Testing Heteroscedasticity on Stochastic Frontier Models

Paulo de Melo Jorge Neto
Professor do Curso de Mestrado em Economia (CAEN) da Universidade Federal do Ceará (UFC) Ph.D. em Economia pela University of Illinois at Champaign-Urbana

Abstract:

This paper develops a Lagrange Multiplier test for heterosdedasticity on stochastic frontier models. The test is developed after a brief description of these models and some estimation methods.

Key-Words:

Stochastic frontier model; Heteroscedasticity; Lagrange multiplier.

1 - INTRODUCTION

This paper develops a test for heteroscedasticity on stochastic frontier model. Heteroscedasticity is expected to occur because larger firms have more factors under their control than smaller firms (CAUDILL & FORD, 1993), what reduces the inefficiency on the large firms and affects the error term that has truncated normal distribution. These authors observed some undesirable properties of the maximum likelihood estimator for the parameters of stochastic frontier model when heteroscedasticity is present.

To test for heteroscedasticity on the error term corresponding to the measurement of inefficiency, we develop a Lagrange Multiplier test-LM test. This approach follows the same method as the Breusch-Pagan test (BREUSCH-PAGAN, 1979) for heteroscedasticity on the traditional linear model with normality.

Following this introduction, the second section describes the stochastic frontier model. The third section shows some of the usual procedure to estimate the parameters of a linear stochastic frontier model and the fourth section develops a LM test to verify the presence of heteroscedasticity on the error term corresponding to the measurement of inefficiency.

The problem of heteroscedasticity may introduce serious problem of inconsistency on the maximum likelihood estimator-MLE. The MLE is not consistent on some models that have non-normal heteroscedastic errors but that consider homoscedasticity to formulate the likelihood function. This sort of inconsistency was identified by HURD (1979) on estimation in truncated samples, by ARABMAZAR & SCHMIDT (1981) on Tobit models and generalized by JARQUE & BERA (1982) on limited dependent models. Once the stochastic frontier model considers a distribution that is the sum of a normal and a truncated normal, it s expected that the MLE may not be consistent when the errors are heteroscedastic. To verify this problem, we show that MLE of a stochastic frontier model is in fact inconsistent when heteroscedasticity occurs and is ignored.

2 - STOCHASTIC FRONTIER MODEL

There is a long discussion about what should be the best procedure to estimate a production function. By definition a production function gives the maximum possible quantity of some output, given quantities of a set of inputs. AIGNER, LOVELL & SCHMIDT (1977) proposed the stochastic frontier model[1]. In this model output is assumed to be bounded from above by a stochastic production function.

$$y_i \leq g(x_i;\hat{a}) + v_i \qquad (1)$$

where $\mathbf{v_i}$ is assumed to be independently and identically distributed as $\mathbf{N(0,s_v^2)}$. The model to be estimated must have positive error term to consider the idea of maximum. This is not possible with an error term normally distributed. Considering a positive term, $\mathbf{u_i}$, to represent the shortfall of output from the frontier, then the frontier model is expressed as:

$$y_i = g(x_i;\hat{a}) + v_i - u_i \qquad (2)$$

The error term $\mathbf{u_i}$ is generally assumed as the absolute value of a normal distribution with expected value equal to zero and variance s_u^2. The distribution of the error term on (2) is the sum of a symmetric normal random variable and a truncated normal. The first represents randomness and components outside the firm control, the second represents technical inefficiency and components under the firm control.[2]

The probability distribution of $\mathbf{u_i}$ is given by scaling down the density so that it integrates to one over the range above zero. The distribution of $\mathbf{u_i}$ is

$$\phi(u_i)=\sqrt{\frac{2}{p}}\frac{1}{s_u}e^{-\left(\frac{u_i}{s_u}\right)^2}$$

where $E[\mathbf{u_i}]=\sqrt{\frac{2}{p}}s_u$ and $V(\mathbf{u_i})=\frac{p-2}{p}s_u^2$.

[1] A Review on the discussion of the frontier model can be found on SCHMIDT (1986) and BAUER (1990).

[2] STEVESON (1980) presented a more general error specification where the truncation is not assumed to be zero but any value.

Letting $\varepsilon = v - u$[3], AIGNER *et al* (1977) suggest that the probability density function of ε is:[4]

$$h(\varepsilon) = (2\sigma^{-1})f(\varepsilon\sigma^{-1})[1 - F(\varepsilon\lambda\sigma^{-1}), \quad -\infty \le \varepsilon \le +\infty \qquad (3)$$

where $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \sigma_u^2 / \sigma_v^2$, and $f(\cdot)$ and $F(\cdot)$ are the standard normal density and distribution functions, respectively. Note that $E[\varepsilon]=-E[u]$ and that $V(\varepsilon)=\frac{\pi-2}{\pi}\sigma_u^2+\sigma_v^2$.

Considering the distribution (3), equation (2) can be estimated by maximum likelihood. As the mean of ε is $-(2/\pi)^{1/2}\sigma_u$, it is necessary a correction on the constant term to use OLS. The corrected OLS is given by using the estimate of σ_u to convert the OLS estimate of the constant. A consistent estimator for σ_u^2 is given in OLSON, SCHMIDT & WALDMAM (1980), where they perform a Monte Carlo study of these estimators on finite sample.

Note that one major importance of the frontier model is the facility to measure inefficiency. One possibility for inefficiency measurement is $E(u/\varepsilon)$, evaluated at the fitted value of ε. An estimator for $E(u/\varepsilon)$ is given in JONDROW, LOVELL & SCHMIDT (1982) and LEE (1983). WALDMAN (1984) examined three alternative estimator of inefficiency. He examined the conditional expectation of the function, a linear prediction that ignores the stochastic nature of the frontier and the best linear prediction and concluded that the conditional expectation is preferred. The author argues that the conditional expectation is the best because it takes advantage of the form of the distribution function.

ROBINSON & NIXON (1991) considered the cost version and introduced heteroscedasticity by assuming that there may be factors affecting the magnitude of λ, the ratio of the inefficiency to the normal standard error. The authors did not consider any variation on σ. It is not clear however, how the ratio can vary and σ be fixed. This is possible only if increasing one variance is compensated by a proportionally decreasing the other.

CAUDILL & FORD (1993) investigated the effects of heteroscedasticity on the parameters in frontier regression models. They claim that larger firms have more factors under their control than smaller firms, **u** should be heteroscedastic. They perform a monte carlo experiment to investigate the biases due to heteroscedasticity in the one-sided error term and observed an overestimation of the intercept and underestimation of the slope coefficients and the two-sided variance.

3 - ESTIMATION PROCEDURE

Before developing the test it is valid to elaborate some comments on the estimation of the parameters $\theta'=(\beta,\sigma_v^2,\sigma_u^2)$ of a linear model. The model to be analyzed is the linear version of equation (2), where $g(x;\beta)=X\beta$. **X** is **(nxk)** matrix and $\boldsymbol{\beta}$ is a **(kx1)** vector. Following the approach of fair (1977) to compute tobit estimator, GREENE (1982) proposed a similar algorithm. LEE (1983) observed that the iterative algorithm suggested by GREENE (1982) for the estimation of stochastic frontier production model does not necessarily solve the likelihood equation. LEE (1983) considered some moments of **u** conditional on $\boldsymbol{\varepsilon}$ that can be introduced on the system of equations (5), (6) and (7). These moments are developed on JONDROW *et al* (1982) and correspond to:

[3] We drop the subscript **i** that represents an observation.
[4] See STEVENSON (1980).

$$E(u_i/å_i) = ó_u ó_v ó^{-1} ã_i - ó_u^{2} ó^{-2} å_i \quad (4)$$

$$E(u_i^{2}/å_i) = ó_u^{2} ó^{-2} (ó_v^{2} - å_i E(u_i/å_i)) \quad (5)$$

where $ó^2 = ó_u^2 + ó_v^2$. Considering â instead of ì and after some algebra:

$$\frac{\partial lnL}{\partial \boldsymbol{b}} = \frac{1}{\boldsymbol{s}_v^{2}} \sum_{i=1}^{n} x_i(\boldsymbol{e}_i + E(u_i/\boldsymbol{e}_i)) = 0 \quad (6)$$

$$\frac{\partial lnL}{\partial \boldsymbol{s}_u^{2}} = \frac{1}{2\boldsymbol{s}_u^{4}} \sum_{i=1}^{n} (E(u_i^{2}/\boldsymbol{e}_i) - \boldsymbol{s}_u^{2}) = 0 \quad (7)$$

$$\frac{\partial lnL}{\partial \boldsymbol{s}_v^{2}} = \frac{1}{2\boldsymbol{s}_v^{4}} \sum_{i=1}^{n} (\boldsymbol{e}_i + E(u_i^{2}/\boldsymbol{e}_i) + 2\boldsymbol{e}_i E(u_i/\boldsymbol{e}_i) - \boldsymbol{s}_v^{2}) = 0 \quad (8)$$

An explicit expression for the variances is obtained from (7) and (8). Equation (6) can be understood as a modified lest square. The procedural is as follows: compute equation (6) with the variance calculated from (7) and (8); proceed with new interaction and continue computing until convergence. LEE (1983) adverted that the Newton method is faster. Note that some robust estimator is necessary to avoid specification problems.[5] If the model is not well specified, that is, if the model does not consider the presence of heteroscedastic errors when this is the case, then one may not initiate the computation with a robust estimator. It is then necessary to estimate the heteroscedastic components.

On the next section, a LM test is developed to examine heteroscedasticity on the stochastic frontier model.

4 - LM TEST FOR HETEROSCEDASTICITY

To consider the presence of heteroscedasticity, let g(x;â)=**Xâ** be the linear version of the stochastic frontier model[6], where **X** is (nxk) matrix and **â** is a (kx1) vector. We use the same specification for heteroscedasticity as in BREUSCH & PAGAN (1979), in which heteroscedasticity is introduced assuming that $\mathbf{ó_u{}^2{}_i = h(á_o + z_i{}'á_{-1}) = h(z_i{}'á)}$. $\mathbf{a_o}$ is a (nx1) vector of ones, $\mathbf{á_{-1}}$ is a (px1) and **a** is a ((p+1)x1) vector and **z** is a (nx(p+1)) matrix whose first column is a column of ones. The function h(·) is assumed to be twice differentiable and **z** may include some elements from **X**. The MLE is obtained by solving the following system:

[5] AMEMIYA (1973) develops an interative method to estimate the consistent parameters of model with truncated dependent variable.

[6] Note that the method used here is valid for the non-linear case.

$$\frac{\partial lnL}{\partial \beta} = \sum_{i=1} \frac{1}{(\sigma^2_{ui} + \sigma^2_v)} \varepsilon_i \vec{x}_i + \sum_{i=1}^{n} \gamma_i \vec{x}_i \left(\frac{\sigma_{ui}}{\sigma_v} \frac{1}{\sqrt{\sigma^2_{ui} + \sigma^2_v}} \right) \tag{9}$$

$$\frac{\partial lnL}{\partial \sigma_v^2} = -\sum_{i=1}^{n} \frac{1}{2(\sigma^2_{ui} + \sigma^2_v)} + \sum_{i-1}^{n} \frac{1}{2(\sigma^2_{ui} + \sigma^2_v)} \varepsilon_i^2 + \sum_{i=1}^{n} \frac{\sigma_{ui}(\sigma^2_{ui} + 2\sigma^2_v)}{2\sigma_v^3 (\sigma^2_{ui} + \sigma^2_v)^{3/2}} \gamma_i \varepsilon_i \tag{10}$$

$$\frac{\partial lnL}{\partial \alpha} = \sum_{i=1} \frac{h'}{2} \left[-\frac{1}{(\sigma^2_{ui} + \sigma^2_v)} + \frac{1}{(\sigma^2_{ui} + \sigma^2_v)} \varepsilon_i^2 - \frac{\sigma_v}{\sigma_{ui}(\sigma^2_{ui} + \sigma^2_v)^{3/2}} \gamma_i \varepsilon_i \right] \vec{z}_{i'} \tag{11}$$

Note that $\frac{\partial lnL}{\partial \beta}$ is a (kx1) vector and that $\frac{\partial lnL}{\partial \alpha}$ is a ((p+1)x1) vector.

The null hypothesis of homoscedasticity is $H_0: \alpha_1 = ... = \alpha_p = 0$. In this case $\sigma_{u\,i}^2 = h(\alpha_o) = \sigma_u^2$. Under the null, the restricted MLE must satisfy equations (6) and (7) the following restricted value of (8):

$$\frac{\partial lnL}{\partial \alpha} = \frac{h'(\alpha_0)}{2\sigma_u^4} \sum_{i=1}^{n} (E(u_i^2 / \varepsilon_i) - \sigma_u^2) z_i = 0 \tag{12}$$

Let **d=(∂lnl/∂θ)** describe the score vector. $\hat{d}$ is partitioned as $d = (d'_{\theta_1}, d'_{\alpha})$, where $\theta_1' = (\beta, \sigma_v^2, \sigma_u^2(\alpha_0))$. From the constrained ML, we have that $\hat{d}_{\theta_1}=0$. Let $\hat{d}_{\alpha}$ be the score vector evaluated under the restriction of the null hypothesis. Let now **I** be the information matrix partioned as

$$\mathbf{I} = \begin{bmatrix} I_{\theta_1\theta_1} & I_{\theta_1\alpha} \\ I_{\alpha\theta_1} & I_{\alpha\alpha} \end{bmatrix}$$

The LM is defined as:[7]

$$LM = \hat{d}_{\alpha'} \hat{I}^{\alpha\alpha} \hat{d}_{\alpha} \tag{13}$$

where $\hat{I}^{\alpha\alpha}$ is the inverse of $\hat{I}_{\alpha\alpha}$.

The high non-linearity on the distribution of the error term makes the computation of the information matrix very difficult. It can be seen from the appendix 1 that the expression for **E((∂lnL/∂θ)(∂lnL/∂θ'))** has a difficult form to compute. To obtain such terms, it is necessary to evaluate $E(\varepsilon^2)$, $E(\varepsilon^3)$, $E(\varepsilon^4)$, $E(\varepsilon^5)$, $E(\gamma)$, $E(\gamma^2)$, $E(\varepsilon\gamma)$, $E(\varepsilon^2\gamma)$, $E(\varepsilon^3\gamma)$, $E(\varepsilon^2\gamma^2)$, $E(\varepsilon\gamma^2)$. Instead of getting expected values, we compute the information matrix without expectation[8].

$$\hat{I} = \sum_i \hat{g}_i \hat{g}_{i'} \tag{14}$$

where $\mathbf{g_i}$ is a ((k+1+(p+1))x1) vector corresponding to the first derivative of a single information (observation) of likelihood function evaluated at the restricted MLE.

$$\hat{I}(\hat{\theta}) = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} \hat{I}_{\hat{\theta}_1\hat{\theta}_1} & \hat{I}_{\hat{\theta}_1\alpha} \\ \hat{I}_{\alpha\hat{\theta}_1} & \hat{I}_{\alpha\alpha} \end{bmatrix} \tag{15}$$

[7] It is assumed that there is no misspecification on the error term distribution. See appendix for detailed calculation.

[8] GREENE (1993) argues that this method produces values closer to data. See also the discussion on DAVIDSON & MACKINNON (1993).

where $b_{33}=\hat{I}_{\alpha\alpha}$. The value of each element is given in the appendix 1.

The LM test can then be evaluated from the following formula:

$$LM=\frac{h'(\hat{\alpha}_0)^2}{4\,\hat{\sigma}_u^8}\Big[\sum_{i=1}^{n}(E(\hat{u}_i^2/\hat{\varepsilon}_i)-\hat{\sigma}_u^2)z_i\Big]'$$
$$x\,[\hat{I}_{\alpha\alpha}-\hat{I}_{\alpha\hat{\theta}_1}\hat{I}^{-1}_{\hat{\theta}_1\hat{\theta}_1}\hat{I}_{\hat{\theta}_1\alpha}] \qquad (16)$$
$$x\Big[\sum_{i=1}^{n}(E(\hat{u}_i^2/\hat{\varepsilon}_i)-\hat{\sigma}_u^2)z_i\Big]$$

It is expected that this statistic follow a χ^2 distribution with p degrees of freedom.[9]

Unfortunately this is a very expensive test. Nothing can be said a priori about the diagonal form of the information matrix to simplify the inverse of $\hat{I}_{\alpha\alpha}$.

Note that this test resembles the value of the traditional Breusch-Pagan test, where the difference comes from the last term of following expression:

$$\frac{\partial lnL}{\partial\alpha}=\frac{h'(\alpha_0)}{2(\sigma_u^2+\sigma_v^2)^2}\sum_{1}^{n}\Big[-(\sigma_u^2+\sigma_v^2)+\varepsilon_i^2$$
$$-\frac{\sigma_v}{\sigma_u(\sigma_u^2+\sigma_v^2)^{1/2}}\gamma_i\varepsilon_i\Big]z_{i'}$$

The last term expresses the influence of the truncated part of the error term. BREUSCH & PAGAN (1979) argued that their test is very simple, one needs only to compute values from an OLS regression and evaluated the test. However, BREUSCH & PAGAN (1980) suggested that when the MLE estimator under the null involves non-linear form, them the attractiveness of LM test seems to disappear.

Note that the high cost to calculate the LM test proposed arises not only because the LM statistics is difficult to calculate, but also because the computation of the MLE is very expensive. LEE (1983) proposed a method to calculate the MLE by interaction. To simplify calculations, a $C(\alpha)$test can be used instead of the LM test. Such a test demands a consistent estimator under the null hypothesis. The consistent estimator is obtained through OLS. In fact, OLSON, SCHMIDT & WALDMAN (1980) described in their monte-carlo study about three estimators of the stochastic frontier production function that the corrected OLS is most efficient on sample smaller than 400. Even for higher sample the additional efficiency of the MLE may not be worth the extra trouble required to compute it.[10]

The formula of the $C(\alpha)$, that has χ^2 distribution with p degrees of freedom, test is given by

$$C(\alpha)=(\hat{d}_{\bar{\theta}_1}-\hat{I}_{\bar{\theta}_1\alpha}\hat{I}^{-1}_{\alpha\alpha}\hat{d}_\alpha)'\hat{I}^{\bar{\theta}_1\bar{\theta}_1}(\hat{d}_{\bar{\theta}_1}-\hat{I}_{\bar{\theta}_1\alpha}\hat{I}^{-1}_{\alpha\alpha}\hat{d}_\alpha)$$

Where $\bar{\theta}_1$ is the OLS estimator.

5 - CONCLUSION

In order to identify the presence of heteroscedasticity, a LM was developed. However, no simple test form was developed given the high non-linearity of maximum likelihood function. Given the computational complexity of the MLE, the $C(\alpha)$ test seems to be a good alternative. A different approach to test heteroscedasticity is found on the information matrix test for misespecification developed by MALLICK (1994).

We did not consider in this paper the possibility of the MLE be inconsistent. Many authors have identified the inconsistency of MLE on model without normally distributed errors. JARQUE & BERA (1982) argued that violation of homocedasticity on the peculiar nature of limited dependent model not only makes MLE inconsistent but also makes tests, as Breusch-Pagan, not applicable. As the frontier model consists of a particular characterization of the error component, a mix of normal and truncated normal, the misestimation problem should be present under heteroscedasticity, of the asymmetric error, using a maximum likelihood

[9] It can be checked that the test is invariant with respect to the functional form of **h**. This is done decomposing the determinants and evaluating the test form.

[10] Their study does not consider any sort of heterosce-dasticity.

approach. We let the analysis of this problem of inconsistency on MLE to further research.

Resumo:

Este trabalho desenvolve um teste do tipo "Lagrange Multiplier" para testar a possibilidade de heterocedasticidade nos modelos de fronteira estocástica. O teste é desenvolvido após uma breve descrição das características desses e de alguns método de estimação.

Palavras-Chaves:

Modelos de Fronteira Estocástica; Heterocedasticidade; Multiplicador de Lagrange.

7 - REFERENCE

AIGNER, D., LOVELL, C., SCHMIDT, P. Formulation and estimation of sthocastic frontier production function models. **Journal of Econometrics**, v. 6, p. 21-37, 1977.

AMEMIYA, T. Regression analysis when the dependent variable is truncated normal. **E-conometrica**, v.41, n. 6, p. 997-1016, 1973.

BERA, A., JARQUE, C., LEE, L. Testing normality assumption in limited dependent variable models. **International Economic Review**, v. 25, n. 3, p. 563-578, 1984.

BAUER, P. Recent developments in econometric estimation of frontiers. **Journal of Econometrics**, v. 46, p. 39-56, 1990.

BREUSCH, T., PAGAN. A. The lagrange multiplier test and its applications to model specification in econometrics. **Review of Economics Studies**, v. 47, p. 239-253, 1980.

_____.A simple test for heteroscedasticity and random coeficient variation. **Econometrica**, v. 47. n. 5, p. 1287-1294, 1979.

CAUDILL, S., FORD, J. Biases in frontier estimation due to heteroscedasticity. **Economic Letters**, v. 41, p. 17-20, 1993.

DAVIDSON, R., MACKINNON, J. **Estimation and inference in econometric**. Oxford: Oxford University Press, 1993.

FAIR, R. A note on the computation of the tobit estimator. **Econometrica**, v. 45, n. 7, p. 1723-1727, 1977.

GODFREY, G. **Misspecification test in econometrics**. the lagrange multiplier principle and other approaches. Cambridge: Cambridge University Press, 1988.

GREENE, W. **Econometrics analysis**. 2. ed. New York: Macmillan, 1993.

_______. Maximun likelihood estimation of stochastic frontier production models. **Journal of Econometrics**. v. 18, p.285-289, 1982.

JARQUE, C.M., BERA, A.K. Eficient test for normality heteroskedasticity and serial independence of regression residuals. **Economic Letters**, v. 6, p. 255-259, 1980.

JONDROW, J., LOVELL, C., SCHMIDT, P. On the estimation of technical inefficiency in the stochastic frontier production function model. **Journal of Econometrics**, v. 19, p. 233-238, 1982.

LEE, L. On maximum likelihood estimation of stochastic frontier production models. **Journal of Econometrics,** v. 23, p. 269-274, 1983.

MALLICK, N. **Specification test for frontier models**. Illinois: University of Illinois at Urbana-Champaing, 1994. (Mimeogr.).

OLSON, J., SCHMIDT, P., WALDMAN, D. A Monte Carlo study of estimators of stochastic frontier production functions. **Journal of Econometrics**, v. 13, p. 67-82, 1980.

ROBISON, M., NIXON, L. Heteroskedasticity inefficiency in a cost frontier model. **Economic Letters**, v. 36, p. 413-417,1991.

SCHMIDT, P. Frontier production functions. **Econometric Reviews**, v. 4, n. 2, p. 289-328, 1986.

SCHMIDT, P., LOVELL, C. Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers. **Journal of Econometrics**, v. 9, p. 343-366, 1979.

STEVENSON, R. Likelihood functions for generalized stochastic frontier estimation. **Journal of Econometrics**, v.13, p. 57-66, 1980.

WALDMAN, D. Properties of technical efficiency estimators in the stochastic frontier model. **Journal of Econometrics**, v. 25, p. 353-364, 1984.

_____. A stationary point for the stochastic frontier likelihood. **Journal of Econometrics**, v. 18, p. 275-279, 1982.

Recebido para publicação em 27.AGO.1999.

APPENDIX 1

The information matrix $\hat{I}\,(\beta, \sigma_v^2, \alpha)$ is computed as:

$$\hat{I}(\hat{\theta}) = \begin{bmatrix} \sum \frac{\partial \ln \hat{L}_i}{\partial \hat{\beta}} \frac{\partial \ln \hat{L}_i}{\partial \hat{\beta}'} & \sum \frac{\partial \ln \hat{L}_i}{\partial \hat{\beta}} \frac{\partial \ln \hat{L}_i}{\partial \hat{\sigma}_v^2} & \sum \frac{\partial \ln \hat{L}_i}{\partial \hat{\beta}} \frac{\partial \ln \hat{L}_i}{\partial \hat{\alpha}^*} \\ \sum \frac{\partial \ln \hat{L}_i}{\partial \hat{\sigma}_v^2} \frac{\partial \ln \hat{L}_i}{\partial \hat{\beta}} & \sum \left(\frac{\partial \ln \hat{L}_i}{\partial \hat{\sigma}_v^2} \right)^2 & \sum \frac{\partial \ln \hat{L}_i}{\partial \hat{\sigma}_v^2} \frac{\partial \ln \hat{L}_i}{\partial \hat{\alpha}^*} \\ \sum \frac{\partial \ln \hat{L}_i}{\partial \hat{\alpha}^*} \frac{\partial \ln \hat{L}_i}{\partial \hat{\beta}} & \sum \frac{\partial \ln \hat{L}_i}{\partial \hat{\alpha}^*} \frac{\partial \ln \hat{L}_i}{\partial \hat{\sigma}_v^2} & \sum \frac{\partial \ln \hat{L}_i}{\partial \hat{\alpha}^*} \frac{\partial \ln \hat{L}_i}{\partial \hat{\alpha}^{*\prime}} \end{bmatrix}$$

In simple form we have:

$$\hat{I}(\hat{\theta}) = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{12}' & b_{22} & b_{23} \\ b_{13}' & b_{23}' & b_{33} \end{bmatrix}$$

Where:

$$b_{11} = \frac{1}{(\hat{\sigma}_v^2)^2} \sum \left[\hat{\varepsilon}_i + \hat{E}(u/\varepsilon_i) \right]^2 X_i X_i'$$

$$b_{12} = \frac{1}{2\hat{\sigma}_v^2 (\hat{\sigma}_u^2)^2} \sum \left[\hat{\varepsilon}_i^2 + \hat{E}(u^2/\varepsilon_i) + 2\hat{\varepsilon}_i \hat{E}(u/\varepsilon_i) - \hat{\sigma}_v^2 \right] \left[\hat{\varepsilon}_i + \hat{E}(u/\varepsilon_i) \right] X_i$$

$$b_{13} = \frac{h'(\alpha_0)}{2\hat{\sigma}_v^2 (\hat{\sigma}_u^2)^2} \sum \left[\hat{E}(u^2/\varepsilon_i) - \hat{\sigma}_u^2 \right] \left[\hat{\varepsilon}_i + \hat{E}(u/\varepsilon_i) \right] Z_i' X_i'$$

$$b_{22} = \frac{1}{4(\hat{\sigma}_u^2)^4} \sum \left[\hat{\varepsilon}_i + \hat{E}(u^2/\varepsilon_i) + 2\hat{\varepsilon}_i \hat{E}(u/\varepsilon_i) - \hat{\sigma}_v^2 \right]^2$$

$$b_{23} = \frac{h'(\alpha_0)}{4(\hat{\sigma}_v^2)^2 (\hat{\sigma}_u^2)^2} \sum \left[\hat{E}(u^2/\varepsilon_i) - \hat{\sigma}_u^2 \right] \left[\hat{\varepsilon}_i^2 + \hat{E}(u^2/\varepsilon_i) + 2\hat{\varepsilon}_i \hat{E}(u/\varepsilon_i) - \hat{\sigma}_v^2 \right]$$

$$b_{33} = \frac{(h'(\alpha_0))^2}{4(\hat{\sigma}_u^2)^4} \sum \left[\hat{E}(u^2/\varepsilon_i) - \hat{\sigma}_u^2 \right]^2 Z_i' Z_i^{*\prime}$$

with:

$$\hat{E}(u/\varepsilon_i) = \frac{\sqrt{\hat{\sigma}_u^2}\sqrt{\hat{\sigma}_v^2}}{\sqrt{\hat{\sigma}_u^2 + \hat{\sigma}_v^2}}\hat{\gamma}_i - \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_v^2}\hat{\varepsilon}_i$$

$$\hat{E}(u^2/\varepsilon_i) = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_v^2}\left(\hat{\sigma}_v^2 - \hat{\varepsilon}_i \hat{E}(u/\varepsilon_i)\right)$$

$$\hat{E}(u/\varepsilon_i) = \frac{\phi\left(\dfrac{\sqrt{\hat{\sigma}_u^2}}{\sqrt{\hat{\sigma}_v^2}\sqrt{\hat{\sigma}_u^2 + \hat{\sigma}_v^2}}\hat{\varepsilon}_i\right)}{1 - \Phi\left(\dfrac{\sqrt{\hat{\sigma}_u^2}}{\sqrt{\hat{\sigma}_v^2}\sqrt{\hat{\sigma}_u^2 + \hat{\sigma}_v^2}}\hat{\varepsilon}_i\right)}$$

Note that $\hat{V}(\hat{\theta}) = \dfrac{\hat{I}(\hat{\theta})}{n}$

The detailed values for each component are given as:

$$\left(\frac{\partial \ln L_i}{\partial \alpha^*}\right)^2 = \frac{h'(\alpha_0)^2}{4\sigma_u^8}\left[E(u^2/\varepsilon_i) - \sigma_u^2\right]^2 Z_i Z_i'$$

$$= \frac{h'(\alpha_0)^2}{4\sigma_u^8 \sigma^3}\left[\sigma_u^2 \sigma_v^2 \sigma \gamma_i^2 - 2\sigma_u^3 \sigma_v \varepsilon_i \gamma_i + \sigma_u^4 \sigma \varepsilon_i^2 - 2\sigma_u^3 \sigma_v \sigma^2 \gamma_i + 2\sigma_u^4 \sigma \varepsilon_i + \sigma_u^4 \sigma\right] Z_i Z_i'$$

$$\frac{\partial \ln L_i}{\partial \alpha}\frac{\partial \ln L_i}{\partial \beta} = \frac{h'(\alpha_0)}{2\sigma_v^2 \sigma_u^4}\left[E(u^2/\varepsilon_i) - \sigma_u^2\right]\left[\varepsilon_i + E(u/\varepsilon_i)\right] Z_i' X_i$$

$$= \frac{h'(\alpha_0)}{2\sigma_v^2 \sigma_u^4}\left[\varepsilon_i\left(\frac{\sigma_v^2 \sigma_u^2}{\sigma^2} - \frac{\sigma_v^2 \sigma_u^4}{\sigma^4} + \frac{\sigma_u^4}{\sigma^2} - \sigma_u^2\right) + \varepsilon_i^3\left(\frac{\sigma_u^4}{\sigma^4} - \frac{\sigma_u^6}{\sigma^6}\right) + \gamma_i\left(\frac{\sigma_v^3 \sigma_u^3}{\sigma^3} - \frac{\sigma_v \sigma_u^3}{\sigma}\right) + \right.$$

$$\left. + \varepsilon_i^2 \gamma_i\left(-\frac{\sigma_v \sigma_u^3}{\sigma^3} + 2\frac{\sigma_v \sigma_u^5}{\sigma^5}\right) + \varepsilon_i \gamma_i^2\left(-\frac{\sigma_v^2 \sigma_u^4}{\sigma^4}\right)\right] Z_i X_i'$$

$$\left(\frac{\partial \ln L_i}{\partial \sigma_v^2}\right)^2 = \frac{1}{4\sigma_v^8}\left[\varepsilon_i^2 + E(u^2/\varepsilon_i) + 2\varepsilon_i E(u/\varepsilon_i) - \sigma_v^2\right]^2$$

$$= \frac{1}{4\sigma_v^8}\left[\varepsilon_i^2\left(1 + 2\frac{\sigma_v^2 \sigma_u^6}{\sigma^6} + 2\frac{\sigma_v^2 \sigma_u^2}{\sigma^2} - 4\frac{\sigma_v^2 \sigma_u^4}{\sigma^4} - 2\frac{\sigma_v^2 \sigma_u^3}{\sigma^4} - 4\frac{\sigma_v \sigma_u^2}{\sigma^2} - 2\sigma_v^2\right) + \right.$$

$$+ \varepsilon_i^4\left(1 + \frac{\sigma_u^8}{\sigma^8} + 2\frac{\sigma_u^4}{\sigma^4} - 4\frac{\sigma_u^2}{\sigma^2} + 4\frac{\sigma_u^4}{\sigma^4}\right) + \varepsilon_i^5\left(-4\frac{\sigma_u^6}{\sigma^6}\right) + \varepsilon_i \gamma_i\left(-2\frac{\sigma_v^3 \sigma_u^5}{\sigma^5} + 4\frac{\sigma_v^3 \sigma_u^3}{\sigma^3} + 2\frac{\sigma_v^3 \sigma_u^3}{\sigma^3} - 4\frac{\sigma_v^3 \sigma_u}{\sigma}\right) +$$

$$+ \varepsilon_i^3 \gamma_i\left(-8\frac{\sigma_v \sigma_u^3}{\sigma^3} + 4\frac{\sigma_v \sigma_u}{\sigma} + 4\frac{\sigma_v \sigma_u^5}{\sigma^5} - 2\frac{\sigma_v \sigma_u^7}{\sigma^7} - 2\frac{\sigma_v \sigma_u^3}{\sigma^3} + 8\frac{\sigma_v \sigma_u^4}{\sigma^4}\right) + \varepsilon_i^2 \gamma_i^2\left(\frac{\sigma_v^2 \sigma_u^6}{\sigma^6} + 4\frac{\sigma_v^2 \sigma_u^2}{\sigma^2} - 4\frac{\sigma_v^2 \sigma_u^4}{\sigma^4}\right) +$$

$$+\left(\frac{\sigma_v^4\sigma_u^4}{\sigma^4}-\sigma_v^4-2\frac{\sigma_v^4\sigma_u^2}{\sigma^2}\right)\Bigg]\Bigg]$$

$$\frac{\partial \ln L_i}{\partial \beta}\frac{\partial \ln L_i}{\partial \beta}=\frac{1}{\sigma_v^4}\left[\varepsilon_i+E(u/\varepsilon_i)\right]^2 X_i X'_i$$

$$=\frac{1}{s_v^4}\left[e_i^2\left(1+\frac{s_u^4}{s^4}\right)+e_i g_i\left(2\frac{s_v s_u}{s}-2\frac{s_v s_u^3}{s^3}\right)+g_i^2\left(\frac{s_v^2 s_u^2}{s^2}\right)\right]X_i X'_i$$

$$\frac{\partial \ln L_i}{\partial \beta}\frac{\partial \ln L_i}{\partial \sigma_v^2}=\frac{1}{2\sigma_v^2\sigma_u^4}\left[\varepsilon_i^2+E(u^2/\varepsilon_i)+2\varepsilon_i E(u/\varepsilon_i)-\sigma_v^2\right]\left[\varepsilon_i+E(u/\varepsilon_i)\right]X_i$$

$$=\frac{1}{2s_v^2 s_u^4}\Bigg[e_i\left(-s_v^2+\frac{s_v^2 s_u^2}{s^2}-\frac{s_v^2 s_u^4}{s^4}+\frac{s_v^2 s_u^2}{s^2}\right)+e_i^3\left(1-\frac{s_u^2}{s^2}+\frac{s_u^4}{s^4}+2\frac{s_u^4}{s^4}\right)+e_i^4\left(-\frac{s_u^6}{s^6}\right)+$$

$$+\gamma_i\left(\frac{\sigma_v^3\sigma_u^4}{\sigma^4}-\frac{\sigma_v^3\sigma_u}{\sigma}\right)+\varepsilon_i\gamma_i^2\left(-\frac{\sigma_v^2\sigma_u^4}{\sigma^4}+2\frac{\sigma_v^2\sigma_u^2}{\sigma^2}\right)+\varepsilon_i^2\gamma_i\left(-\frac{\sigma_v\sigma_u^3}{\sigma_u^3}+\frac{\sigma_v\sigma_u}{\sigma}+\frac{\sigma_v\sigma_u^5}{\sigma^5}-4\frac{\sigma_v\sigma_u^3}{\sigma^3}\right)\Bigg]X_i$$

$$\frac{\partial \ln L_i}{\partial \alpha^*}\frac{\partial \ln L_i}{\partial \sigma_v^2}=\frac{h'(\alpha_0)}{4\sigma_v^4\sigma_u^4}\left[E(u^2/\varepsilon_i)-\sigma_u^2\right]\left[\varepsilon_i^2+E(u^2/\varepsilon_i)+2\varepsilon_i E(u/\varepsilon_i)-\sigma_v^2\right]Z_i$$

$$=\frac{h'(a_0)}{4s_v^4 s_u^4}\Bigg[e_i^2\left(2\frac{s_u^6 s_v^2}{s^6}-2\frac{s_u^4 s_v^2}{s^4}-\frac{s_u^4}{s^3}+2\frac{s_u^4}{s^2}-s_u^2\right)+e_i^3\left(-\frac{s_u^2}{s^2}\right)+e_i^4\left(\frac{s_u^8}{s^8}-2\frac{s_u^6}{s^6}\right)+$$

$$+\varepsilon_i\gamma_i\left(-2\frac{\sigma_v^4\sigma_u^5}{\sigma^5}+2\frac{\sigma_v^3\sigma_u^3}{\sigma^3}+\frac{\sigma_v\sigma_u^3}{\sigma^2}-2\frac{\sigma_v\sigma_u^3}{\sigma}\right)+\varepsilon_i^2\gamma_i\left(\frac{\sigma_v\sigma_u}{\sigma}\right)+\varepsilon_i^3\gamma_i\left(-2\frac{\sigma_u^7\sigma_v}{\sigma^7}+4\frac{\sigma_v\sigma_u^5}{\sigma^5}\right)+$$

$$+\varepsilon_i^2\gamma_i^2\left(\frac{\sigma_v^2\sigma_u^6}{\sigma^6}-2\frac{\sigma_v^2\sigma_u^4}{\sigma^4}\right)+\left(\frac{\sigma_v^4\sigma_u^4}{\sigma^4}-\frac{\sigma_v^2\sigma_u^2}{\sigma}+\sigma_v\sigma_u\right)\Bigg]Z_i$$