $u$  in our framework. In addition to the notational simplicity, there is no loss of generality to use the probit choice model as a maintained hypothesis. If an alternative probability choice model generated by a distribution function  $F(\varepsilon)$  rather than the probit model is found out to be the appropriate one to use, this choice model can be incorporated into our basic model as a maintained hypothesis with  $z_i\gamma$  and  $\varepsilon$  in the probit model replaced by the transformed values  $J_1(z_i\gamma)$  and  $J_1(\varepsilon)$ ; see Lee [1980].

Heteroscedasticity in the disturbances  $u_i$  can be incorporated into our basic model. Suppose that  $\sigma_i^2 = E(u_i^2)$  is a function of exogenous variables vector  $w_i$  of dimension  $m$ . Without loss of generality, we assume further that  $\sigma_i^2 = \alpha_1 + w_i\alpha_2$  is a linear function of  $w_i$ . When  $\alpha_2 = 0$ , the disturbances  $u_i$  are homoscedastic. There are several different ways that heteroscedasticity can be incorporated in our basic model. One of the approach that is relatively simple is to assume that  $c_1$  and  $c_2$  are constants for all  $i$  in the density function  $g(u)$  in (2.5) but  $c_{0i} = \alpha_1 + w_i\alpha_2$  is a function of  $w_i$ . This specification is justified since the variance of  $u_i$  under the general Pearson distribution is  $E(u_i^2) = c_{0i}/(1 - 3c_2)$ . This parameterization has been used in

Bera and Jarque [1980] in their testing of normality and heteroscedastic regression residuals in the standard linear model.

### 3. SPECIFICATION TESTS FOR THE CENSORED REGRESSION MODELS

There are several hypotheses that are of interest in this model. The basic model specified in the previous section is our universe model and will be denoted as Model  $M_U$ . The following list of hypotheses and the corresponding models will be considered in this and the subsequent sections.

$H_N$ : The marginal distributions of  $u_i$ 's are normal,  $N(0, \sigma_i^2)$ , with  $\sigma_i^2 = \alpha_1 + z_i \alpha_2$ .

$H_{NH}$ : The disturbances  $u_i$  are normally distributed  $N(0, \sigma^2)$  and are homoscedastic.

$H_{IN}$ : The disturbances  $u_i$  are normally distributed  $N(0, \sigma_i^2)$  and are independent with all  $\epsilon_j$ .

$H_{INH}$ : The disturbances  $u_i$  are homoscedastic, normally distributed  $N(0, \sigma^2)$ , and are independent with all  $\epsilon_j$ .

The corresponding models will be denoted as  $M_N$ ,  $M_{NH}$ ,  $M_{IN}$ , and  $M_{INH}$ . In the models  $M_{IN}$  and  $M_{INH}$ , since the choice equation is independent with the regression equation, there is no selection bias in the dependent variable of the regression equation and the values of  $y_1$  are only randomly missing. The model  $M_{INH}$  is the most restrictive one. The other two hypotheses  $H_N$  and  $H_{NH}$  specify normal distribution, and normal and

homoscedastic disturbances in the censored regression model. There are, of course, other combinations of hypotheses that are of interest but, because it is complicated to set up the tests, they will not be considered here. The following diagram indicates the obvious relationships of the models in a lattice structure. The lower level models in the tree are special cases for the upper ones.

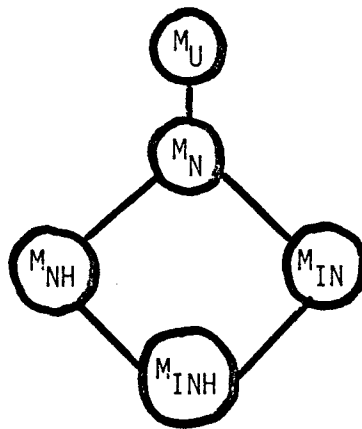


Diagram 1: Relationships among the models

To test the hypotheses, we will consider the Lagrangean multiplier test approach developed in Aitchison and Silvey [1958], and Silvey [1959]. The Lagrangian multiplier statistic is identical to the efficient score statistic in Rao [1973]. Its applications in econometrics have been considered in Breusch and Pagan [1979, 1980] and Engle [1978], among others. This approach is attractive for our models since the computation of the maximum likelihood estimates for the basic model  $M_U$  will be very complicated. The log likelihood function for the basic model is

$$\begin{aligned}
L_u = \sum_{i=1}^N \left\{ (1 - I_i) \ln(1 - \Phi(z_i \gamma)) + I_i q(y_i - x_i \beta) + I_i \ln \Phi \left[ (z_i \gamma \right. \right. \\
\left. \left. - \rho J_2(y_i - x_i \beta)) / \sqrt{1 - \rho^2} \right] \right. \\
\left. - I_i \ln \int_{-\infty}^{\infty} \exp q(u) du \right\} \quad (3.1)
\end{aligned}$$

where

$$q(u) \equiv \int \frac{c_1 - u}{\alpha_1 + w\alpha_2 - c_1 u + c_2 u^2} du \quad (3.2)$$

and

$$J_2(u) = \Phi^{-1} \left( \int_{-\infty}^u \exp q(t) dt / \int_{-\infty}^{\infty} \exp q(t) dt \right) \quad (3.3)$$

The vector of parameters in the log likelihood function  $L_u$  is  $\theta' = (\rho, \beta', \gamma', \alpha_1, \alpha_2', c_1, c_2)$ . To derive the Lagrangean multiplier test for a specific hypothesis in our list, we need to derive the first-order derivatives of  $L_u$  and then evaluate them at the restricted parameters under the specific hypothesis. Let

$$\left. \frac{\partial L_u}{\partial \theta_i} \right|_H$$

denote the first-order derivative of  $L_u$  with respect to  $\theta_i$ , evaluated under the hypothesis  $H$ . To simplify the notations, let  $\phi(\cdot)$  denote the

standard normal density function and  $\sigma_i^2 = \alpha_1 + w_i \alpha_2$ . Furthermore, define a  $9 \times 1$  vector  $v(i)' = (v_1(i), \dots, v_9(i))$  as

$$v_1(i) = I_i - \Phi(z_i \gamma) \quad (3.4)$$

$$v_2(i) = I_i u_i + \rho \sigma_i \phi(z_i \gamma) \quad (3.5)$$

$$v_3(i) = I_i u_i^2 - \sigma_i^2 (\phi(z_i \gamma) - \rho^2 z_i \gamma \phi(z_i \gamma)) \quad (3.6)$$

$$v_4(i) = I_i u_i^3 - \rho \sigma_i^3 (\rho^2 - 3 - \rho^2 (z_i \gamma)^2) \phi(z_i \gamma) \quad (3.7)$$

$$v_5(i) = I_i u_i^4 - \sigma_i^4 [3\phi(z_i \gamma) - \rho^2 (6 - 3\rho^2 + \rho^2 (z_i \gamma)^2) z_i \gamma \phi(z_i \gamma)] \quad (3.8)$$

$$v_6(i) = I_i \frac{\phi\left[\frac{z_i \gamma - \rho u_i / \sigma_i}{\sqrt{1 - \rho^2}}\right]}{\phi\left[\frac{z_i \gamma - \rho u_i / \sigma_i}{\sqrt{1 - \rho^2}}\right]} - \sqrt{1 - \rho^2} \phi(z_i \gamma) \quad (3.9)$$

$$v_7(i) = I_i u_i \frac{\phi\left[\frac{z_i \gamma - \rho u_i / \sigma_i}{\sqrt{1 - \rho^2}}\right]}{\phi\left[\frac{z_i \gamma - \rho u_i / \sigma_i}{\sqrt{1 - \rho^2}}\right]} - \rho (1 - \rho^2)^{1/2} \sigma_i z_i \gamma \phi(z_i \gamma) \quad (3.10)$$

$$v_8(i) = I_i u_i^2 \frac{\phi\left[\frac{z_i \gamma - \rho u_i / \sigma_i}{\sqrt{1 - \rho^2}}\right]}{\phi\left[\frac{z_i \gamma - \rho u_i / \sigma_i}{\sqrt{1 - \rho^2}}\right]} - (1 - \rho^2)^{1/2} (1 - \rho^2 + \rho^2 (z_i \gamma)^2) \sigma_i^2 \phi(z_i \gamma) \quad (3.11)$$



$$\begin{aligned}
v_9(i) &= I_i u_i^3 \frac{\phi \left[ \frac{(z_i \gamma - \rho u_i / \sigma_i) / \sqrt{1 - \rho^2}}{\sigma_i} \right]}{\Phi \left[ \frac{(z_i \gamma - \rho u_i / \sigma_i) / \sqrt{1 - \rho^2}}{\sigma_i} \right]} \\
&\quad - \rho (1 - \rho^2)^{1/2} [3(1 - \rho^2) + \rho^2 (z_i \gamma)^2] \sigma_i^3 z_i \gamma \phi(z_i \gamma)
\end{aligned} \tag{3.12}$$

It can be easily shown under  $H_N$  that all the components of  $v(i)$  are random variables with zero mean,<sup>2</sup> i.e.,  $E_{H_N}(v(i)) = 0$ . It is straightforward to show that

$$\left. \frac{\partial L_u}{\partial \rho} \right|_{H_N} = \sum_{i=1}^N \left[ \frac{\rho}{(1 - \rho^2)^{3/2}} (z_i \gamma) v_6(i) - \frac{1}{(1 - \rho^2)^{3/2} \sigma_i} v_7(i) \right] \tag{3.13}$$

$$\left. \frac{\partial L_u}{\partial \beta} \right|_{H_N} = \sum_{i=1}^N \left[ \frac{x_i'}{\sigma_i^2} v_2(i) + \frac{\rho x_i'}{(1 - \rho^2)^{1/2} \sigma_i} v_6(i) \right] \tag{3.14}$$

$$\left. \frac{\partial L_u}{\partial \gamma} \right|_{H_N} = \sum_{i=1}^N \left[ \frac{\phi(z_i \gamma) z_i'}{1 - \Phi(z_i \gamma)} v_1(i) + \frac{z_i'}{(1 - \rho^2)^{1/2}} v_6(i) \right] \tag{3.15}$$

$$\left. \frac{\partial L_u}{\partial \alpha_1} \right|_{H_N} = \sum_{i=1}^N \left[ -\frac{1}{2\sigma_i^2} v_1(i) + \frac{1}{2\sigma_i^4} v_3(i) + \frac{\rho}{2(1 - \rho^2)^{1/2} \sigma_i^3} v_7(i) \right] \tag{3.16}$$

$$\begin{aligned} \left. \frac{\partial L_u}{\partial \alpha_2} \right|_{H_N} &= \sum_{i=1}^N \left[ -\frac{w_i^!}{2\sigma_i^2} v_1(i) + \frac{w_i^!}{2\sigma_i^4} v_3(i) \right. \\ &\quad \left. + \frac{\rho w_i^!}{2(1-\rho^2)^{1/2} \sigma_i^3} v_7(i) \right] \end{aligned} \quad (3.17)$$

$$\begin{aligned} \left. \frac{\partial L_u}{\partial c_1} \right|_{H_N} &= \sum_{i=1}^N \left[ \frac{1}{\sigma_i^2} v_2(i) - \frac{1}{3\sigma_i^4} v_4(i) + \frac{\rho}{3(1-\rho^2)^{1/2} \sigma_i} v_6(i) \right. \\ &\quad \left. - \frac{\rho}{3(1-\rho^2)^{1/2} \sigma_i^3} v_8(i) \right] \end{aligned} \quad (3.18)$$

$$\begin{aligned} \left. \frac{\partial L_u}{\partial c_2} \right|_{H_N} &= \sum_{i=1}^N \left[ -\frac{3}{4} v_1(i) + \frac{1}{4\sigma_i^4} v_5(i) + \frac{3\rho}{4(1-\rho^2)^{1/2} \sigma_i} v_7(i) \right. \\ &\quad \left. - \frac{\rho}{4(1-\rho^2)^{1/2} \sigma_i^3} v_9(i) \right] \end{aligned} \quad (3.19)$$

The above expressions are derived by evaluating  $(\partial L_u)/(\partial \theta)$  at  $c_1 = c_2 = 0$ . These derivations can be written in more compact notations. Define a matrix  $X(i)$  consisting of weights of  $v(i)$  in the equations (3.13)-(3.19) and zero in the proper positions. We have

$$\left. \frac{\partial L_u}{\partial \theta} \right|_{H_N} = \sum_{i=1}^N X(i)v(i) \quad (3.20)$$

Under  $H_N$ , the information matrix is

$$E_{H_N} = \begin{pmatrix} \frac{\partial L_u}{\partial \theta} & \frac{\partial L_u}{\partial \theta'} \end{pmatrix} = \sum_{i=1}^N X(i)\Omega(i)X(i)' \quad (3.21)$$

where  $\Omega(i) = E_{H_N}(v(i)v(i)')$  is a  $9 \times 9$  covariance matrix of  $v(i)$ . The test of the hypothesis  $H_N$  is based on the gradient vector in (3.20) and the information matrix (3.21). Partition the matrix  $\Omega(i)$  into a block matrix as

$$\Omega(i) = \begin{bmatrix} \Omega_1(i) & \Omega_2(i) \\ \Omega_2(i)' & \Omega_3(i) \end{bmatrix}$$

where  $\Omega_1(i)$  is the  $5 \times 5$  covariance matrix of the first five components of  $v(i)$ ,  $\Omega_3(i)$  is the  $4 \times 4$  covariance matrix of the last four components of  $v(i)$  and  $\Omega_2(i)$  is the cross covariance of those components.  $\Omega_1(i)$  and  $\Omega_2(i)$  can be derived analytically by some recursive formulae for  $E_{H_N}(I_i u_i^\ell)$ ,  $\ell = 1, \dots, 8$  and

$$E_{H_N} \left[ I_i u_i^\ell \frac{\phi \left[ \frac{z_i \gamma - \rho u_i / \sigma_i}{\sqrt{1 - \rho^2}} \right]}{\Phi \left[ \frac{z_i \gamma - \rho u_i / \sigma_i}{\sqrt{1 - \rho^2}} \right]} \right],$$

$\ell = 1, \dots, 7$ . Since, under  $H_N$ ,  $u_i = \sigma_i \rho \epsilon_i + v_i$  where  $v_i$  is  $N(0, (1 - \rho^2)\sigma_i^2)$  and is independent with  $\epsilon_i$ ,  $E_{H_N}(I_i u_i^\ell)$  can be derived in terms of  $E_{H_N}(I_i \epsilon_i^\ell)$  and  $E_{H_N}(v_i)$  from the binomial expansion,

$$E_{H_N}(I_i u_i^\ell) = \sum_{r=0}^{\ell} C_r^\ell \left( \alpha_1 + w_i \alpha_2 \right)^{r/2} \rho^r E_{H_N}(I_i \epsilon_i^r) E_{H_N}(v_i^{\ell-r}), \quad \ell \geq 1 \quad (3.22)$$

where

$$C_r^{\ell} \equiv \frac{\ell!}{r!(\ell-r)!}$$

The expectation of  $I_i \varepsilon_i$  under  $H_N$  is

$$E_{H_N}(I_i \varepsilon_i) = -\phi(z_i \gamma) \quad (3.23)$$

see, Johnson and Kotz [1972], and the expectations of  $I_i \varepsilon_i^r$ ,  $r \geq 2$ , can be derived from the following recursive formulae,

$$\begin{aligned} E_{H_N}(I_i \varepsilon_i^r) &= \int_{-\infty}^{z_i \gamma} \varepsilon^r \phi(\varepsilon) d\varepsilon \\ &= -\varepsilon^{r-1} \phi(\varepsilon) \Big|_{-\infty}^{z_i \gamma} + (r-1) \int_{-\infty}^{z_i \gamma} \varepsilon^{r-2} \phi(\varepsilon) d\varepsilon \\ &= -\left(z_i \gamma\right)^{r-1} \phi(z_i \gamma) + (r-1) E_{H_N}(I_i \varepsilon_i^{r-2}), \end{aligned}$$

$$r \geq 2 \quad (3.24)$$

The moments of the normal variable  $v_i$  are well known;

$$\begin{aligned} E_{H_N}(v_i^r) &= 1 \cdot 3 \cdot 5 \dots (r-1) \sigma_i^r (1 - \rho^2)^{r/2} && \text{for even } r, \\ &= 0 && \text{for odd } r \end{aligned}$$

$$(3.25)$$

Under  $H_N$ , the joint density of  $(u_i, I_i = 1)$  is

$$\psi(u_i, 1) = \frac{1}{\sigma_i} \phi\left(\frac{u_i}{\sigma_i}\right) \Phi\left[\frac{(z_i\gamma - \rho u_i/\sigma_i)/\sqrt{1-\rho^2}}{\sqrt{1-\rho^2}}\right]$$

and therefore

$$\begin{aligned} E_{H_N} \left( I_i u_i^\ell \frac{\phi\left[\frac{(z_i\gamma - \rho u_i/\sigma_i)/\sqrt{1-\rho^2}}{\sqrt{1-\rho^2}}\right]}{\Phi\left[\frac{(z_i\gamma - \rho u_i/\sigma_i)/\sqrt{1-\rho^2}}{\sqrt{1-\rho^2}}\right]} \right) \\ = \int_{-\infty}^{\infty} \frac{u^\ell}{\sigma_i} \phi\left[\frac{(z_i\gamma - \rho u/\sigma_i)/\sqrt{1-\rho^2}}{\sqrt{1-\rho^2}}\right] \phi(u/\sigma_i) du \\ = \sqrt{1-\rho^2} \phi(z_i\gamma) \int_{-\infty}^{\infty} u^\ell \frac{1}{\sigma_i \sqrt{1-\rho^2}} \phi\left(\frac{u - \sigma_i \rho z_i \gamma}{\sigma_i \sqrt{1-\rho^2}}\right) du \\ = \sqrt{1-\rho^2} \phi(z_i\gamma) E_S(u^\ell) \end{aligned} \quad (3.26)$$

where  $E_S(u^\ell)$  denotes the  $\ell^{\text{th}}$  moment of  $u$  around zero with respect to the normal density function with mean  $\sigma_i \rho z_i \gamma$  and variance  $\sigma_i^2(1-\rho^2)$ .  $E_f(u^\ell)$  can also be evaluated by recursive formula,

$$E_S(u^\ell) = \begin{cases} 1 \cdot 3 \cdot 5 \dots (\ell-1) \sigma_i^\ell (1-\rho^2)^{\ell/2} \sum_{r=1}^{\ell} (-1)^r C_r^{\ell} \sigma_i^r \rho^r (z_i\gamma)^r E_S(u^{\ell-r}), & \text{for even } \ell \\ -\sum_{r=1}^{\ell} (-1)^r C_r^{\ell} \sigma_i^r \rho^r (z_i\gamma)^r E_S(u^{\ell-r}), & \text{for odd } \ell \end{cases} \quad (3.27)$$

for  $\ell = 1, 2, \dots$ . The analytical expressions for the submatrix  $\Omega_3(i)$ , however, are difficult to be derived but can be estimated by the sample moments  $S(i)$

$$S(i) = \begin{bmatrix} v_6^2(i) & v_6(i)v_7(i) & v_6(i)v_8(i) \\ * & v_7^2(i) & v_7(i)v_8(i) \\ * & * & v_8^2(i) \end{bmatrix}$$

Given that the samples are random, it is clear from equations (3.20) and (3.21) that the asymptotic properties of the test with  $\Omega_3(i)$  being estimated by  $S(i)$  will be the same.

Let  $I_\theta$  be the identity matrix with the same dimension  $k_1 + k_2 + m + 4$  as the parameter vector  $\theta' = (\rho, \beta', \gamma', \alpha_1, \alpha_2', c_1, c_2)$ . Let  $J_N$  be the submatrix of  $I_\theta$  consisting of the last two columns of  $I_\theta$ . The Lagrangean multiplier test for  $H_N$  is

$$\left( \sum_{i=1}^N v(i)' X'(i) \right) J_N J_N' \left( \sum_{i=1}^N X(i) \Omega(i) X(i)' \right)^{-1} J_N J_N' \left( \sum_{i=1}^N X(i) v(i) \right) \quad (3.28)$$

evaluated at the constrained maximum likelihood estimates of  $(\rho, \beta', \gamma', \alpha_1, \alpha_2')$  in the model  $M_N$ . The statistic (3.28) is asymptotically Chi-square distributed with 2 degrees of freedom. Asymptotic Chi-square test statistic can also be derived for any consistent estimates instead of maximum likelihood estimates as described in Breusch and Pagan [1980]

which has been called the pseudo Lagrangean multiplier test. The pseudo Lagrangean multiplier test is of interest for our model since consistent two-stage estimates are available in Amemiya [1973], Heckman [1976] and Lee [1979], among others, for the homoscedastic disturbance case. The formulation for the pseudo Lagrangean multiplier is also based on the gradient vector (3.20) and the details are referred to the article of Breusch and Pagan [1980].

Similar Lagrangean multiplier tests can also be derived for the more restrictive hypotheses  $H_{NH}$ ,  $H_{IH}$  and  $H_{INH}$ . Let  $J_{NH}$  be the submatrix of  $I_{\theta}$  consisting of the last  $m + 2$  columns corresponding to the sub-vector  $(\alpha_2', c_1, c_2)$  of  $\theta$ . The Lagrangean multiplier test for  $H_{NH}$  is

$$\left( \sum_{i=1}^N v(i)' X(i)' \right) J_{NH} J_{NH}' \left( \sum_{i=1}^N X(i) \Omega(i) X(i)' \right)^{-1} J_{NH} J_{NH}' \left( \sum_{i=1}^N X(i) v(i) \right) \quad (3.29)$$

evaluated at the constrained MLE of  $(\rho, \beta', \gamma', \alpha_1)$  in the model  $M_{NH}$ . This statistic is asymptotically chi-square distributed with  $m + 2$  degrees of freedom.

#### 4. TESTS FOR SELECTION BIAS, HOMOSCEDASTICITY AND NORMALITY

Selection bias is present in the regression equation (2.1) if and only if the disturbances  $u_i$  and  $\varepsilon_i$  are correlated. The testing of no selection bias in our model is to test that the correlation coefficient  $\rho$  equals zero. It is computationally involved to set up a marginal Lagrangean multiplier test for testing  $\rho$  in our basic model since the gradient vector of the corresponding likelihood function involves

complicated indefinite integrals. It is of interest to consider the joint test of selection bias, normality and homoscedasticity.

Under the hypothesis  $H_{INH}$ , the first order derivatives in (3.13) - (3.19) can be greatly simplified. To simplify notations, let  $\phi_i$  and  $\Phi_i$  denote the normal density and probability functions evaluated at  $z_i\gamma$ .

We have

$$\left. \frac{\partial L_u}{\partial \rho} \right|_{H_{INH}} = \sum_{i=1}^N \left( - \frac{\phi_i}{\Phi_i} \right) \frac{1}{\sqrt{\alpha_1}} I_i u_i \quad (4.1)$$

$$\left. \frac{\partial L_u}{\partial \beta} \right|_{H_{INH}} = \sum_{i=1}^N \frac{x_i'}{\alpha_1} I_i u_i \quad (4.2)$$

$$\left. \frac{\partial L_u}{\partial \gamma} \right|_{H_{INH}} = \sum_{i=1}^N \frac{\phi_i}{\Phi_i (1 - \Phi_i)} z_i' (I_i - \Phi_i) \quad (4.3)$$

$$\left. \frac{\partial L_u}{\partial \alpha_1} \right|_{H_{INH}} = \sum_{i=1}^N \frac{1}{2\alpha_1^2} I_i (u_i^2 - \alpha_1) \quad (4.4)$$

$$\left. \frac{\partial L_u}{\partial \alpha_2} \right|_{H_{INH}} = \sum_{i=1}^N \frac{w_i'}{2\alpha_1^2} I_i (u_i^2 - \alpha_1) \quad (4.5)$$

$$\left. \frac{\partial L_u}{\partial c_1} \right|_{H_{INH}} = \sum_{i=1}^N \left[ \frac{1}{\alpha_1} I_i u_i - \frac{1}{3\alpha_1^2} I_i u_i^3 \right] \quad (4.6)$$

$$\left. \frac{\partial L_u}{\partial c_2} \right|_{H_{INH}} = \sum_{i=1}^N \left[ - \frac{3}{4} (I_i - \Phi_i) + \frac{1}{4\alpha_1^2} (I_i u_i^4 - 3\alpha_1^2 \Phi_i) \right] \quad (4.7)$$



It is clear from the above expressions that the Lagrangean multiplier test for  $H_{INH}$  will utilize the information in the first four moments of  $I_i u_i$  and  $I_i - \phi(z_i \gamma)$ . Under  $H_{INH}$ , the information matrix of  $\frac{\partial L_u}{\partial \theta}$  is a diagonal block matrix;

$$E_{H_{INH}} \begin{pmatrix} \frac{\partial L_u}{\partial \theta} & \frac{\partial L_u}{\partial \theta'} \end{pmatrix} = \begin{bmatrix} D_1 & & 0 \\ & D_2 & \\ 0 & & D_3 \end{bmatrix} \quad (4.8)$$

where

$$D_1 = \sum_{i=1}^N \begin{bmatrix} \frac{\phi_i^2}{\phi_i} & -\frac{1}{\sqrt{\alpha_1}} \phi_i x_i \\ -\frac{1}{\sqrt{\alpha_1}} \phi_i x_i' & \frac{1}{\alpha_1} \phi_i x_i' x_i \end{bmatrix} \quad (4.9)$$

$$D_2 = \sum_{i=1}^N \frac{\phi_i^2}{\phi_i (1 - \phi_i)} z_i' z_i \quad (4.10)$$

$$D_3 = \sum_{i=1}^N \phi_i \begin{bmatrix} \frac{1}{2\alpha_1^2} & \frac{1}{2\alpha_1^2} w_i & 0 & \frac{3}{2\alpha_1} \\ \frac{1}{2\alpha_1^2} w_i' & \frac{1}{2\alpha_1^2} w_i' w_i & 0 & \frac{3}{2\alpha_1} w_i' \\ 0 & 0 & \frac{2}{3\alpha_1} & 0 \\ \frac{3}{2\alpha_1} & \frac{3}{2\alpha_1} w_i & 0 & 6 \end{bmatrix} \quad (4.11)$$

It is straightforward to derive the inverse matrices of  $D_1$ ,  $D_2$  and  $D_3$ . To write the Lagrangean multiplier statistics in the matrix notations, denote  $D_\phi = \text{Diag} [\phi_i]$  and  $D_{\phi'} = \text{Diag} [\phi_i']$  as two  $N \times N$  diagonal matrices with elements  $\phi_i$  and  $\phi_i'$ , respectively. Let  $\mathbf{1}'_N = (1, \dots, 1)$  be a  $N \times 1$  vector consisting of unity,  $W' = [w_1', w_2', \dots, w_N']$  and  $X' = [x_1', x_2', \dots, x_N']$  be the two matrices of exogenous variables. Furthermore, let  $\zeta_0' = (I_1 - \phi_1, \dots, I_N - \phi_N)$ ,  $\zeta_1' = (I_1 u_1, \dots, I_N u_N)$ ,  $\zeta_2' = (I_1 u_1^2 - \alpha_1 \phi_1, \dots, I_N u_N^2 - \alpha_1 \phi_N)$ ,  $\zeta_3' = (I_1 u_1^3, \dots, I_N u_N^3)$  and  $\zeta_4' = (I_1 u_1^4 - 3\alpha_1^2 \phi_1, \dots, I_N u_N^4 - 3\alpha_1^2 \phi_N)$ . The Lagrangean multiplier statistic  $S_{INH}$  for testing the hypothesis  $H_{INH}$  is the sum of three components,

$$S_{INH} = S_{I|NH} + S_{H|IN} + S_{N|IH} \quad (4.12)$$

where

$$S_{I|NH} = \frac{1}{\alpha_1} \zeta_1' D_\phi D_\phi^{-1} \ell_N \left[ \ell_N' D_\phi^2 D_\phi^{-1} \ell_N - \ell_N' D_\phi X (X' D_\phi X)^{-1} X' D_\phi \ell_N \right]^{-1} \ell_N' D_\phi D_\phi^{-1} \zeta_1 \quad (4.13)$$

$$S_{H|IN} = \frac{1}{2} \left( \frac{1}{\alpha_1} \zeta_2 - \zeta_0 \right)' W \left[ W' D_\phi W - W' D_\phi \ell_N (\ell_N' D_\phi \ell_N)^{-1} \ell_N' D_\phi W \right]^{-1} W' \left( \frac{1}{\alpha_1} \zeta_2 - \zeta_0 \right) \quad (4.14)$$

and

$$S_{N|IH} = \frac{3}{2\alpha_1} \left( \zeta_1 - \frac{1}{3\alpha_1} \zeta_3 \right)' \ell_N \left( \ell_N' D_\phi \ell_N \right)^{-1} \ell_N' \left( \zeta_1 - \frac{1}{3\alpha_1} \zeta_3 \right) + \frac{1}{24} \left( \frac{1}{\alpha_1^2} \zeta_4 - 3\zeta_0 \right)' \ell_N \left( \ell_N' D_\phi \ell_N \right)^{-1} \ell_N' \left( \frac{1}{\alpha_1^2} \zeta_4 - 3\zeta_0 \right) \quad (4.15)$$

which are evaluated at the MLE of  $\beta$ ,  $\gamma$  and  $\alpha_1$  in the model  $M_{INH}$ . The restricted MLE of  $\beta$  and  $\sigma^2$  in the model  $M_{INH}$  are

$$\hat{\beta} = \left( \sum_{i=1}^N I_i x_i' x_i \right)^{-1} \sum_{i=1}^N I_i x_i' y_i$$

and

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^N I_i (y_i - x_i \hat{\beta})^2}{\sum_{i=1}^N I_i}$$

The MLE of  $\gamma$  in  $M_{INH}$  is the probit MLE.

Each component in the Lagrangean multiplier test statistic  $S_{INH}$  is a statistic which can provide some interesting interpretations. As shown in the appendix, each component statistic is a conditional Lagrangean multiplier test statistic for one hypothesis conditional on the others. The component  $S_{I|NH}$  is a Lagrangean multiplier test statistic for testing the independence of the choice equation (2.2) and the regression equation (2.1), in the model  $M_{NH}$ . Thus it is a test statistic for the non-existence of selection bias conditional on the assumptions that the disturbances in the regression equation (2.1) is homoscedastic and nominally distributed. On the other hand, the statistic  $S_{H|IN}$  is a conditional Lagrangean multiplier test statistic for testing homoscedasticity conditional on the disturbances  $u_i$  being normally distributed and no selection bias. The statistic  $S_{N|IH}$  is a conditional Lagrangean multiplier test statistic for testing normality conditional on homoscedasticity and no selection bias. These three conditional statistics are apparently orthogonal to each other in our basic model under the joint hypothesis  $H_{INH}$ . All the statistics are asymptotically chi-square distributed. The degrees of freedom of  $S_{I|NH}$ ,  $S_{N|IH}$  and  $S_{H|IN}$  are, respectively, one, two and  $m$  where  $m$  is the dimension of the vector  $w_i$ . Consequently, the test statistic  $S_{INH}$  is asymptotically chi-square distributed with  $m + 3$  degrees of freedom. As the joint test statistic is a sum of three conditional test statistics, we can access the proportional contribution

of each conditional test statistic to the overall statistic in the joint test of selection bias, homoscedasticity and normality hypotheses.

There are other interesting relationships of these statistics with some significance tests in some regression equations. Consider the following equation derived from (2.1) under the model  $M_{NH}$ .

$$I_i y_i = I_i x_i \beta + I_i u_i, \quad i = 1, \dots, N \quad (4.16)$$

As contrary to equation (2.1), all the values of the dependent variables in (4.16) are observable.<sup>3</sup> Since  $I_i$  and  $u_i$  are correlated, we have  $E(I_i u_i) = -\rho \alpha_1^{1/2} \phi_i$ . It implies that

$$I_i y_i = I_i x_i \beta - \sigma_{u\varepsilon} \phi(z_i \gamma) + \xi_i \quad (4.17)$$

where  $\sigma_{u\varepsilon} = \rho \alpha_1^{1/2}$  is the covariance of  $\varepsilon_i$  and  $u_i$  and  $\xi_i = I_i u_i + \sigma_{u\varepsilon} \phi(z_i \gamma)$ . The residuals  $\xi_i$  have zero mean by construction. Let  $\hat{\gamma}$  be the probit MLE of  $\gamma$ . Substituting  $\hat{\gamma}$  into the equation (4.17), it becomes

$$I_i y_i = I_i x_i \beta - \sigma_{u\varepsilon} \phi(z_i \hat{\gamma}) + \tilde{\xi}_i \quad (4.18)$$

where  $\tilde{\xi}_i = \xi_i - \sigma_{u\varepsilon} (\phi(z_i \gamma) - \phi(z_i \hat{\gamma}))$  is the modified residual in the equation. Since  $\sigma_{u\varepsilon} = \rho \alpha_1^{1/2}$  where  $\alpha_1$  is the variance of  $u_i$ ,  $\sigma_{u\varepsilon}$  is zero if and only if  $\rho = 0$ . An approach to test the absence of selection bias is to test  $\sigma_{u\varepsilon} = 0$  in equation (4.18). It should be noted that ordinary least squares approach does not provide consistent estimates of  $\beta$  and  $\sigma_{u\varepsilon}$  in the estimation of the regression equation (4.18) since the regressor  $I_i x_i$  and the disturbance  $\xi_i$  are correlated.<sup>4</sup> A consistent approach is to use an instrumental variable (I.V.) approach to estimate this equation.

A useful instrumental variable for  $I_i x_i$  is  $\phi(z_i \hat{\gamma}) x_i$ . Under the hypothesis,  $H_{I|NH}$ :  $\rho = 0$ , conditional on normality and homoscedasticity,  $\tilde{\xi}_i$  equals  $\xi_i$  and is heteroscedastic disturbance with variance  $\alpha_1 \phi(z_i \gamma)$ . An appropriate estimation approach is thus a weighted IV approach. To simplify notations, let  $\hat{\phi}_i = \phi(z_i \hat{\gamma})$  and  $\hat{\Phi}_i = \Phi(z_i \hat{\gamma})$ . We have

$$\begin{aligned} \begin{pmatrix} \tilde{\beta} \\ \tilde{\sigma}_{u\epsilon} \end{pmatrix}_{IV} &= \left[ \sum_{i=1}^N \frac{1}{\hat{\phi}_i} \begin{pmatrix} \hat{\phi}_i x_i' \\ \hat{\phi}_i \end{pmatrix} (I_i x_i \quad \hat{\phi}_i) \right]^{-1} \sum_{i=1}^N \frac{1}{\hat{\phi}_i} \begin{pmatrix} \hat{\phi}_i x_i' \\ \hat{\phi}_i \end{pmatrix} I_i y_i \\ &= \begin{bmatrix} \sum_{i=1}^N I_i x_i' x_i & \sum_{i=1}^N \hat{\phi}_i x_i' \\ \sum_{i=1}^N \frac{\hat{\phi}_i}{\hat{\phi}_i} I_i x_i & \sum_{i=1}^N \frac{\hat{\phi}_i^2}{\hat{\phi}_i} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^N I_i x_i' y_i \\ \sum_{i=1}^N \frac{\hat{\phi}_i}{\hat{\phi}_i} I_i y_i \end{bmatrix} \end{aligned} \quad (4.19)$$

It can be easily shown under some regular conditions as in Lee et al. [1980] that the IV estimator is asymptotically normal with the asymptotic variance,

$$\text{var} \begin{pmatrix} \tilde{\beta} \\ \tilde{\sigma}_{u\epsilon} \end{pmatrix}_{IV} = \alpha_1 \begin{bmatrix} \sum_{i=1}^N \phi_i x_i' x_i & \sum_{i=1}^N \phi_i x_i' \\ \sum_{i=1}^N \phi_i x_i & \sum_{i=1}^N \frac{\phi_i^2}{\phi_i} \end{bmatrix}^{-1} \quad (4.20)$$

under the hypothesis  $H_{I|NH}$ . A test of this hypothesis is to test the significance of  $\sigma_{u\epsilon}$  different from zero with the IV estimator  $\tilde{\sigma}_{u\epsilon}$  and

its associated asymptotic distribution. From equation (4.19), we have

$$\begin{aligned} \tilde{\sigma}_{u\epsilon} = & \left[ \sum_{i=1}^N \frac{\hat{\phi}_i^2}{\hat{\phi}_i} - \left( \sum_{i=1}^N \frac{\hat{\phi}_i}{\hat{\phi}_i} I_i x_i \right) \left( \sum_{i=1}^N I_i x_i' x_i \right)^{-1} \left( \sum_{i=1}^N \hat{\phi}_i x_i' \right) \right]^{-1} \\ & \cdot \sum_{i=1}^N \frac{\hat{\phi}_i}{\hat{\phi}_i} (I_i y_i - I_i x_i \hat{\beta}) \end{aligned} \quad (4.21)$$

where

$$\hat{\beta} = \left( \sum_{i=1}^N I_i x_i' x_i \right)^{-1} \left( \sum_{i=1}^N I_i x_i y_i \right)$$

and the statistic

$$\begin{aligned} \frac{1}{\hat{\alpha}_1} \left[ \sum_{i=1}^N \frac{\hat{\phi}_i}{\hat{\phi}_i} (I_i y_i - I_i x_i \hat{\beta}) \right] \left[ \sum_{i=1}^N \frac{\hat{\phi}_i^2}{\hat{\phi}_i} - \sum_{i=1}^N \hat{\phi}_i x_i \left( \sum_{i=1}^N \hat{\phi}_i x_i' x_i \right)^{-1} \right. \\ \left. \cdot \sum_{i=1}^N \hat{\phi}_i x_i' \right]^{-1} \sum_{i=1}^N \frac{\hat{\phi}_i}{\hat{\phi}_i} (I_i y_i - I_i x_i \hat{\beta}) \end{aligned} \quad (4.22)$$

where  $\hat{\alpha}_1 = \sum_{i=1}^N I_i (y_i - x_i \hat{\beta})^2 / \sum_{i=1}^N I_i$ , is asymptotically chi-square distributed with one degree of freedom. It is obvious that this statistic is exactly the same as the conditional statistic  $S_{I|NH}$  evaluated at the MLE of  $(\beta, \gamma, \alpha_1)$  in the model  $M_{INH}$ .

The above analysis proved that the conditional Lagrangean multiplier test for selection bias is exactly the same as the test of

significance of  $\sigma_{u\epsilon}$  in equation (4.18). Equation (4.18) is different from the following equation for which the two-stage estimation in Heckman [1976] and Lee [1979] was developed,

$$y_i = x_i\beta - \sigma_{u\epsilon} \frac{\phi_i}{\Phi_i} + \zeta_i \quad (4.23)$$

where  $E(\zeta_i | I_i = 1) = 0$ . The test of significance of  $\sigma_{u\epsilon}$  from the modified equation,

$$y_i = x_i\beta - \sigma_{u\epsilon} \frac{\hat{\phi}_i}{\hat{\Phi}_i} + \tilde{\zeta}_i \quad (4.24)$$

is also equivalent to the above conditional Lagrangean multiplier test for selection bias.<sup>5/</sup> This can be easily shown as follows. Without loss of generality, let us assume that the first  $N_1 (< N)$  observations  $y_i$  are the non-censored observations of  $y_{1i}$ . The familiar two-stage estimates of  $\beta$  and  $\sigma_{u\epsilon}$  are

$$\begin{pmatrix} \hat{\beta} \\ \hat{\sigma}_{u\epsilon} \end{pmatrix}_T = \left[ \sum_{i=1}^{N_1} \begin{pmatrix} x_i' \\ -\frac{\hat{\phi}_i}{\hat{\Phi}_i} \end{pmatrix} \begin{pmatrix} x_i \\ -\frac{\hat{\phi}_i}{\hat{\Phi}_i} \end{pmatrix} \right]^{-1} \sum_{i=1}^{N_1} \begin{pmatrix} x_i' y_i \\ -\frac{\hat{\phi}_i}{\hat{\Phi}_i} y_i \end{pmatrix} \quad (4.25)$$

and it follows



$$\begin{aligned}
(\hat{\sigma}_{u\epsilon})_T = & - \left[ \sum_{i=1}^{N_1} \frac{\hat{\phi}_i^2}{\hat{\phi}_i^2} - \left( \sum_{i=1}^{N_1} \frac{\hat{\phi}_i^2}{\hat{\phi}_i^2} x_i \right) \left( \sum_{i=1}^{N_1} x_i' x_i \right)^{-1} \right. \\
& \left. \cdot \left[ \sum_{i=1}^{N_1} \frac{\hat{\phi}_i}{\hat{\phi}_i} x_i' \right] \right]^{-1} \sum_{i=1}^{N_1} \frac{\hat{\phi}_i}{\hat{\phi}_i} (y_i - x_i \hat{\beta})
\end{aligned} \tag{4.26}$$

where

$$\hat{\beta} = \left( \sum_{i=1}^{N_1} x_i' x_i \right)^{-1} \sum_{i=1}^{N_1} x_i' y_i$$

is the OLS estimate of  $\beta$  with the non-censored observations  $y_i$ . This estimate  $\hat{\beta}$  is also the MLE in the model  $M_{I|NH}$ , since

$$\left( \sum_{i=1}^{N_1} x_i' x_i \right)^{-1} \sum_{i=1}^{N_1} x_i' y_i = \left( \sum_{i=1}^N I_i x_i' x_i \right)^{-1} \sum_{i=1}^N I_i x_i' y_i$$

Under the hypothesis  $H_{I|NH}$ , the disturbances  $\tilde{\zeta}_i$  are homoscedastic and the asymptotic variance of  $(\hat{\sigma}_{u\epsilon})_T$  is

$$\alpha_1 \left[ \sum_{i=1}^{N_1} \frac{\hat{\phi}_i^2}{\hat{\phi}_i^2} - \sum_{i=1}^{N_1} \frac{\hat{\phi}_i}{\hat{\phi}_i} x_i \left( \sum_{i=1}^{N_1} x_i' x_i \right)^{-1} \sum_{i=1}^{N_1} \frac{\hat{\phi}_i}{\hat{\phi}_i} x_i' \right]^{-1} \tag{4.27}$$

which is asymptotically equivalent to

$$\alpha_1 \left[ \sum_{i=1}^N \frac{\hat{\phi}_i^2}{\hat{\phi}_i} - \sum_{i=1}^N \hat{\phi}_i x_i \left[ \sum_{i=1}^N \hat{\phi}_i x_i' x_i \right]^{-1} \sum_{i=1}^N \hat{\phi}_i x_i' \right]^{-1}$$

for the censored sampling since the sample  $i$  in (4.24) has probability  $\phi_i$  being included in the analysis and  $N_1$  is random.

It follows that a chi-square statistic for the test of significance of

$$\sigma_{uc} \frac{1}{\hat{\alpha}_1} \left[ \sum_{i=1}^{N_1} \frac{\hat{\phi}_i}{\hat{\phi}_i} (y_i - x_i \hat{\beta}) \right] \left[ \sum_{i=1}^{N_1} \frac{\hat{\phi}_i^2}{\hat{\phi}_i^2} - \sum_{i=1}^{N_1} \frac{\hat{\phi}_i}{\hat{\phi}_i} x_i \left( \sum_{i=1}^{N_1} x_i x_i' \right)^{-1} \sum_{i=1}^{N_1} \frac{\hat{\phi}_i}{\hat{\phi}_i} x_i \right]^{-1} \cdot \left[ \sum_{i=1}^{N_1} \frac{\hat{\phi}_i}{\hat{\phi}_i} (y_i - x_i \hat{\beta}) \right] \quad (4.28)$$

is asymptotically equivalent to the Lagrangean multiplier statistic  $S_{I|NH}$  in (4.22).

The conditional Lagrangean multiplier statistic  $S_{H|IN}$  can also be interpreted as a test of significance for  $\alpha_2 = 0$  in the following regression equation

$$I_i \hat{u}_i^2 = \alpha_1 I_i + I_i w_i \alpha_2 + \zeta_i \quad i = 1, \dots, N \quad (4.29)$$

where  $\hat{u}_i = y_i - x_i \hat{\beta}$  is the MLE residual in the model  $M_{INH}$  and  $\text{var}(\zeta_i) = 2\alpha_1^2 \hat{\phi}_i$ . The equation can be estimated by the IV approach. The appropriate instrumental variables are  $\hat{\phi}_i$  and  $\hat{\phi}_i w_i$  for  $I_i$  and  $I_i w_i$ , respectively. Since the disturbances  $\zeta_i$  are heteroscedastic, we have the following weighted IV estimator:

$$\begin{aligned} \begin{pmatrix} \tilde{\alpha}_1 \\ \tilde{\alpha}_2 \end{pmatrix}_{IV} &= \left\{ \sum_{i=1}^N \frac{1}{\hat{\phi}_i} \begin{bmatrix} \hat{\phi}_i \\ \hat{\phi}_i w_i' \end{bmatrix} \begin{bmatrix} I_i & I_i w_i \end{bmatrix} \right\}^{-1} \sum_{i=1}^N \frac{1}{\hat{\phi}_i} \begin{bmatrix} \hat{\phi}_i \\ \hat{\phi}_i w_i' \end{bmatrix} I_i \hat{u}_i^2 \\ &= \left\{ \sum_{i=1}^N \begin{bmatrix} I_i & I_i w_i \\ I_i w_i' & I_i w_i' w_i \end{bmatrix} \right\}^{-1} \sum_{i=1}^N \begin{bmatrix} I_i \hat{u}_i^2 \\ I_i \hat{u}_i^2 w_i' \end{bmatrix} \end{aligned} \quad (4.30)$$

which, in turn, is exactly the OLS estimator of  $\alpha_1$  and  $\alpha_2$  from equation (4.29). The asymptotic covariance matrix of this estimator is

$$\text{var} \begin{pmatrix} \tilde{\alpha}_1 \\ \tilde{\alpha}_2 \end{pmatrix}_{IV} = 2\alpha_1^2 \begin{bmatrix} \sum_{i=1}^N \phi_i & \sum_{i=1}^N \phi_i w_i \\ \sum_{i=1}^N \phi_i w_i' & \sum_{i=1}^N \phi_i w_i' w_i \end{bmatrix}^{-1}$$

Since

$$\begin{aligned} (\tilde{\alpha}_2)_{IV} &= \left[ \sum_{i=1}^N I_i w_i' w_i - \sum_{i=1}^N I_i w_i' \left( \sum_{i=1}^N I_i \right)^{-1} \sum_{i=1}^N I_i w_i \right]^{-1} \\ &\quad \cdot \sum_{i=1}^N \left( I_i \hat{u}_i^2 - I_i \hat{\alpha}_1 \right) w_i' \end{aligned}$$

where

$$\hat{\alpha}_1 = \left( \sum_{i=1}^N I_i \right)^{-1} \sum_{i=1}^N I_i \hat{u}_i^2$$

the following asymptotic chi-square statistic is a statistic for the test of significance for  $\alpha_2 = 0$

$$\begin{aligned} \frac{1}{2\hat{\alpha}_1^2} \sum_{i=1}^N (I_i \hat{u}_i^2 - I_i \hat{\alpha}_1) w_i \left[ \sum_{i=1}^N \hat{\phi}_i w_i' w_i - \sum_{i=1}^N \hat{\phi}_i w_i' \left( \sum_{i=1}^N \hat{\phi}_i \right)^{-1} \right. \\ \left. \cdot \sum_{i=1}^N \hat{\phi}_i w_i \right]^{-1} \sum_{i=1}^N (I_i \hat{u}_i^2 - I_i \hat{\alpha}_1) w_i \quad (4.31) \end{aligned}$$

Obviously, this statistic is exactly the conditional Lagrangean multiplier statistic  $S_{H|IN}$  for the testing of heteroscedasticity.

So far in our analysis of the test of selection bias, we have implicitly assumed that the Lagrangean multiplier  $\lambda_\rho$  corresponding to the constraint  $\rho = 0$ , evaluated at the constrained MLE under the hypothesis  $H_{INH}$ , is not identically zero. If it were zero, the Lagrangean multiplier test statistic would have no power for the testing of this hypothesis. This case does happen for the component  $S_{I|NH}$  when  $\sigma_{u\epsilon}$  is not estimable in equation (4.18). One of the cases is that  $z_i$  contains only the constant term and the exogenous variable  $x_i$  contains a constant term.<sup>6</sup> For this example, by the orthogonality of the least squares residual vector with the vectors of regressors,  $\sum_{i=1}^N I_i \hat{u}_i = 0$  where  $\hat{u}_i = I_i y_i - I_i x_i \hat{\beta}$ , and it is obvious that the first-order derivative

$$\left. \frac{\partial L_u}{\partial \rho} \right|_{H_{INH}}$$

evaluated at the MLE  $\hat{\beta}$  and  $\hat{\alpha}_1$  of the model  $M_{INH}$ , and also  $(\partial L_{NH})/(\partial \rho) \Big|_{H_{I|NH}}$  in Appendix A, are always zero for any sample size. Hence the Lagrangean multiplier tests for testing the hypothesis  $H_{INH}$  or  $H_{I|NH}$  do not have power and should not be used. This case provides probably the first example to confirm the conjecture of Silvey [1959, p. 399]. In the model  $M_{NH}$ , the parameter  $\rho$  is identifiable since the conditional odd moments of the uncensored disturbances conditional on  $I = 1$  will not be zero if  $\rho \neq 0$ . The likelihood ratio test for  $\rho = 0$  in the model  $M_{NH}$  is

applicable. Hence, for this case, the Lagrangean multiplier test is not equivalent, for the large samples to the likelihood ratio test. This difference occurs since the inference of the likelihood ratio test is based upon the levels of the likelihood and the Lagrangean multiplier test is based upon the first derivatives of the likelihood. At the point  $\hat{\beta}$ ,  $\hat{\alpha}_1$  and  $\rho = 0$ , the likelihood value is a stationary value in the loglikelihood function  $L_{NH}$  of the model  $M_{NH}$ .

In the test of the hypothesis  $H_{INH}$ , the Lagrangean multiplier statistic  $S_{INH}$  is decomposed into these orthogonal conditional Lagrangean multiplier statistics. For the test of the less restrictive hypothesis  $H_{IN}$  in  $M_U$  which allows heteroscedastic disturbances, the Lagrangean multiplier test statistic  $S_{IN}$  will also be decomposed into two components  $S_{I|N}$  and  $S_{N|I}$  where  $S_{I|N}$  is the Lagrangean multiplier test statistic for testing the hypothesis  $\rho = 0$  conditional on normality, and  $S_{N|I}$  is the Lagrangean multiplier test statistic for testing normality conditional on  $\rho = 0$ . Let  $\Omega = \text{diag} [\alpha_1 + w_1\alpha_2]$  be a  $N \times N$  diagonal matrix with element  $\alpha_1 + w_i\alpha_2$  in its  $i^{\text{th}}$  diagonal position. Let  $\zeta_4^* = (I_1 u_1^4 - 3(\alpha_1 + w_1\alpha_2)^2 \phi_1, \dots, I_N u_N^4 - 3(\alpha_1 + w_N\alpha_2)^2 \phi_N)$ . The Lagrangean multiplier statistic for the testing of the hypothesis  $H_{IN}$  is

$$S_{IN} = S_{I|N} + S_{N|I} \quad (4.32)$$

where

$$S_{I|N} = \zeta_1^{\prime \Omega^{-1/2}} D_{\phi} D_{\phi}^{-1} \ell_N [\ell_N^{\prime} D_{\phi}^2 D_{\phi}^{-1} \ell_N - \ell_N^{\prime} D_{\phi} \Omega^{-1/2} X \\ \cdot (X^{\prime} \Omega^{-1/2} D_{\phi} \Omega^{-1/2} X)^{-1} X^{\prime} \Omega^{-1/2} D_{\phi} \ell_N]^{-1}$$

$$\cdot \lambda_N D_{\Phi}^{-1} D_{\Phi} \Omega^{-1/2} \zeta_1 \quad (4.33)$$

and

$$\begin{aligned} S_{N|I} &= \frac{3}{2} \left( \zeta_1 - \frac{1}{3} \Omega^{-1} \zeta_3 \right)' \Omega^{-1} \lambda_N \left( \lambda_N' \Omega^{-1/2} D_{\Phi} \Omega^{-1/2} \lambda_N \right)^{-1} \\ &\quad \cdot \lambda_N' \Omega^{-1} \left( \zeta_1 - \frac{1}{3} \Omega^{-1} \zeta_3 \right) \\ &\quad + \frac{1}{24} \left( \Omega^{-2} \zeta_4^* - 3 \zeta_0 \right)' \lambda_N M^{-1} \lambda_N' \left( \Omega^{-2} \zeta_4^* - 3 \zeta_0 \right) \quad (4.34) \end{aligned}$$

with

$$\begin{aligned} M &= 4 \lambda_N' D_{\Phi} \lambda_N - 3 \lambda_N' D_{\Phi} \Omega^{-1} \lambda_N \left( \lambda_N' \Omega^{-1} D_{\Phi} \Omega^{-1} \lambda_N \right)^{-1} \lambda_N' D_{\Phi} \Omega^{-1} \lambda_N \\ &\quad - 3 \left[ 1 - \lambda_N' D_{\Phi} \Omega^{-1} \lambda_N \left( \lambda_N' \Omega^{-2} D_{\Phi} \lambda_N \right)^{-1} \right]^2 \lambda_N' D_{\Phi} \Omega^{-1} W \\ &\quad \cdot \left[ W' \Omega^{-1} D_{\Phi} \Omega^{-1} W - W' \Omega^{-1} D_{\Phi} \lambda_N \left( \lambda_N' \Omega^{-1} D_{\Phi} \Omega^{-1} \lambda_N \right)^{-1} \right. \\ &\quad \left. \cdot \lambda_N' D_{\Phi} \Omega^{-1} W \right]^{-1} W' \Omega^{-1} D_{\Phi} \lambda_N \end{aligned}$$

The statistic  $S_{IN}$  is evaluated at the MLE of the model  $M_{IN}$  and is asymptotically chi-square distributed with three degrees of freedom.

## 5. CONCLUSION

In this article, we have derived some Lagrangean multiplier tests for the normal distributional and homoscedastic disturbances assumptions

in the censored regression model. We assume that the distribution of the disturbances is a member of the Pearson family. The Pearson family of distributions is attractive since it contains distributions with various shapes and the normal distribution as a special case. We have also considered the joint test of the absence of selection bias, homoscedasticity and normality assumptions. The Lagrangean multiplier statistic for testing such joint hypotheses is the sum of three conditional Lagrangean multiplier statistics. Each of them is a test of one hypothesis conditional on the other two hypotheses. They can also be interpreted as tests of significance of coefficients in some regression equations. Those tests are derived from some instrumental variables estimators. It has been shown that the test of significance based on the two-stage estimator in Heckman [1976] is also a conditional Lagrangean multiplier test for selection bias in the censored regression model. We have also pointed out that, for some very unusual cases, e.g., the choice equation does not have explanatory variables or the choice probabilities are constant for all individuals, the conditional Lagrangean multiplier tests for selectivity bias have no power and are not equivalent, for large samples, to the likelihood ratio tests. This provides, probably, the first set of examples to confirm to the conjecture in Silvey [1959, p. 399] that there may be cases for which the maximum likelihood ratio principle is superior to the Lagrangean multiplier principle.

APPENDIX A: CONDITIONAL LAGRANGEAN MULTIPLIER TESTS  
FOR SELECTIVITY BIAS, HOMOSCEDASTICITY  
AND NORMALITY

In this appendix, we will derive the conditional Lagrangean multiplier test statistics for each hypothesis conditional on the other two hypotheses being true.

Conditional Lagrangean Multiplier Test for Selection Bias:

First let us consider the test of the hypothesis  $H_{I|NH}: \rho = 0$  in the model  $M_{NH}$  where the disturbances in equation (2.1) are homoscedastic and are normally distributed. The loglikelihood function for the model  $M_{NH}$  is

$$L_{NH} = \sum_{i=1}^N \left\{ (1 - I_i) \ln(1 - \Phi(z_i \gamma)) - \frac{1}{2\alpha_1} I_i (y_i - x_i \beta)^2 - \frac{1}{2} I_i \ln(2\pi\alpha_1) + I_i \ln \phi \left[ \frac{z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta)}{\sqrt{1 - \rho^2}} \right] \right\} \quad (A.1)$$

The first-order derivatives are

$$\frac{\partial L_{NH}}{\partial \beta} = \sum_{i=1}^N \left\{ \frac{1}{\alpha_1} I_i (y_i - x_i \beta) x_i' + \frac{\rho}{\sqrt{\alpha_1 (1 - \rho^2)}} I_i \frac{\phi \left[ \frac{z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta)}{\sqrt{1 - \rho^2}} \right]}{\Phi \left[ \frac{z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta)}{\sqrt{1 - \rho^2}} \right]} x_i' \right\}$$



$$\frac{\partial L_{NH}}{\partial \gamma} = \sum_{i=1}^N \left\{ - (1 - I_i) \frac{\phi(z_i \gamma)}{1 - \Phi(z_i \gamma)} z_i' + \frac{1}{\sqrt{1 - \rho^2}} I_i \right.$$

$$\left. \cdot \frac{\phi \left[ \left[ z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta) \right] / \sqrt{1 - \rho^2} \right]}{\Phi \left[ \left[ z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta) \right] / \sqrt{1 - \rho^2} \right]} z_i' \right\}$$

$$\frac{\partial L_{NH}}{\partial \rho} = \sum_{i=1}^N I_i \frac{\phi \left[ \left[ z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta) \right] / \sqrt{1 - \rho^2} \right]}{\Phi \left[ \left[ z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta) \right] / \sqrt{1 - \rho^2} \right]}$$

$$\cdot \frac{\left( \rho z_i \gamma - \frac{1}{\sqrt{\alpha_1}} (y_i - x_i \beta) \right)}{(1 - \rho^2)^{3/2}}$$

$$\frac{\partial L_{NH}}{\partial \alpha_1} = \sum_{i=1}^N \left\{ \frac{1}{2\alpha_1} I_i (y_i - x_i \beta)^2 - \frac{1}{2\alpha_1} I_i + I_i \right.$$

$$\left. \cdot \frac{\phi \left[ \left[ z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta) \right] / \sqrt{1 - \rho^2} \right]}{\Phi \left[ \left[ z_i \gamma - \frac{\rho}{\sqrt{\alpha_1}} (y_i - x_i \beta) \right] / \sqrt{1 - \rho^2} \right]} \frac{\rho (y_i - x_i \beta)}{2\alpha_1 \sqrt{\alpha_1} (1 - \rho^2)} \right\}$$

Let  $\phi_i = \phi(z_i\gamma)$  and  $\Phi_i = \Phi(z_i\gamma)$ . Evaluated at  $\rho = 0$ , the first-order derivatives are reduced to

$$\left. \frac{\partial L_{NH}}{\partial \rho} \right|_{H_I|NH} = - \sum_{i=1}^N \frac{1}{\sqrt{\alpha_1}} \frac{\phi_i}{\Phi_i} I_i u_i$$

$$\left. \frac{\partial L_{NH}}{\partial \beta} \right|_{H_I|NH} = \sum_{i=1}^N \frac{1}{\alpha_1} x_i' I_i u_i$$

$$\left. \frac{\partial L_{NH}}{\partial \gamma} \right|_{H_I|NH} = \sum_{i=1}^N \frac{\phi_i}{\Phi_i(1-\Phi_i)} z_i'(I_i - \Phi_i)$$

$$\left. \frac{\partial L_{NH}}{\partial \alpha_1} \right|_{H_I|NH} = \sum_{i=1}^N \left\{ -\frac{1}{2\alpha_1} (I_i - \Phi_i) + \frac{1}{2\alpha_1} \left[ \frac{1}{\alpha_1} I_i u_i^2 - \Phi_i \right] \right\}$$

and the information matrix is

$$E_{H_{I|NH}} \begin{pmatrix} \frac{\partial L_{NH}}{\partial \theta} & \frac{\partial L_{NH}}{\partial \theta'} \end{pmatrix} =$$

$$\begin{bmatrix} \sum_{i=1}^N \frac{\phi_i^2}{\Phi_i} & - \sum_{i=1}^N \frac{1}{\sqrt{\alpha_1}} \phi_i x_i & 0 & 0 \\ - \sum_{i=1}^N \frac{1}{\sqrt{\alpha_1}} \phi_i x_i' & \sum_{i=1}^N \frac{1}{\alpha_1} \phi_i x_i' x_i & 0 & 0 \\ 0 & 0 & \sum_{i=1}^N \frac{\phi_i^2}{\Phi_i(1-\Phi_i)} z_i' z_i & 0 \\ 0 & 0 & 0 & \frac{1}{2\alpha_1^2} \sum_{i=1}^N \Phi_i \end{bmatrix}$$

where  $\theta' = (\rho, \beta', \gamma', \alpha_1)$ . It follows that the conditional Lagrangean multiplier test statistic for  $\rho = 0$  is

$$\frac{1}{\alpha_1} \zeta_1' D_\phi D_\phi^{-1} \lambda_N \left[ \lambda_N' D_\phi^2 D_\phi^{-1} \lambda_N - \lambda_N' D_\phi X (X' D_\phi X)^{-1} X' D_\phi \lambda_N \right]^{-1} \lambda_N' D_\phi D_\phi^{-1} \zeta_1 \quad (\text{A.2})$$

evaluated at the restricted MLE in the model  $M_{INH}$ , where the matrices are the same ones as defined in the main text. The statistic (A.2) is exactly equal to the statistic  $S_{I|NH}$  in (4.13).

Conditional Lagrangean Multiplier Test for Homoscedasticity:

The loglikelihood function for the model  $M_{IN}$  where the disturbances  $u_i$  are normally distributed and are independent with  $\varepsilon_i$  is

$$L_{IN} = \sum_{i=1}^N \left\{ (1 - I_i) \ln(1 - \Phi(z_i \gamma)) - \frac{1}{2(\alpha_1 + w_i \alpha_2)} I_i (y_i - x_i \beta)^2 \right. \\ \left. - \frac{1}{2} I_i \ln 2\pi - \frac{1}{2} I_i \ln(\alpha_1 + w_i \alpha_2) \right. \\ \left. + I_i \ln \Phi(z_i \gamma) \right\} \quad (A.3)$$

Taking the first-order derivatives of  $L_{IN}$  and then evaluating them at the hypothesis  $H_{H|IN}$ :  $\alpha_2 = 0$ , we have

$$\left. \frac{\partial L_{IN}}{\partial \beta} \right|_{H_{H|IN}} = \sum_{i=1}^N \frac{1}{\alpha_1} x_i' I_i u_i$$

$$\left. \frac{\partial L_{IN}}{\partial \gamma} \right|_{H_{H|IN}} = \sum_{i=1}^N \frac{\phi_i}{\Phi_i(1 - \Phi_i)} z_i' (I_i - \Phi_i)$$

$$\left. \frac{\partial L_{IN}}{\partial \alpha_1} \right|_{H_{H|IN}} = \sum_{i=1}^N \left\{ -\frac{1}{2\alpha_1} (I_i - \Phi_i) + \frac{1}{2\alpha_1} \left[ \frac{1}{\alpha_1} I_i u_i^2 - \Phi_i \right] \right\}$$

$$\left. \frac{\partial L_{IN}}{\partial \alpha_2} \right|_{H_{H|IN}} = \sum_{i=1}^N \left\{ -\frac{1}{2\alpha_1} w_i' (I_i - \Phi_i) + \frac{1}{2\alpha_1} w_i' \left[ \frac{1}{\alpha_1} I_i u_i^2 - \Phi_i \right] \right\}$$

and the corresponding information matrix is

$$E_{H|IN} \begin{pmatrix} \frac{\partial L_{IN}}{\partial \theta} & \frac{\partial L_{IN}}{\partial \theta'} \end{pmatrix} = \begin{bmatrix} \sum_{i=1}^N \frac{1}{\alpha_1} \phi_i x_i' x_i & 0 & 0 & 0 \\ 0 & \sum_{i=1}^N \frac{\phi_i^2}{\phi_i(1-\phi_i)} z_i' z_i & 0 & 0 \\ 0 & 0 & \frac{1}{2\alpha_1^2} \sum_{i=1}^N \phi_i & \frac{1}{2\alpha_1^2} \sum_{i=1}^N \phi_i w_i \\ 0 & 0 & \frac{1}{2\alpha_1^2} \sum_{i=1}^N \phi_i w_i' & \frac{1}{2\alpha_1^2} \sum_{i=1}^N \phi_i w_i' w_i \end{bmatrix}$$

It follows that the Lagrangean multiplier test for  $\alpha_2 = 0$  is

$$\frac{1}{2} \left( \frac{1}{\alpha_1} \zeta_2 - \zeta_0 \right)' W \left[ W' D_{\Phi} W - W D_{\Phi} \ell_N (\ell_N' D_{\Phi} \ell_N)^{-1} \ell_N' D_{\Phi} W \right]^{-1} W' \left( \frac{1}{\alpha_1} \zeta_2 - \zeta_0 \right) \quad (\text{A.4})$$

which is exactly the statistic  $S_{H|IN}$  in (4.14).

#### Conditional Lagrangean Multiplier Test for Normality:

The loglikelihood function for the model  $M_{IH}$ , where the disturbances  $u_i$  are homoscedastic and  $u_i$  and  $\epsilon_i$  are independent, is

$$L_{IH} = \sum_{i=1}^N \left\{ (1 - I_i) \ln(1 - \Phi(z_i \gamma)) + I_i \ln \Phi(z_i \gamma) \right. \\ \left. + I_i q(y_i - x_i \beta) - I_i \ln \int_{-\infty}^{\infty} \exp q(u) du \right\} \quad (A.5)$$

where

$$q(u) = \int \frac{c_1 - u}{\alpha_1 - c_1 u + c_2 u^2} du$$

The first-order derivatives of  $L_{IH}$  evaluated at the hypothesis  $H_{N|IH}$ :

$c_1 = c_2 = 0$  are

$$\left. \frac{\partial L_{IH}}{\partial \beta} \right|_{H_{N|IH}} = \sum_{i=1}^N \frac{1}{\alpha_1} x_i' I_i u_i$$

$$\left. \frac{\partial L_{IH}}{\partial \gamma} \right|_{H_{N|IH}} = \sum_{i=1}^N \frac{\phi_i}{\phi_i (1 - \phi_i)} z_i' (I_i - \phi_i)$$

$$\left. \frac{\partial L_{IH}}{\partial \alpha_1} \right|_{H_{N|IH}} = \sum_{i=1}^N \left\{ \frac{1}{2\alpha_1} \left[ \frac{1}{\alpha_1} I_i u_i^2 - \phi_i \right] - \frac{1}{2\alpha_1} (I_i - \phi_i) \right\}$$

$$\left. \frac{\partial L_{IH}}{\partial c_1} \right|_{H_{N|IH}} = \sum_{i=1}^N \left\{ \frac{1}{\alpha_1} I_i u_i - \frac{1}{3\alpha_1^2} I_i u_i^3 \right\}$$

$$\left. \frac{\partial L_{IH}}{\partial c_2} \right|_{H_{N|IH}} = \sum_{i=1}^N \left\{ \frac{3}{4} \left[ \frac{1}{3\alpha_1^2} I_i u_i^4 - \phi_i \right] - \frac{3}{4} (I_i - \phi_i) \right\}$$

and the corresponding information matrix is

$$\begin{bmatrix} \frac{1}{\alpha_1} \sum_{i=1}^N \phi_i x_i' x_i & 0 & 0 & 0 & 0 \\ 0 & \sum_{i=1}^N \frac{\phi_i^2}{\phi_i(1-\phi_i)} z_i' z_i & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2\alpha_1^2} \sum_{i=1}^N \phi_i & 0 & \frac{3}{2\alpha_1} \sum_{i=1}^N \phi_i \\ 0 & 0 & 0 & \frac{2}{3\alpha_1} \sum_{i=1}^N \phi_i & 0 \\ 0 & 0 & \frac{3}{2\alpha_1} \sum_{i=1}^N \phi_i & 0 & 6 \sum_{i=1}^N \phi_i \end{bmatrix}$$

It follows that the Lagrangean multiplier test statistic for  $c_1 = 0$  and  $c_2 = 0$  is

$$\begin{aligned} & \frac{3}{2\alpha_1} \left( \zeta_1 - \frac{1}{3\alpha_1} \zeta_3 \right)' \lambda_N \left( \lambda_N' D_{\phi} \lambda_N \right)^{-1} \lambda_N' \left( \zeta_1 - \frac{1}{3\alpha_1} \zeta_3 \right) \\ & + \frac{1}{24} \left( \frac{1}{\alpha_1} \zeta_4 - 3\zeta_0 \right)' \lambda_N \left( \lambda_N' D_{\phi} \lambda_N \right)^{-1} \lambda_N' \left( \frac{1}{\alpha_1} \zeta_4 - 3\zeta_0 \right) \end{aligned} \quad (\text{A.6})$$

evaluated at the restricted MLE in the model  $M_{INH}$ . This statistic is exactly the statistic  $S_{N|IH}$  in (4.15).

## APPENDIX B: LAGRANGEAN MULTIPLIER TESTS FOR PROBIT MODEL

Lagrangean Multiplier tests can also be derived for testing the normality and homoscedasticity assumptions in the probit model. We will assume that the true distribution is a member of the Pearson-type distribution.

Probit Model

Consider the choice equation  $y_i^* = z_i\gamma - \varepsilon_i$ ,  $i = 1, \dots, N$  in (2.2). Since only dichotomous indicators are observable, we adopt the normalization  $\alpha_1 = 1$  for convenience. With this normalization, the general Pearson distribution of  $\varepsilon$  allowing heteroscedasticity is

$$f(\varepsilon) = \exp \left[ \int \frac{c_1 - \varepsilon}{1 + w\alpha - c_1\varepsilon + c_2\varepsilon^2} d\varepsilon \right] / \int_{-\infty}^{\infty} \exp \left[ \int \frac{c_1 - t}{1 + w\alpha - c_1t + c_2t^2} dt \right] dt \quad (\text{B.1})$$

Let  $F(z\gamma; 1 + w\alpha, c_1, c_2) = \int_{-\infty}^{z\gamma} f(\varepsilon) d\varepsilon$  denote the probability that  $I = 1$ . The loglikelihood function is

$$L_D = \sum_{i=1}^N \left\{ I_i \ln F(z_i\gamma; 1 + w_i\alpha, c_1, c_2) + (I - I_i) \ln [1 - F(z_i\gamma; 1 + w_i\alpha, c_1, c_2)] \right\} \quad (\text{B.2})$$

For the cases that all exogenous variables are discrete, chi-square goodness of fitness for testing normality can be derived from the



contingency table. Thus the above approach based on the general Pearson density function is of interest only for the cases that some of the regressors of  $z_i$  are continuous variables. The joint test for normality and homoscedasticity is to test the hypothesis  $H_{NH}$ :  $c_1 = c_2 = 0$  and  $\alpha = 0$ . Denote  $\phi_i = \phi(z_i; \gamma)$  and  $\Phi_i = \Phi(z_i; \gamma)$ . Define a vector  $X(i)$  as

$$X(i) = \begin{bmatrix} z_i' \\ -\frac{1}{2} (z_i; \gamma) w_i' \\ -\frac{1}{3} [(z_i; \gamma)^2 - 1] \\ -\frac{1}{4} (z_i; \gamma) [3 + (z_i; \gamma)^2] \end{bmatrix}$$

The first-order derivatives evaluated at the hypothesis  $H_{NH}$  of the likelihood  $L_D$  are

$$\frac{\partial L_D}{\partial \theta} \Big|_{H_{NH}} = \sum_{i=1}^N \frac{\phi_i}{\phi_i(1-\phi_i)} X(i)(I_i - \Phi_i) \quad (B.3)$$

where  $\theta' = (\gamma', \alpha', c_1, c_2)$  is the parameter vector. The information matrix under  $H_{NH}$  is

$$E_{H_{NH}} \left( \frac{\partial L_D}{\partial \theta} \frac{\partial L_D}{\partial \theta'} \right) = \sum_{i=1}^N \frac{\phi_i^2}{\phi_i(1-\phi_i)} X(i)X'(i) \quad (B.4)$$

Let  $I_\theta$  be an identity matrix of dimension  $k_2 + m + 2$  and  $J_{NH}$  be a submatrix of  $I_\theta$  consisting of the columns corresponding to the subvector  $(\alpha', c_1, c_2)$ . A Lagrangean multiplier test for the hypothesis  $H_{NH}$  is

$$\left( \sum_{i=1}^N (I_i - \phi_i) \frac{\phi_i}{\phi_i(1-\phi_i)} X'(i) \right) J_{NH} J'_{NH} \left( \sum_{i=1}^N \frac{\phi_i^2}{\phi_i(1-\phi_i)} X(i)X'(i) \right)^{-1} \\ \cdot J_{NH} J'_{NH} \left( \sum_{i=1}^N \frac{\phi_i}{\phi_i(1-\phi_i)} X(i)(I_i - \phi_i) \right) \quad (B.5)$$

evaluated at the probit MLE of  $\gamma$ . The degree of freedom of this asymptotic chi-square statistic is the dimension of  $w$  plus two, namely,  $m + 2$ .

## FOOTNOTES

\* I appreciate having financial support from the National Science Foundation under Grant SES-8006481 to the University of Minnesota. I have greatly benefited from the conversations with Dr. K. R. Sawyer and R. J. Olsen. Any errors are of my own.

1. The general solution of the differential equation (2.3) is

$$\ln g(u) = \int \frac{a+u}{b_0 + b_1 u + b_2 u^2} du + c$$

where  $c$  is a constant. Since

$$\int_{-\infty}^{\infty} g(u) du = 1$$

$e^{-c}$  equals the denominator in (2.4).

2. These have been derived according to the formulae in (3.22) and (3.26).
3. This is so because  $I_i y_i$  is equal to zero when  $y_{1i}$  is censored.
4. It can be easily shown that  $E_{H_{NH}}(I_i \xi_i) = (\Phi(z_i \gamma) - 1) \sigma_{uE} \phi(z_i \gamma)$ .
5. I am indebted to Randall Olsen for this observation.
6. More general cases correspond to the situations that the variable  $\Phi(z_i \hat{\gamma}) x_i$  and  $\phi(z_i \hat{\gamma})$  are linearly dependent. These cases will not occur if either  $x_i$  do not contain dummy variable or  $z_i$  contains some continuous variables.

## REFERENCES

- Aigner, D. J., C. A. K. Lovell, and P. Schmidt [1977], "Formulation and Estimation of Stochastic Frontier Production Function Models," Journal of Econometrics 6, pp. 21-37.
- Aitchison, J., and S. D. Silvey [1958], "Maximum-Likelihood Estimation of Parameters Subject to Restraints," Annals of Mathematical Statistics 29, pp. 813-828.
- Amemiya, T. [1973], "Regression Analysis When the Dependent Variable Is Truncated Normal," Econometrica 41, pp. 997-1016.
- Amemiya, T. [1974], "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal," Econometrica 42, pp. 999-1012.
- Bera, A. K., and C. M. Jarque [1980], "An Efficient Large-Sample Test for Normality and Homoscedasticity of Regression Residuals," manuscript, Australian National University.
- Breusch, T. S., and A. R. Pagan [1979], "A Simple Test for Heteroscedasticity and Random Coefficient Variation," Econometrica 47, pp. 1287-1294.
- Breusch, T. S., and A. R. Pagan [1980], "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics," Review of Economic Studies XLVII, pp. 239-253.
- Elderton, W. P., and N. L. Johnson [1969], Systems of Frequency Curves, Cambridge University Press, Cambridge, London.
- Engle, R. F. [1978], "The Lagrange Multiplier Test as a Regression Diagnostic," Discussion Paper 78-6, University of California, San Diego.
- Goldberger, A. S. [1980], "Abnormal Selection Bias," Discussion paper No. 8006, Social Systems Research Institute, University of Wisconsin, Madison.
- Heckman, J. [1976], "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," Annals of Economic and Social Measurement 5, pp. 475-492.
- Huang, C. J., and B. W. Bolch [1974], "On the Testing of Regression Disturbances for Normality," Journal of American Statistical Association 69, pp. 330-335.
- Hurd, M. [1979], "Estimation in Truncated Samples When There Is Heteroscedasticity," Journal of Econometrics 11, pp. 247-258.

- Johnson, N. L., and S. Kotz [1970], Continuous Univariate Distributions--1, Houghton Mifflin Co., Boston.
- \_\_\_\_\_[1972], Distributions in Statistics: Continuous Multivariate Distributions, John Wiley & Sons, Inc., New York.
- Lee, L. [1979], "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables," Econometrica 47, pp. 977-996.
- Lee, L. F. [1980], "Generalized Econometric Models with Selectivity," manuscript, University of Florida.
- Lee, L. F., G. S. Maddala, and R. P. Trost [1980], "Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity," Econometrica 48, pp. 491-503.
- Maddala, G. S., and F. D. Nelson [1975], "Specification Errors in Limited Dependent Variable Models," NBER Working Paper 96.
- Mardia, K. V. [1970], Families of Bivariate Distributions, Hafner Pub. Co., Connecticut.
- Nelson, F. D. [1979], "The Effect of and A Test for Misspecification in the Censored-Normal Model," Social Science Working Paper 291, California Institute of Technology.
- Olsen, R. J. [1979], "Tests for the Presence of Selectivity Bias and Their Relation to Specifications of Functional Form and Error Distribution," Yale University, Institution for Social and Policy Studies Working Paper 812.
- Rao, C. R. [1973], Linear Statistical Inference and Its Applications, 2nd edition, John Wiley & Sons, New York.
- Silvey, S. D. [1959], "The Lagrangean Multiplier Test," Annals of Mathematical Statistics 30, pp. 389-407.
- White, H., and G. M. MacDonald [1980], "Some Large Sample Tests for Non-normality in the Linear Regression Model," Journal of American Statistical Association 75, pp. 16-28.