# Testing for Heteroskedasticity in Bivariate Probit Models

by

Thomas Thorn
B.Sc., University of Victoria, 2009

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF ARTS

in the Department of Economics

# Supervisory Committee

Testing for Heteroskedasticity in Bivariate Probit Models

by

Thomas Thorn
B.Sc., University of Victoria, 2011

**Supervisory Committee**

Dr. David E. Giles, (Department of Economics)
**Supervisor**

Dr. Judith A. Clarke, (Department of Economics)
**Departmental Member**

# Abstract

**Supervisory Committee**
Dr. David E. Giles, (Department of Economics)
**Supervisor**
Dr. Judith A. Clarke, (Department of Economics)
**Departmental Member**

Two score tests for heteroskedasticity in the errors of a bivariate Probit model are developed, and numerous simulations are performed. These tests are based on an outer product of the gradient estimate of the information matrix, and are constructed using an artificial regression. The empirical sizes of both tests are found to be well-behaved, settling down to the nominal size under the asymptotic distribution as the sample size approaches 1000 observations. Similarly, the empirical powers of both tests increase quickly with sample size. The largest improvement in power occurs as the sample size increases from 250 to 500. An application with health care data from the German Socioeconomic Panel is performed, and strong evidence of heteroskedasticity is detected. This suggests that the maximum likelihood estimator for the standard bivariate Probit model will be inconsistent in this particular case.

# Table of Contents

# List of Tables

# List of Figures

# Acknowledgments

I am extremely grateful for the unending support of my supervisor, Dr. David E. Giles. I would like to thank Steven Jones for pushing me to apply to university, all those years ago, and my brother Alan for always providing moral support. I would also like to express my appreciation to Dr. Judith A. Clarke for her numerous helpful suggestions and Lief Bluck for arranging the required computing facilities. Finally, I am indebted to the Social Sciences and Humanities Research Council of Canada for financial assistance.

# Chapter 1 Introduction

When analyzing economic data, economists often wish to estimate a model with a discrete or qualitative dependent variable. In such instances, a 'qualitative dependent variable' model, such as a Logit or Probit model, is often used to obtain appropriate coefficient estimates. Empirical economists have found qualitative dependent variable models to be an essential tool for examining a wide variety of issues ranging from bankruptcy to labour force participation. In the absence of such models, a wide variety of problems would remain outside the scope of econometric estimation and inference. Equally, these models have also found widespread application in other disciplines.

An empirical economist may wish to model two equations whose errors may be correlated in some way. In an ordinary least squares framework, the economist in question would likely use a Seemingly Unrelated Regression (SUR) model. Using a SUR model would allow the economist to test parameter restrictions that involve different equations while increasing the efficiency of estimation. The gains associated with using SUR models have led to their widespread use in the applied econometric literature. The gains associated with using a SUR model, as opposed to separately estimating two equations, are exactly analogous to the gains in using a bivariate qualitative dependent variable model over two separate qualitative dependent variable models. While Logit and Probit models are widely used in single-equation econometric analysis, Probit models are estimated far more often in the two-equation (bivariate) case. The bivariate, two-equation, Probit model is frequently used in empirical economics, particularly when examining cross-sectional data. It is widely recognized, however, that heteroskedasticity is often present with cross-sectional data. This presents a major problem because heteroskedasticity of the errors renders the maximum likelihood estimators of any qualitative dependent variable model, including the bivariate Probit model, inconsistent. There are methods, of course, for taking into account heteroskedasticity when obtaining coefficient estimates, eliminating this inconsistency. These methods, however, require that one knows the specific form of the heteroskedasticity that is present in the model that is being estimated. If the heteroskedasticity is incorrectly specified, the coefficient estimator will generally still be rendered inconsistent. Unfortunately, no tests for heteroskedasticity in bivariate Probit models currently exist, and this is what this thesis sets out to address.

There are therefore two major motivations for developing tests for heteroskedasticity for a bivariate Probit model. First, an empirical researcher may be uncertain as to whether or not heteroskedasticity is present in a model and will therefore wish to test for its presence. Second, an empirical researcher may be aware that heteroskedasticity is present in the model they are estimating. In this case they will wish to test whether or not the form of heteroskedasticity they have assumed while estimating the model has dealt with the problem adequately.

This thesis begins with a brief literature review. Two new (score) tests for heteroskedasticity are then derived, and a computational model is presented for examining the empirical size and power of these tests in a Monte Carlo experiment. The results of this set of statistical simulations are then presented and interpreted. The tests are then applied to a particular set of health care data, after which a discussion of future areas of research is followed by a brief conclusion. Finally, three appendices provide the R code used in the computational model, additional empirical size results, and tables showing the empirical rejection rates of the new tests under a variety of sample sizes and parameter values.

# Chapter 2 Literature Review

The impact of heteroskedasticity on the standard estimators for single-equation Probit models has been thoroughly explored in the existing literature. Yatchew and Griliches (1985), for example, derived exact expressions for the impact of such misspecification on Probit models. Most relevant to this thesis, these authors found that although heteroskedasticity leads to inconsistent parameter estimators, small amounts of heteroskedasticity result in the parameter estimators being only slightly rescaled. This can be solved, of course, by estimating the parameters while taking into account any heteroskedasticity. Unfortunately, however, the form of heteroskedasticity is typically not known in advance. In addition, if a researcher assumes a specific form of heteroskedasticity, they may assume an erroneous specification. In this instance, there is value in testing the corrected model to see if any heteroskedasticity remains. Yatchew and Griliches further find that other model misspecifications result in inconsistent estimators.

A number of model specification tests for single equation Probit and Logit models already exist in the literature. Thomas (1993), for instance, developed a Lagrange Multiplier (LM) test for the validity of assuming the logistic distribution in the case of the Logit model. This LM test was shown to both have favorable empirical size and power properties relative to similar information matrix and a Regression Equation Specification Error Test. More generally, Smith (1989) derived a broad test for distributional misspecification for parametric qualitative dependent variable models, given a specific null distribution. Both of these sets of tests are important because distributional misspecification will result in asymptotically biased estimators (Smith, 1989). In the same vein, Murphy (2007) derived a score test for normality in bivariate Probit models. This test is also rather important because bivariate Probit estimators are also inconsistent in the event of non-normality. Monte Carlo simulations performed by Murphy illustrated that his score test had an empirical size that was close to its nominal value and strong power properties.

Through the development of two artificial regressions, Davidson and MacKinnon (1984) were able to generate five different test statistics for various model misspecification issues associated with single-equation Logit and Probit models. The first of these artificial regressions involves regressing a gradient vector of the log likelihood function, against a vector of ones in the form

$$\boldsymbol{\iota} = G(\widetilde{\boldsymbol{\beta}})\boldsymbol{b} + \boldsymbol{errors},$$

where $\boldsymbol{\iota}$ is a vector of ones, $\boldsymbol{b}$ is a coefficient vector, $\widetilde{\boldsymbol{\beta}}$ refers to the maximum likelihood estimator of the coefficient vector of a Probit model, and a typical element of the log-likelihood gradient vector $G_{ij}(\widetilde{\boldsymbol{\beta}};\ y_i)$ is

$$G_{ij}(\widetilde{\boldsymbol{\beta}};\ y_i) = \left[y_i\Phi_2(\boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})^{-1} + (y_i - 1)\Phi_2(-\boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})^{-1}\right] \varphi_2(\boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}})\, \boldsymbol{x}_{\mathrm{ij}}\ .$$

$\Phi_2$ refers to the standard normal cumulative distribution function, while $\varphi_2$ refers to the standard normal probability density function and $\boldsymbol{x}_i'\widetilde{\boldsymbol{\beta}}$ refers to the fitted values obtained from estimating a Probit model for $\boldsymbol{y}_{\mathrm{i}}$ with the covariate matrix $\boldsymbol{x}_{\mathbf{ij}}$. This artificial regression produces an LM-statistic, hereafter denoted $LM_1$, and an F-statistic denoted $F_1$. They also considered a second artificial regression:

$$\boldsymbol{r}(\widetilde{\boldsymbol{\beta}}) = \boldsymbol{R}(\widetilde{\boldsymbol{\beta}})\boldsymbol{c} + \boldsymbol{errors},$$

where $\boldsymbol{c}$ is a coefficient vector, a typical element of $r_i(\tilde{\boldsymbol{\beta}}; y_i)$ is

$$r_i(\tilde{\boldsymbol{\beta}}; y_i) = y_i \left[\frac{\Phi_2(-\boldsymbol{x}_i\tilde{\boldsymbol{\beta}})}{\Phi_2(\boldsymbol{x}_i\tilde{\boldsymbol{\beta}})}\right]^{1/2} + (y_i - 1)\left[\frac{\Phi_2(\boldsymbol{x}_i\tilde{\boldsymbol{\beta}})}{\Phi_2(-\boldsymbol{x}_i\tilde{\boldsymbol{\beta}})}\right]^{1/2},$$

and a typical element of $R_{ij}(\tilde{\boldsymbol{\beta}})$ is

$$R_{ij}(\tilde{\boldsymbol{\beta}}) = \left[\Phi_2(\boldsymbol{x}_i\tilde{\boldsymbol{\beta}})\Phi_2(-\boldsymbol{x}_i\tilde{\boldsymbol{\beta}})\right]^{-1/2} \varphi_2(\boldsymbol{x}_i\tilde{\boldsymbol{\beta}})\, \boldsymbol{x}_{ij}.$$

This second artificial regression produces three test statistics: $F_2$, $LM_2$ and $nR^2$. Davidson and MacKinnon also performed a number of simulations to determine the empirical size and power of these two sets of tests. These simulations revealed that the members of the second set of tests have a better empirical size than those of the first set, which performed poorly with small sample sizes. Examining the power properties of these tests, the authors found that $LM_1$ has equal or superior power properties to those of the likelihood ratio or $LM_2$ tests, both in terms of raw and size-adjusted rejection rates. The one exception to $LM_1$'s strong power properties is, in fact, in the case most relevant to this paper—when testing for heteroskedasticity. In this instance, size-adjusted rejection rates for $LM_1$ were worse than those of $LM_2$ and the likelihood ratio test. The authors attribute this to the non-linear nature of the Probit model making the size-corrected critical values unreliable. When raw rejection rates were used, the likelihood ratio test displayed higher power for a sample size of 50 (followed by $LM_1$ and $LM_2$, respectively). $LM_1$ dominated the other tests for sample sizes of 100 and 200. Since, when testing for heteroskedasticity, $LM_1$ had a higher empirical size and generally had higher raw empirical power than $LM_2$ and the size-corrected rejection rates are unreliable, it is impossible say to which set of tests has better power properties.

Murphy (1994) extended Davidson and MacKinnon's second artificial regression to test for model misspecification problems such as omitted variables and neglected heteroskedasticity for univariate discrete choice models. Murphy (1996) also developed tests for omitted variables, heteroskedasticity and asymmetry in univariate ordered Logit models. These tests are also based on Davidson and MacKinnon's $LM_2$ and $nR^2$ test statistics. Since Murphy does not provide any statistical simulations in either of his papers, the empirical sizes and powers of his tests are unknown in small samples. Fabbri, Monfardini and Radice (2004) further adapted Davidson and MacKinnon's $LM_1$ test to check for exogeneity in bivariate Probit models. Although these authors build the same artificial regression that Davidson and MacKinnon use to generate both the $F_1$ and $LM_1$ test statistics, the former authors focus only on the $LM_1$ test statistic. In addition to this test, they also construct conditional moment, likelihood ratio and Wald test. The empirical size of the $LM_1$ test in this context was close to the nominal size under a variety of different data generating processes, provided the sample size had reached 1000 or 2000, depending on the particular data generating process. Similarly, the power properties of this test were strong relative to the other tests that the authors examined.

# Chapter 3 Score Tests for Bivariate Probit

To develop a test for heteroskedasticity, we must first specify the underlying bivariate Probit model. The general specification of the bivariate model is taken from Greene (2008), but with a variance-covariance matrix reflecting the possibility of heteroskedasticity in each equation's error term.

$$y_1^* = \boldsymbol{x}_1\boldsymbol{\beta}_1 + \epsilon_1 \qquad y_1 = 1 \quad if\ y_1^* > 0, 0\ otherwise,$$

$$y_2^* = \boldsymbol{x}_2\boldsymbol{\beta}_2 + \epsilon_2 \qquad y_2 = 1 \quad if\ y_2^* > 0, 0\ otherwise,$$

$$E(\epsilon_1 | \boldsymbol{x}_1, \boldsymbol{x}_2) = E(\epsilon_2 | \boldsymbol{x}_1, \boldsymbol{x}_2) = 0,$$

$$\Sigma_i = \begin{bmatrix} \exp(2\,\boldsymbol{z}_{i1}\boldsymbol{\gamma}_1) & \rho \exp(\boldsymbol{z}_{i2}\boldsymbol{\gamma}_1 + \boldsymbol{z}_{i2}\boldsymbol{\gamma}_2) \\ \rho \exp(\boldsymbol{z}_{i1}\boldsymbol{\gamma}_1 + \boldsymbol{z}_{i2}\boldsymbol{\gamma}_2) & \exp(2\,\boldsymbol{z}_{i2}\boldsymbol{\gamma}_2) \end{bmatrix} \ for\ i \in 1, \dots, n.$$

where $\Sigma_i$ is the variance-covariance matrix for observation i, and ρ is the correlation between the two error terms. Let $\boldsymbol{x}_1$ be a regressor matrix of dimensions (n x $k_1$); $\boldsymbol{x}_{i1}$ be an entry in that regressor matrix of dimensions (1 x $k_1$); $\boldsymbol{\beta}_1$ be a coefficient vector of dimensions ($k_1$ x 1); $\boldsymbol{x}_2$ be a regressor matrix of dimensions (n x $k_2$); $\boldsymbol{x}_{i2}$ be an entry in that regressor matrix of dimensions (1 x $k_2$); $\boldsymbol{\beta}_2$ be a coefficient vector of dimensions ($k_2$ x 1); $\boldsymbol{z}_1$ be a covariate matrix of dimensions (n x $m_1$); $\boldsymbol{z}_{i1}$ be an entry in that covariate matrix of dimensions (1 x $m_1$); $\boldsymbol{\gamma}_1$ be a coefficient vector of dimensions ($m_1$ x 1); $\boldsymbol{z}_2$ be a covariate matrix of dimensions (n x $m_2$); $\boldsymbol{z}_{i2}$ be an entry in that covariate matrix of dimensions (1 x $m_2$); and finally let $\boldsymbol{\gamma}_2$ be a coefficient vector of dimensions ($m_2$ x 1). $y_1^*$ and $y_2^*$ are unobserved latent variables of dimensions (n x 1).

To ensure that all the elements of the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ vectors are identifiable, we must specify that the $\boldsymbol{z}_{i1}$ and $\boldsymbol{z}_{i2}$ values in the variance-covariance matrix do not contain an intercept

(Davidson and MacKinnon, 1984, p.247). This form of the variance co-variance matrix is used to ensure that the variance does not take negative values. The hypothesis being tested is:

$H_0: \boldsymbol{\gamma}_1 = 0 \ and \ \boldsymbol{\gamma}_2 = 0,$

$H_A: H_0$ is not true .

This hypothesis will be tested using a Lagrange Multiplier test. There are several motivations for using an LM test rather than a Wald test or Likelihood Ratio Test (LRT). First, the LM test only requires that the model is properly specified under the null hypothesis in order for the associated maximum likelihood coefficient estimators to be consistent. For a Wald test, the model must be correctly specified under the alternative hypothesis for this property to hold. Similarly, for a likelihood ratio test to have this property, the model must be correctly specified under both the null and alternative hypotheses. To correctly specify the alternative hypothesis, one not only needs to know the components of the $\boldsymbol{z}$-matrices, one must also know the functional form of the heteroskedasticity. If either the functional form or the $\boldsymbol{z}$-matrices are mis-specified, the maximum likelihood estimators will be inconsistent under the alternative hypothesis. As a result, Wald or LRT tests will be meaningless, even as the sample size increases. An LM test, on the other hand, will experience no size distortion even if the $\boldsymbol{z}$-matrices are mis-specified or the incorrect functional form is used when describing the heteroskedasticity. LM tests will, however, be far less powerful in this situation. Finally, LM tests are less costly computationally. Overall, it is clear that, for this particular testing problem, LM tests are likely to be a superior choice.

LM tests are calculated under the null hypothesis. Under the null hypothesis, the variance-covariance matrix will simplify to

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

The degrees of freedom associated with this test depend on the dimensions of $\mathbf{z}_{i1}$ and $\mathbf{z}_{i2}$. Since $\boldsymbol{z}_{i1}$ is of dimension (1 x $m_1$) and $\boldsymbol{z}_{i2}$ is (1 x $m_2$) and it will be recalled that the constant term is omitted, the null requires ($m_1 + m_2$) restrictions be imposed on the parameters of the model. This number will be ($k_1 + k_2 - 2$) if $\boldsymbol{z}_{ij}$ is equal to $\mathbf{x}_{ij}$ (j = 1,2), for example.

As noted earlier, Davidson and MacKinnon (1984) developed a series of LM tests for model misspecification in univariate Probit models. This thesis builds on an adapted version of the first of their artificial regressions,

$$\boldsymbol{\iota} = G(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\gamma}}, \tilde{\rho})\boldsymbol{b} + \boldsymbol{errors},$$

where $G(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\gamma}}, \tilde{\rho})$ is a gradient vector described on page 13 below, to model an unknown form of heteroskedasticity. This artificial regression builds an LM and an F statistic by estimating the information matrix using the outer product of the gradient vector to obtain the Hessian matrix. The log likelihood function is originally specified as:

$$logL = \sum_{i=1}^{n} y_{i1}y_{i2}\ln(P_{11}) + \ y_{i1}(1-y_{i2})\ln(P_{10}) + \ (1-y_{i1})y_{i2}\ln(P_{01})$$

$$+ \ (1-y_{i1})(1-y_{i2})\ln(P_{00})$$

where

$$P_{00} = \Pr[y_{i1} = 0, y_{i2} = 0] = \int_{\theta_1}^{\infty}\int_{\theta_2}^{\infty} \varphi_2(x_{i1}, x_{i2}, \rho_i, z_{i1}, z_{i2})dx_{i2}dx_{i1},$$

$$P_{01} = \Pr[y_{i1} = 0, y_{i2} = 1] = \int_{\theta_1}^{\infty}\int_{-\infty}^{\theta_2} \varphi_2(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \rho_i, \boldsymbol{z}_{i1}, \boldsymbol{z}_{i2}) dx_{i2} dx_{i1} \ ,$$

$$P_{10} = \Pr[y_{i1} = 1, y_{i2} = 0] = \int_{-\infty}^{\theta_1}\int_{\theta_2}^{\infty} \varphi_2(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \rho_i, \boldsymbol{z}_{i1}, \boldsymbol{z}_{i2}) dx_{i2} dx_{i1} \ ,$$

$$P_{11} = \Pr[y_{i1} = 1, y_{i2} = 1] = \int_{-\infty}^{\theta_1}\int_{-\infty}^{\theta_2} \varphi_2(\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \rho_i, \boldsymbol{z}_{i1}, \boldsymbol{z}_{i2}) dx_{i2} dx_{i1},$$

$$\theta_1 = \frac{\boldsymbol{x}_{i1}'\boldsymbol{\beta}_1}{\exp\,(\boldsymbol{\gamma}_1\boldsymbol{z}_{i1})} \ , \ \theta_2 = \frac{\boldsymbol{x}_{i2}'\boldsymbol{\beta}_2}{\exp(\boldsymbol{\gamma}_2\boldsymbol{z}_{i2})}$$

and $\varphi_2$ is the bivariate normal probability density function and $\rho_i, \boldsymbol{z}_{i1}$ and $\boldsymbol{z}_{i2}$determine its variance covariance matrix. In order to obtain the gradient it is advantageous to transform the problem as follows:

$$q_{ij} = 2\, y_{ij} - 1\ ; \qquad \text{for j = 1,2}$$

$$w_{ij} = q_{ij}\, \boldsymbol{x}_{ij}'\boldsymbol{\beta}_j \,/\, \exp\,(\boldsymbol{\gamma}_j \boldsymbol{z}_{ij})$$

$$\rho_i^* = q_{i1}q_{i2}\rho \ .$$

As a result of this transformation, the log likelihood function becomes

$$lnL = \sum_{i=1}^{n} \ln \Phi_2(\, w_{i1}, w_{i2}, \rho_i^*),$$

where $\Phi_2$ is the bivariate normal cumulative distribution function, such that

$$Prob(Y_1 = y_{i1}, Y_2 = y_{i2} \mid \boldsymbol{x}_1, \boldsymbol{x}_2) = \Phi_2(w_{i1}, w_{i2}, \rho_i^*) \ .$$

One can then break this cumulative distribution function into the conditional and marginal distributions

$$\Phi_2(w_{i1}, w_{i2}, \rho_i^*) = \Phi(w_{i1})\Phi(w_{i2} \mid w_{i1}, \rho_i^*)$$

and

$$\Phi_2(w_{i1}, w_{i2}, \rho_i^*) = \Phi(w_{i2})\Phi(w_{i1} \mid w_{i2}, \rho_i^*) \ .$$

Since $w_{i1}$ depends on $\boldsymbol{\beta}_1$ and $\boldsymbol{\gamma}_1$, it will be easier to take the derivatives with respect to $\boldsymbol{\beta}_1$ and $\boldsymbol{\gamma}_1$ using the first joint and marginal distribution, whereas the derivatives with respect to $\boldsymbol{\beta}_2$ and $\boldsymbol{\gamma}_2$ will be easier to take with the second joint and marginal distribution.

Accordingly, the derivatives of the bivariate normal cumulative distribution function are

$$\frac{\partial \Phi_2}{\partial \boldsymbol{\beta}_1} = \varphi(w_{i1})\, \Phi(w_{i1}|\, w_{i2}, \rho_i^*)\, \frac{\partial w_{i1}}{\partial \boldsymbol{\beta}_1},$$

$$\frac{\partial \Phi_2}{\partial \boldsymbol{\beta}_2} = \varphi(w_{i2})\, \Phi(w_{i2}|\, w_{i1}, \rho_i^*)\, \frac{\partial w_{i2}}{\partial \boldsymbol{\beta}_2},$$

$$\frac{\partial \Phi_2}{\partial \boldsymbol{\gamma}_1} = \varphi(w_{i1})\, \Phi(w_{i1}|\, w_{i2}, \rho_i^*)\, \frac{\partial w_{i1}}{\partial \boldsymbol{\gamma}_1},$$

and

$$\frac{\partial \Phi_2}{\partial \boldsymbol{\gamma}_2} = \varphi(w_{i2})\, \Phi(w_{i2}|\, w_{i1}, \rho_i^*)\, \frac{\partial w_{i2}}{\partial \boldsymbol{\gamma}_2},$$

where

$$\frac{\partial w_{ij}}{\partial \beta_j} = q_{ij}\, \boldsymbol{x}_{ij}' \,/\, \exp\,(\boldsymbol{\gamma}_j \boldsymbol{z}_{ij}),$$

$$\frac{\partial w_{ij}}{\partial \gamma_j} = q_{ij}\, \boldsymbol{x}_{ij}' \boldsymbol{\beta}_j (-\boldsymbol{z}_{ij}) \,/\, \exp\,(\boldsymbol{\gamma}_j \boldsymbol{z}_{ij}) \text{ and}$$

$$\Phi(w_{i1}|\, w_{i2}, \rho_i^*) = \Phi\left[\frac{w_{i2} - \rho_j^{*2} w_{i1}}{\sqrt{1-\rho_j^{*2}}}\right].$$

The derivative of the bivariate normal cumulative distribution function with respect to the parameter for the correlation between the two error terms is

$$\frac{\partial \Phi_2}{\partial \rho} = \varphi_2(w_{i1}, w_{i2}, \rho_i^*)\, \frac{\partial ln\rho^*}{\partial \rho},$$

where

$$\frac{\partial ln\rho_i^*}{\partial \rho} = q_{i1} q_{i2}\,.$$

Returning to the log-likelihood function, $lnL = \sum_{i=1}^{n} \ln \Phi_2( w_{i1}, w_{i2}, \rho_i^*)$, we can then trivially assemble a gradient vector consisting of the elements

$$\frac{\partial lnL}{\partial \beta_j} = \sum_{i=1}^{n} \frac{1}{\Phi_2} \frac{\partial \Phi_2}{\partial w_{ij}} \frac{\partial w_{ij}}{\partial \beta_j} ,$$

$$\frac{\partial lnL}{\partial \gamma_j} = \sum_{i=1}^{n} \frac{1}{\Phi_2} \frac{\partial \Phi_2}{\partial w_{ij}} \frac{\partial w_{ij}}{\partial \gamma_j} \text{ and}$$

$$\frac{\partial lnL}{\partial \rho} = \sum_{i=1}^{n} \frac{1}{\Phi_2} \frac{\partial \Phi_2}{\partial \rho} .$$

It can then be shown that the gradient vector under the null hypothesis,

$$G(\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}, \tilde{\rho}) = \begin{bmatrix} \frac{\partial lnL}{\partial \beta_j} & \frac{\partial lnL}{\partial \gamma_j} & \frac{\partial lnL}{\partial \rho} \end{bmatrix},$$

is comprised of the three partial derivatives

$$\frac{\partial lnL}{\partial \beta_j} = \sum_{i=1}^{n} \left( \frac{q_{ij} h_{ij}}{\Phi_2} \right) \boldsymbol{x}_{ij} ,$$

$$\frac{\partial lnL}{\partial \gamma_j} = - \sum_{i=1}^{n} \left( \frac{q_{ij} h_{ij}}{\Phi_2} \right) (\boldsymbol{x}'_{ij} \boldsymbol{\beta}_j) \boldsymbol{z}_{ij} \text{ and}$$

$$\frac{\partial lnL}{\partial \rho} = \sum_{i=1}^{n} \left( \frac{q_{i1} q_{i2} \varphi_2}{\Phi_2} \right) ,$$

where

$$h_{i1} = \varphi(w_{i1}) \Phi \left[ \frac{w_{i2} - \rho_j^{*2} w_{i1}}{\sqrt{1 - \rho_j^{*2}}} \right] .$$

It should be noted that $\varphi$ and $\Phi$ refer to the univariate standard normal probability density and cumulative distribution functions, respectively. With this gradient vector, the artificial regression,

$$\boldsymbol{\iota} = G(\widetilde{\boldsymbol{\beta}}, \widetilde{\boldsymbol{\gamma}}, \tilde{\rho}) \boldsymbol{b} + \boldsymbol{errors},$$

can take place. The goal of this regression is to create the two test statistics

$$LM_1 = \boldsymbol{\iota}' \boldsymbol{G}(\widetilde{\boldsymbol{\beta}}, \tilde{\rho}, \widetilde{\boldsymbol{\gamma}}) [\boldsymbol{G}(\widetilde{\boldsymbol{\beta}}, \tilde{\rho}, \widetilde{\boldsymbol{\gamma}})' \boldsymbol{G}(\widetilde{\boldsymbol{\beta}}, \tilde{\rho}, \widetilde{\boldsymbol{\gamma}})]^{-1} \boldsymbol{G}(\widetilde{\boldsymbol{\beta}}, \tilde{\rho}, \widetilde{\boldsymbol{\gamma}})' \boldsymbol{\iota}$$

and

$$F_1 = (\frac{n-SSR}{k1+k2-2})/(\frac{SSR}{n-(k1+k2)}) ,$$

where SSR refers to the sum of squared residuals. $LM_1$ is asymptotically chi-square distributed with ($m_1$ + $m_2$) degrees of freedom, while $F_1$ is asymptotically F-distributed with ($m_1$ + $m_2$) and (n – ($k_1$ + $k_2$ + 1)) degrees of freedom. Since the dependent variable of the artificial regression is a vector of ones, y'y is equal to the sample size n. (n-SSR) is therefore the explained sum of squares of the artificial regression. This is interesting to note because the $LM_1$ test statistic is also the explained sum of squares. We should therefore not expect the size and power properties of $LM_1$ and $F_1$ to differ substantially. It should also be noted that both of the test statistics rely on estimating the information matrix with the outer product of the gradient vector. It has been found that information matrices constructed from the outer product of the gradient vector may have poor small-sample properties (Murphy 1996, p.137), but of course they are computationally convenient because they do not require the calculation of the second derivatives. The performance of our particular tests in samples of moderate size is taken up in the next chapter.

# Chapter 4 A Simulation Experiment

## 4.1 Experiment Design

Although the asymptotic properties of the two tests are known, their small-sample properties are unknown. Determining the power properties of the two new tests requires the generation of heteroskedastic errors in the bivariate Probit model. In the specification of the error structure, however, the variance of equation j is assumed to be $\exp(2\boldsymbol{\gamma}_j\mathbf{z}_{ij})$. Since the $\mathbf{z}_{ij}$ values will typically be unknown, the application of this test statistic may be problematic. For the purposes of this thesis, I will therefore focus on the case where $\mathbf{z}_{ij}$ is equal to $\mathbf{x}_{ij}$. This case is of particular interest for applied research because, in many cases, it is sensible to assume that the variance of the dependent variable may be dependent upon the regressors being used. It should again be noted that for the model parameters to remain identifiable, the constant should be omitted from the $\mathbf{z}$-matrix (Davidson and MacKinnon, 1984). In this case the relevant entry in the gradient vector becomes

$$\frac{\partial lnL}{\partial \gamma_j} = -\sum_{i=1}^{n}\left(\frac{q_{ij}h_{ij}}{\Phi_2}\right)(\boldsymbol{x}'_{ij}\boldsymbol{\beta}_j)\boldsymbol{x}_{ij}\ .$$

The degrees of freedom of the $LM_1$ and $F_1$ test statistics will therefore depend on the number of restrictions put in place for the null hypothesis, $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_2 = 0$ to hold.

A standard Monte Carlo experiment was performed to determine the empirical sizes of the tests, and to explore the empirical power of both the $F_1$ and $LM_1$ tests in finite samples. The $\boldsymbol{x}$-values used were independent standard normal variates, with the same values being used in all simulations. That is, the regressors were "fixed in repeated samples." The regressors of both equations were set to be equal, although this does not necessarily have to be the case. The same random number seed was also used throughout to ensure that the results are as comparable as possible. The following true coefficient vectors were used in all of the experiments:

$$\boldsymbol{\beta}_1 = (0.25, 0.5)', \ \ \boldsymbol{\beta}_2 = (1, 0.5)'.$$

These values were chosen to ensure that the model was not perfectly identified. Power was examined by setting $\gamma_1 = \gamma_2$, and iterating the γ-values outwards from $\gamma_1 = \gamma_2 = 0$ by reasonable increments until the resulting empirical power was sufficiently close to 100%. All simulations are performed in the R statistical environment (R Development Core Team, 2010). The 'corpcor' package was used to invert the information matrix (Opgen-Rhein, Schaefer and Strimmer, 2010). The 'VGAM' package was used to obtain maximum likelihood estimates for the β coefficients and ρ (Yee, 2011). Finally, the 'mtvnorm' package was used to obtain bivariate cumulative distribution and probability density values (Genz *et al.*, 2011).

## 4.2 Empirical Size

The empirical sizes of both tests were determined at 1%, 5% and 10% nominal significance levels. The results for 5% and 10% are found below, while the 1% values can be found in Appendix B.

**Table 1 - Empirical Size (%) at α = 5%, 1000 Repetitions**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| n | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 250 | 9.8 | 9.9 | 11.5 | 11.6 | 11.4 | 11.8 | 10.5 | 10.5 | 10.5 | 10.6 |
| 500 | 7.4 | 7.5 | 7.2 | 7.3 | 7.0 | 7.0 | 7.9 | 8.1 | 6.4 | 6.4 |
| 1000 | 6.2 | 6.2 | 6.3 | 6.3 | 6.2 | 6.2 | 5.6 | 5.6 | 6.4 | 6.4 |
| 2500 | 5.4 | 5.3 | 5.8 | 5.8 | 5.7 | 5.7 | 5.0 | 5.0 | 6.0 | 6.0 |
| 5000 | 5.8 | 5.8 | 6.0 | 6.0 | 4.6 | 4.6 | 6.1 | 6.1 | 5.5 | 5.5 |
| 7500 | 5.5 | 5.5 | 5.5 | 5.5 | 6.9 | 6.9 | 4.4 | 4.4 | 6.1 | 6.1 |

**Table 2 – Empirical Size (%) at α = 10%, 1000 Repetitions**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| n | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 250 | 17.4 | 17.4 | 18.6 | 18.8 | 17.6 | 17.7 | 17.9 | 17.9 | 18.8 | 18.8 |
| 500 | 14.9 | 14.9 | 13.1 | 13.2 | 13.4 | 13.4 | 14.4 | 14.4 | 11.9 | 11.9 |
| 1000 | 10.8 | 10.8 | 11.9 | 11.9 | 10.8 | 10.8 | 12.1 | 12.1 | 12.2 | 12.2 |
| 2500 | 12.0 | 11.9 | 11.7 | 11.7 | 11.0 | 11.0 | 11.6 | 11.6 | 10.8 | 10.8 |
| 5000 | 10.5 | 10.5 | 10.8 | 10.8 | 9.5 | 9.5 | 10.3 | 10.3 | 10.3 | 10.3 |
| 7500 | 10.7 | 10.7 | 10.9 | 10.9 | 12.0 | 12.0 | 8.7 | 8.7 | 10.5 | 10.5 |

These results suggest that there is a large improvement in the empirical size of the tests, in terms of converging to the nominal size, as the sample size increases from 250 to 1000. After this point, however, the improvements are far more modest. It should be noted that as the sample size increases, sometimes the empirical size gets slightly worse. This is a reflection of the fact that only 1000 Monte Carlo repetitions were performed. A larger number of repetitions provided more reliable results. The results below were obtained with 5000 repetitions. It is clear that the pattern observed above has disappeared.

**Table 3 – Empirical Size (%) at α = 5%, 5000 repetitions**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| n | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 250 | 10.68 | 10.78 | 10.80 | 10.88 | 11.24 | 11.40 | 10.12 | 10.24 | 10.72 | 10.78 |
| 1000 | 6.10 | 6.08 | 6.46 | 6.50 | 5.82 | 5.84 | 6.14 | 6.20 | 5.78 | 5.78 |
| 2500 | 5.52 | 5.50 | 6.02 | 6.02 | 5.04 | 5.04 | 5.50 | 5.48 | 5.46 | 5.50 |

**Table 4 – Empirical Size (%) at α = 10%, 5000 repetitions**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| n | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 250 | 17.58 | 17.62 | 16.94 | 16.98 | 17.46 | 17.52 | 16.92 | 16.92 | 17.88 | 17.90 |
| 1000 | 11.60 | 11.58 | 11.78 | 11.78 | 11.58 | 11.58 | 12.02 | 12.02 | 11.60 | 11.60 |
| 2500 | 11.00 | 10.98 | 11.62 | 11.62 | 10.36 | 10.36 | 10.82 | 10.80 | 10.72 | 10.74 |

There is no clear relationship between ρ and the empirical size, other than that the empirical size tends to be closest to its asymptotic value when ρ is equal to zero.

## 4.3 Empirical Power

The empirical power of both tests was determined with 1000 Monte Carlo repetitions and $\boldsymbol{\gamma}$-values iterated outwards until the powers of the tests were essentially 100%. Figures 1 to 3 illustrate how the power of the $LM_1$ test statistic changes as the correlation between the two equations in the bivariate Probit model changes, given a particular sample size. Conversely, Figures 4 to 6 illustrate how the power of the $LM_1$ test statistic changes as the sample size increases, given a particular correlation between the errors of the two equations. Tables containing additional power values can be found in Appendix C. In all of these tables we have set $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_2$, in order to make the results more inter-comparable. At the end of Appendix C, several simulations are reported where $\boldsymbol{\gamma}_1 \neq \boldsymbol{\gamma}_2$.

As Appendix C illustrates, $F_1$ and $LM_1$ have similar empirical power properties, although $LM_1$ test tends to have slightly higher power. This thesis will therefore focus on the power properties of the $LM_1$ test statistic. The power properties of the $F_1$ test can, however, be found in Appendix C.

Figures 1 to 3 illustrate that regardless of the sample size, the tests' powers increase as the correlation between the two error terms increases. This is a sensible result, given that as the randomly generated errors move from being perfectly correlated to being perfectly negatively correlated, the generated **y**-values will move from being very similar to being very different, given a particular set of **x**-values. This would make any heteroskedasticity more apparent, and presumably easier to detect.

Figures 4 to 6 illustrate that as sample size increases from 250 to 1000, the power of the tests improves dramatically, regardless of the value of $\rho$. This improvement is much more dramatic as the sample size increases from 250 to 500, but there is still a great deal of improvement when the sample size further increases from 500 to 1000.

**Figure 1 - Empirical Power Curve of $LM_1$ with Sample Size 250**

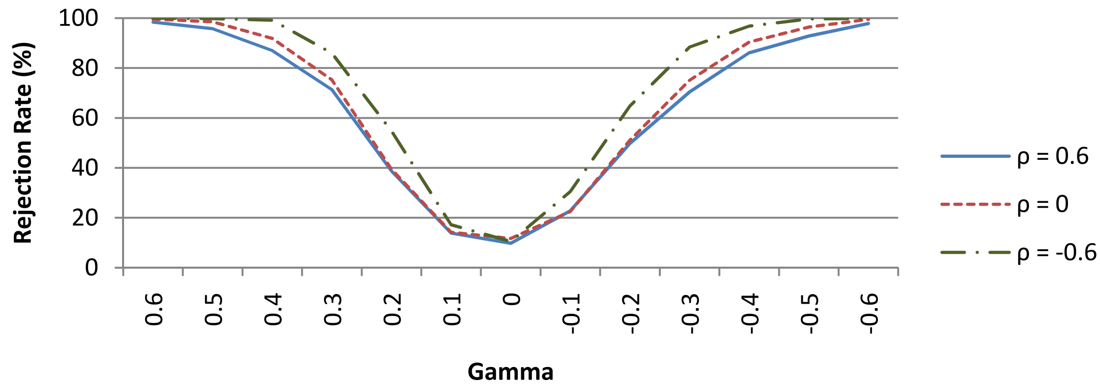



**Figure 2 - Empirical Power Curve of $LM_1$ with Sample Size 500**

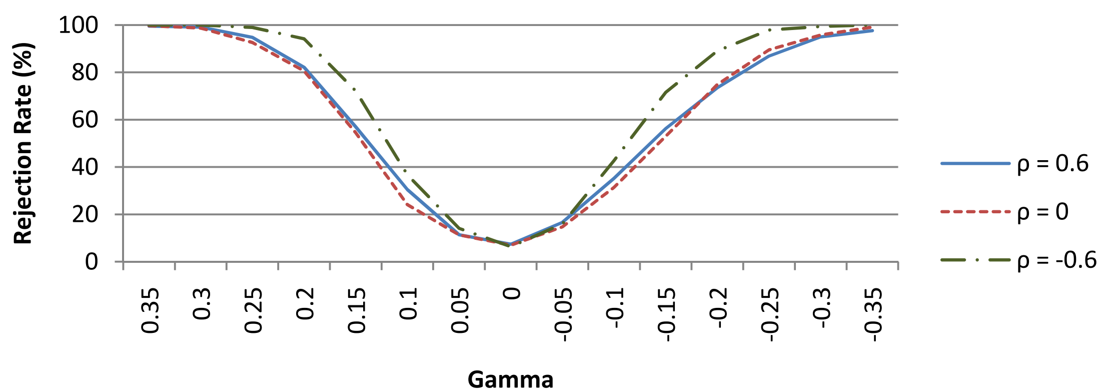



**Figure 3 - Empirical Power Curve of $LM_1$ with Sample Size 1000**

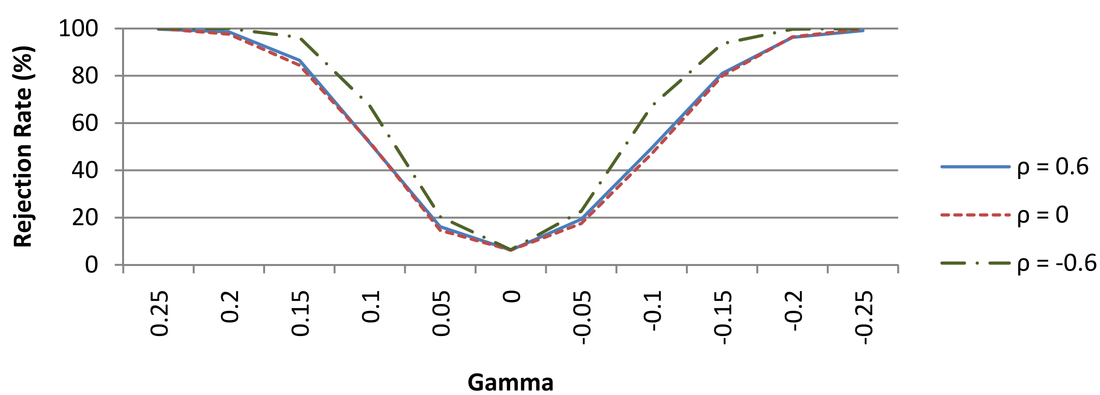


**Figure 4 - Empirical Power Curve of $LM_1$ with ρ = 0.6**

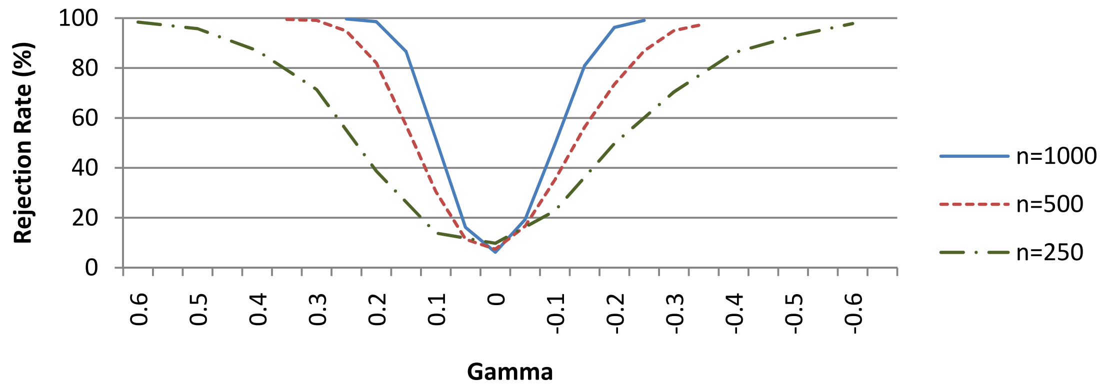



**Figure 5 - Empirical Power Curve of $LM_1$ with ρ = 0**

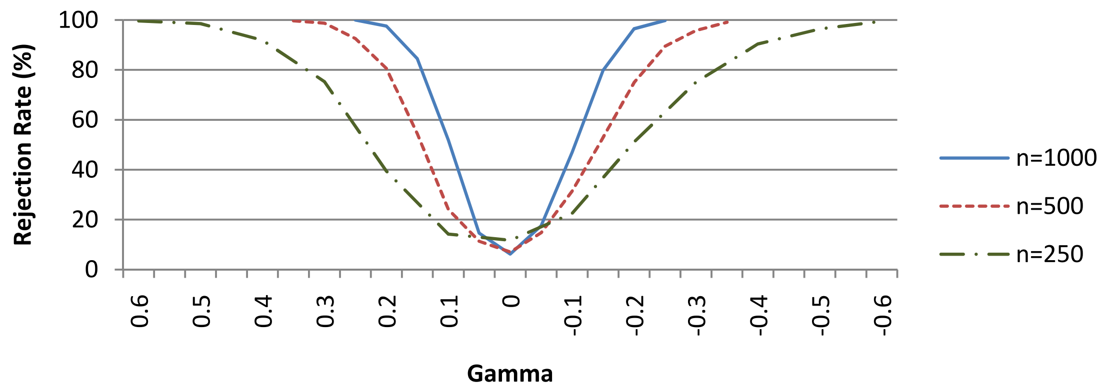
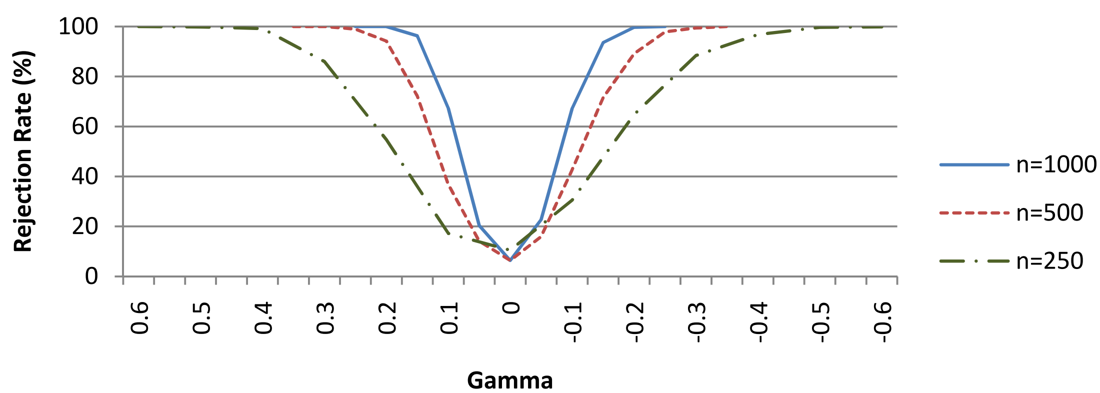



**Figure 6 - Empirical Power Curve of $LM_1$ with ρ = -0.6**



# Chapter 5 An Application

Our two new tests were applied to a German Socioeconomic Panel dataset used in Greene (2008) to estimate a standard bivariate Probit model. This dataset was obtained from Riphahn, Wambach and Million (2003). The data track the number of hospital and doctor visits made by 7,380 men and women aged 25 to 65, totalling 27,326 observations. Greene obtained the latter number of observations by breaking the panel dataset into a cross-sectional dataset. As a result, it is likely that there is some contemporaneous correlation between observations, and this is not modelled in the specification used below. The two-equation model being estimated is as follows:

$doc_i^* = \beta_{10} + \beta_{11}female_i + \beta_{12}age_i + \beta_{13}income_i + \beta_{14}educ_i + \beta_{15}married_i + \beta_{15}kids_i + \varepsilon_{1i}$ ,

$doc_i = 1$, if $doc_i^* > 0$; 0 otherwise,

$hosp_i^* = \beta_{20} + \beta_{21}female_i + \beta_{22}age_i + \beta_{23}income_i + \beta_{24}educ_i + \beta_{25}married_i + \beta_{25}kids_i + \varepsilon_{2i}$ ,

$hosp_i = 1$, if $hosp_i^* > 0$; 0 otherwise,

$$Prob(doc = doc_i, hosp = hosp_i \mid \boldsymbol{x}_1, \boldsymbol{x}_2) = \Phi_2(w_{i1}, w_{i2}, \rho_i^*) .$$

Here, "doc" equals one if an individual has visited the doctor in the past year and equals zero otherwise, "hosp" equals one if that individual has visited the hospital and equals zero otherwise. "Female" and "married" are equal to one if an individual is female or married, respectively. "Educ" refers to the number of years of education, "age" refers to the respondent's age, and "income" is the respondent's nominal monthly household net income. $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ refer to the regressor matrix in the first and second equations, respectively. $w_{i1}, w_{i2}$ and $\rho_i^*$ refer to the dependent and independent variables after the variable transformation outlined on page 11. A bivariate Probit model is estimated, rather than two univariate Probit models, because the error terms of the two equations are likely related: the random process pushing people to visit the doctor is likely related to the random process pushing people to visit the hospital. The use of a bivariate Probit model, rather than two individual Probit models, allows the economist estimating this model to impose parameter restrictions across equations while gaining efficiency of estimation.

The resulting estimated coefficients and standard errors were almost identical to those in Greene (Greene, 2008, p.822), even though he used different software. All of the coefficients were statistically significant except for Income and Kids on the Hospital equation.

**Table 5 - Estimated Bivariate Probit Model**

| | Doctor | | Hospital | |
|---|---|---|---|---|
| Constant | -0.1243 | (0.0581) | -1.3385 | (0.0838) |
| Female | 0.3551 | (0.0160) | 0.1050 | (0.0220) |
| Age | 0.1188 | (0.0008) | 0.0046 | (0.0011) |
| Income | -0.1337 | (0.0456) | 0.0444 | (0.0639) |
| Kids | -0.1524 | (0.0183) | -0.0152 | (0.0256) |
| Education | -0.0148 | (0.0036) | -0.0219 | (0.0522) |
| Married | 0.0735 | (0.0206) | -0.0479 | (0.0279) |
| ρ | 0.2981 | (0.0139) | | |

The $LM_1$ and $F_1$ test statistics for this model were constructed under the assumption that $\mathbf{z}_{ij}$ was equal to $\mathbf{x}_{ij}$—that is, that the heteroskedasticity was a function of the regressors. There are six variables in $\mathbf{z}_{ij}$ for each of the two equations ($j$ =1,2) - hence $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ are both (1x6) vectors. The null hypothesis therefore requires twelve restrictions on the parameters. The $LM_1$ test statistic is asymptotically $\chi^2$ distributed with 12 degrees of freedom under the null; while the $F_1$ test statistic has 12 and (27,326 -15) degrees of freedom asymptotically, under the null. The resulting $LM_1$ and $F_1$ statistics were 693.32

and 355.66, respectively. Regardless of what reasonable significance level one chooses, the null hypothesis must be strongly rejected—there is heteroskedasticity present in the errors of these models. The associated maximum likelihood coefficient estimator based on this model will therefore be inconsistent.

**Table 6 - Reestimated Bivariate Probit Model**

| | Doctor | | Hospital | |
|---|---|---|---|---|
| Constant | 0.4600 | (6.5735) | 0.0891 | (1.6010) |
| Female | 0.1313 | (2.6092) | 0.0161 | (0.5957) |
| Age | 0.0041 | (0.1645) | -0.0004 | (0.0194) |
| Income | -0.0227 | (0.1333) | -0.0962 | (0.0014) |
| Kids | -0.0551 | (2.7298) | -0.0027 | (0.6172) |
| Education | -0.0051 | (0.2592) | -0.0035 | (0.1393) |
| Married | 0.0276 | (4.8852) | -0.0081 | (0.6487) |
| ρ | 0.0351 | (0.0120) | | |

When the model is re-estimated with heteroskedasticity accounted for, the significance of the coefficient estimates changes dramatically. Only ρ and the coefficient on income in the hospital equation remain significant. The magnitude of the coefficients also exhibits a great degree of change. In this particular example, the application of the tests for heteroskedasticity proves to be of vital importance. The test results provide compelling evidence that the parameter estimates are likely to be inconsistent. Even though an extremely large sample is being used, the original Probit estimates are of little interest and should not be used. The re-estimated model indicates that there is a clear problem with the mis-specification, and this requires further study.

# Chapter 6 Areas for Future Research

The most interest area for future research is in attempting to adapt the second of Davidson and MacKinnon's artificial regressions for use in bivariate Probit models. These test statistics are problematic because their derivation requires the interpretation of the log likelihood function as a weighted least squares regression. Specifically, one can view the maximized log likelihood,

$$\frac{\sum_{i=1}^{n}(y_i - F(\boldsymbol{x}_i'\boldsymbol{\beta}))f(\boldsymbol{x}_i'\boldsymbol{\beta})\mathbf{x}_i}{F(\boldsymbol{x}_i'\boldsymbol{\beta})(1-F(\boldsymbol{x}_i'\boldsymbol{\beta}))},$$

as a weighted least squares regression of $F(\boldsymbol{x}_i'\boldsymbol{\beta})$ on $y_i$, with weights equal to $[F(\boldsymbol{x}_i'\boldsymbol{\beta})(1 - F(\boldsymbol{x}_i'\boldsymbol{\beta}))]^{-0.5}$ (Davidson and MacKinnon, 2004, pp.455-456). It can be shown that one can rearrange the maximized log likelihood function in the bivariate case to be

$$\sum_{i=1}^{n}\left[\frac{y_{i1} - \frac{(1-y_{i2})P_{10} + y_{i2}P_{11}}{\omega_{i1}}}{P_{00}P_{01}P_{10}P_{11}}\right][(1-\boldsymbol{y}_{i2})(P_{01}P_{11}) + \boldsymbol{y}_{i2}(P_{10}P_{00})]\omega_{i1}\, f(\boldsymbol{x}_{i1}'\boldsymbol{\beta}_1)\mathbf{x}_{i1} \text{ and}$$

$$\sum_{i=1}^{n}\left[\frac{y_{i1} - \frac{(1-y_{i2})P_{10} + y_{i2}P_{11}}{\omega_{i1}}}{P_{00}P_{01}P_{10}P_{11}}\right][(1-\boldsymbol{y}_{i2})(P_{01}P_{11}) + \boldsymbol{y}_{i2}(P_{10}P_{00})]\omega_{i1}\, f(\boldsymbol{x}_{i1}'\boldsymbol{\beta}_1)(\boldsymbol{x}_{i1}'\boldsymbol{\beta}_1)\boldsymbol{z}_{i1}$$

where $\omega_{i1} = [(1-\boldsymbol{y}_{i2})(P_{10} + P_{00}) + \boldsymbol{y}_{i2}(P_{01} + P_{11})]$,

$$\sum_{i=1}^{n}\left[\frac{y_{i2} - \frac{(1-y_{i1})P_{01} + y_{i1}P_{11}}{\omega_{i2}}}{P_{00}P_{01}P_{10}P_{11}}\right][(1-\boldsymbol{y}_{i1})(P_{10}P_{11}) + \boldsymbol{y}_{i1}(P_{01}P_{00})]\omega_{i2}\, f(\boldsymbol{x}_{i2}'\boldsymbol{\beta}_2)\mathbf{x}_{i2} \text{ and}$$

$$\sum_{i=1}^{n}\left[\frac{y_{i2} - \frac{(1-y_{i1})P_{01} + y_{i1}P_{11}}{\omega_{i2}}}{P_{00}P_{01}P_{10}P_{11}}\right][(1-\boldsymbol{y}_{i1})(P_{10}P_{11}) + \boldsymbol{y}_{i1}(P_{01}P_{00})]\omega_{i2}\, f(\boldsymbol{x}_{i2}'\boldsymbol{\beta}_2)(\boldsymbol{x}_{i2}'\boldsymbol{\beta}_2)\boldsymbol{z}_{i2}$$

where $\omega_{i2} = [(1-\boldsymbol{y}_{i1})(P_{01} + P_{00}) + \boldsymbol{y}_{i1}(P_{10} + P_{11})]$, and finally

$$\sum_{i=1}^{n}\left[\frac{y_{i2} - \frac{(y_{i1}-1)P_{01} + y_{i1}P_{11}}{\omega_{i3}}}{P_{00}P_{01}P_{10}P_{11}}\right][(1-\boldsymbol{y}_{i1})(P_{10}P_{11}) + \boldsymbol{y}_{i1}(P_{01}P_{00})]\omega_{i3}\, q_{i1}q_{i2}\varphi_2$$

where $\omega_{i3} = [(\boldsymbol{y}_{i1} - 1)(P_{01} + P_{00}) + \boldsymbol{y}_{i1}(P_{10} + P_{11})]$, and the $P_{xx}$ values are as described previously.

At first glance one problem becomes apparent—as regressions, this specification would involve both $y_{i1}$ and $y_{i2}$ as dependent variables. Estimating these equations will therefore require the simultaneous estimation of all five likelihood equations. Taking a step back, it is clear in the univariate case that the re-arranged log likelihood function, when interpreted as a weighted least squares regression, provides an estimate of the information matrix. It is not clear, however, that a systems model estimating these five equations will provide an accurate estimate of the information matrix. Exploring the potential for a test statistic based on these regressions is an area for future study.

Another area of research to explore in the future is developing other misspecification tests for bivariate Probit models. Potential misspecification tests include testing for omitted variable bias and testing the assumption of a Probit specification. Finally, these tests could be extended to bivariate Logit models. Although bivariate Logit models are used far less often in the econometrics literature than are bivariate Probit models, it would still be worthwhile to develop a set of tests for practitioners who prefer bivariate Logit over bivariate Probit. A final area for future research is to compare the performance of the LM tests derived here with the computationally more burdensome LRT. This would provide additional insight into the performance of the LM tests, while offering practitioners an alternative test.

# Chapter 7 Conclusion

Two tests for heteroskedasticity in the errors of a bivariate Probit model have been developed in this thesis. Upon analysis through computer simulation, their empirical size and power properties were explored. Both tests were shown to have sizes that quickly approached their nominal asymptotic values as the sample size increased. Similarly, both tests displayed power properties that quickly improved with the sample size. It was simple to apply both of these LM tests to real-world health care data, and the result was decisive. It is clear that these tests will be of great value to empirical economists who wish to estimate bivariate Probit models, and to test their model's specification.

# Bibliography

Davidson, R. & J. G. MacKinnon (1984). Convenient specification tests for logit and probit models. *Journal of Econometrics* 25, 241-262.

Davidson, R. & J. G. MacKinnon (2004). *Econometric Theory and Methods*, Oxford University Press, Oxford.

Fabbri, D., C. Monfardini and R. Radice (2004). Testing exogeneity in the bivariate probit model: Monte Carlo evidence and an application to health economics. Department Working Paper No. 514. Department of Economics, University of Bologna, Italy.

Genz, A., Bretz, F., Miwa, T., Mi. X., Leisch, F., Scheipl, F., *et al*.. (2011). mvtnorm: Multivariate normal and t distributions. *R package version 0.9-96*. Retrieved from http://cran.r-project.org/package=mvtnorm

Greene, W. E. (2008), *Econometric Analysis*, 6th ed., Prentice Hall, Englewood Cliffs, New Jersey.

Murphy, A. (1994). Artificial regression based LM tests of misspecification for discrete choice models. *Economic and Social Review*, 26, 69-74.

Murphy, A. (1996). Simple LM tests for misspecification for ordered logit models. *Economics Letters*, 52, 137-141.

Murphy, Anthony (2007). Score tests of normality in bivariate probit models. *Economics Letters*, 95, 374-379.

Opgen-Rhein, R., Schaefer, J., Strimmer, K. (2010). *corpcor: Efficient estimation of covariance and (partial) correlation*. R package version 1.5.7.

R Development Core Team (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Riphahn, Wambach and Million (2003). Incentive effects in the demand for health care: a bivariate panel count data estimation. *Journal of Applied Econometrics,*18, 387-405.

Smith R. J. (1989). On the use of distributional misspecification checks in limited dependent variable models. *Economic Journal*, 99, 178-192.

Thomas, J. M. (1993). On testing the logistic assumption in binary dependent variable models. *Empirical Economics*, 18, 381-392.

Yatchew, A. and Z. Griliches (1985). Specification error in probit models. *Review of Economics and Statistics*, 67, 134-139.

Yee, TW (2011). *VGAM: vector generalized linear and additive models*. R package version 0.8-2, URL http://CRAN.R-project.org/package=VGAM.

# Appendix A R Code

```
#####################################################
#
#  MC Loop for Bivariate Probit Heteroskedasticity Test
#
#  Notes:
#
#        1) This code tests under the null hypothesis that there is no heteroskedasticity
#        present in this model.  To include heteroskedasticity, change the value of
#        gam1 and gam2
#        to not equal zero
#
#        2) If you wish to load x-data rather than have it randomly generated before the
#        Monte Carlo loop begins, you must change the the appropriate location
#        in the code  (below the line: #import or generate normally distributed x-data)
#
#####################################################

library(mvtnorm)
library(VGAM)
library(corpcor)

#set random seed
seed<-5937293
set.seed(seed)

#set the sample size
num <- 1000
#set the number of repetitions
loops <- 1000

#set the critical values
signifLevel <- 0.01
chicrit1 <- qchisq((1-signifLevel),2)
fcrit1 <- qf((1-signifLevel),2,num-5)
signifLevel <- 0.05
chicrit5 <- qchisq((1-signifLevel),2)
fcrit5 <- qf((1-signifLevel),2,num-5)
signifLevel <- 0.10
chicrit10 <- qchisq((1-signifLevel),2)
fcrit10 <- qf((1-signifLevel),2,num-5)
```

```
#set the true values of the intercept and slope coefficient
#b01,b11 / b02 b12
btrue <- matrix(c(0.25, 0.5, 1, 0.5),ncol=2)

#set the true value of the correlation coefficient
rho <- 0.3

#import or generate normally distributed x-data
xvals <- cbind(rep(1,num),rnorm(num,0))
#xvals <- read.csv("C:/Users/tthorn/xvals.csv",header=TRUE,row.names=1)
#xvals <- cbind(rep(1,num),xvals[1:num,1])
xnum <- ncol(xvals)

#create arrays to store results
lm1Array <- array(data=NA,dim=loops)
f1Array <- array(data=NA,dim=loops)
gradientvec <- array(data=NA,dim=c(num, (2*xnum)+(2*(xnum-1))+1))
y <- array(data=0,dim=c(num,2))
y1 <- array(data=NA,dim=c(num,2))

#######################################################
#
#  Start MC Loop
#
######################################################
for (k in 1:loops){

#generate an empty nx2 array to store the errors
errors <- array(data=NA,dim=c(num,2))

######################################################
#   Generate Errors
######################################################
gam1 <- 0
gam2 <- 0
error1a <- rnorm(num,0)
error2a <- rnorm(num,0)
error1 <- exp(gam1*xvals[,2])*error1a
error2 <- exp(gam2*xvals[,2])*(error1*rho + error2a*((1-(rho^2))^0.5))
errors <- cbind(error1,error2)

#######################################################
#determine y values
y1 <- t(btrue%*%t(xvals)) + errors
```

```
#convert to binary
y <- array(data=0,dim=c(num,2))
y[(y1>=0)] = 1

#set up q1 and q2
q = 2*y - 1

#use the VGAM package to fit the bivariate Probit model
#to the randomly generated data
rhostart = log((rho+1)/(1-rho))
fit <- vglm(y~xvals[,2], binom2.rho,lrho="rhobit",init.rho=rho,epsilon=1e-
10,coefstart=c(btrue[1,1],btrue[2,1],rhostart,btrue[1,2],btrue[2,2]),maxit=5000,
na.action=na.pass)

#form yhats
xbhat = predict(fit,type="terms", matrix=TRUE)
yhat  = cbind(xbhat[,1]+attr(xbhat,"constant")[1],xbhat[,2]+attr(xbhat,"constant")[2])

#obtain the estimated value for rho, the correlation coefficient
rhobit <- coef(fit,matrix=TRUE)[1,3]
rhohat <- (exp(rhobit)-1)/(1+exp(rhobit))

#set up w and rhostar
w <- q*yhat
rhostar <- q[,1]*q[,2]*rhohat

#Start loop to create gradient vector
for(i in 1:num){

sigmahat <- matrix(c(1,rhostar[i],rhostar[i],1),ncol=2)
cdfmvn <- pmvnorm(lower=c(-Inf,-Inf), upper=c(w[i,1],w[i,2]),sigma=sigmahat)
pdfmvn <- dmvnorm(c(w[i,1],w[i,2]),sigma=sigmahat)

######
# dlogl/db1  and  dlogl/dgam1
######

insidevar <- (w[i,2] - rhostar[i]*w[i,1])/((1-rhostar[i]^2)^0.5)
dldw1 <- (dnorm(w[i,1])* pnorm(insidevar)/cdfmvn)

bgrads <- q[i,1]*dldw1*xvals[i,]
gamgrads <- -q[i,1]*dldw1*yhat[i,1]*xvals[i,]

gradientvec[i,1] <- bgrads[1]
gradientvec[i,2] <- bgrads[2]
```

```
gradientvec[i,6] <- gamgrads[2]

######
# dlogl/db2  and  dlogl/dgam2
######

insidevar <- (w[i,1] - rhostar[i]*w[i,2])/((1-rhostar[i]^2)^0.5)
dldw2 <- (dnorm(w[i,2])*pnorm(insidevar)/cdfmvn)

bgrads <- q[i,2]*dldw2*xvals[i,]
gamgrads <- -q[i,2]*dldw2*yhat[i,2]*xvals[i,]

gradientvec[i,3] <- bgrads[1]
gradientvec[i,4] <- bgrads[2]
gradientvec[i,7] <- gamgrads[2]

######
# dlogl/drho
######

pgrad <- q[i,1]*q[i,2]*pdfmvn/cdfmvn
gradientvec[i,5] <- pgrad
}
#End loop that creates gradient

#######################################################
#  Create the LM1 test statistic
#######################################################
ones <- matrix(1,c(num,1))
lm1Array[k] <-
t(ones)%*%gradientvec%*%pseudoinverse(t(gradientvec)%*%gradientvec,tol=1e-
30)%*%t(gradientvec)%*%ones

#######################################################
#  Create the F1 F-statistic.
#######################################################
ssr <- t(ones)%*%ones -
t(ones)%*%gradientvec%*%pseudoinverse(t(gradientvec)%*%gradientvec,tol=1e-
30)%*%t(gradientvec)%*%ones
f1Array[k] <- ((num-ssr)/2)/(ssr/(num-2))




}
```

```
##################################################
#END MC LOOP
##################################################

#calculate the number of rejections at 1,5 and 10%.  Ignore NA values
f1rejections1 <- sum(sort(f1Array)>=fcrit1)/length(sort(f1Array))
lm1rejections1 <- sum(sort(lm1Array)>=chicrit1)/length(sort(lm1Array))
f1rejections5 <- sum(sort(f1Array)>=fcrit5)/length(sort(f1Array))
lm1rejections5 <- sum(sort(lm1Array)>=chicrit5)/length(sort(lm1Array))
f1rejections10 <- sum(sort(f1Array)>=fcrit10)/length(sort(f1Array))
lm1rejections10 <- sum(sort(lm1Array)>=chicrit10)/length(sort(lm1Array))

#add up the number of NA values
f1NAs <- sum(is.na(f1Array))
lm1NAs <- sum(is.na(lm1Array))

f1Array
lm1Array

f1NAs
lm1NAs

f1rejections1
lm1rejections1

f1rejections5
lm1rejections5

f1rejections10
lm1rejections10
```

# Appendix B Additional Empirical Size Tables

**Table 7 - Empirical Size (%) at α = 1%, 1000 Repetitions**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| n | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 250 | 2.7 | 2.7 | 4.7 | 4.8 | 4.4 | 4.4 | 4.1 | 4.3 | 3.5 | 3.7 |
| 500 | 1.9 | 2.0 | 1.5 | 1.5 | 1.6 | 1.6 | 2.0 | 2.0 | 1.6 | 1.6 |
| 1000 | 1.9 | 1.9 | 1.1 | 1.2 | 0.9 | 1.0 | 1.3 | 1.3 | 1.6 | 1.6 |
| 2500 | 1.5 | 1.4 | 1.1 | 1.1 | 1.3 | 1.3 | 1.2 | 1.2 | 1.0 | 1.0 |
| 5000 | 1.1 | 1.1 | 1.3 | 1.3 | 1.0 | 1.0 | 1.5 | 1.5 | 1.1 | 1.1 |
| 7500 | 1.2 | 1.2 | 1.0 | 1.0 | 1.6 | 1.6 | 0.9 | 0.9 | 1.6 | 1.6 |

**Table 8 - Empirical Size (%) at α = 1%, 5000 Repetitions**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| n | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 250 | 3.32 | 3.44 | 3.76 | 3.84 | 3.54 | 3.66 | 3.06 | 3.14 | 3.30 | 3.40 |
| 1000 | 1.50 | 1.48 | 1.38 | 1.44 | 1.14 | 1.18 | 1.46 | 1.48 | 1.44 | 1.44 |
| 2500 | 1.24 | 1.22 | 1.44 | 1.44 | 1.28 | 1.28 | 1.16 | 1.14 | 0.96 | 0.96 |

# Appendix C Empirical Power Tables

**Table 9 - Percent Rejection Rates (%) at α = 1%, Sample Size 1000**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.25 | 99.5 | 99.5 | 98.9 | 98.9 | 99.3 | 99.3 | 99.6 | 99.6 | 100.0 | 100.0 |
| 0.20 | 93.7 | 93.7 | 92.6 | 92.6 | 91.3 | 91.4 | 96.3 | 96.3 | 99.5 | 99.5 |
| 0.15 | 67.5 | 67.5 | 65.0 | 65.0 | 64.1 | 64.2 | 74.3 | 74.7 | 87.5 | 87.6 |
| 0.10 | 26.8 | 26.8 | 23.3 | 23.4 | 26.1 | 26.2 | 30.1 | 30.1 | 41.1 | 41.1 |
| 0.05 | 4.5 | 4.5 | 5.5 | 5.5 | 3.7 | 3.8 | 6.2 | 6.2 | 7.3 | 7.2 |
| 0 | 1.9 | 1.9 | 1.1 | 1.2 | 0.9 | 1.0 | 1.3 | 1.3 | 1.6 | 1.6 |
| -0.05 | 8.3 | 8.1 | 6.0 | 6.0 | 6.0 | 6.0 | 6.6 | 6.6 | 9.0 | 9.1 |
| -0.10 | 27.8 | 27.9 | 24.3 | 24.4 | 25.2 | 25.3 | 33.9 | 34.1 | 43.1 | 43.1 |
| -0.15 | 59.5 | 59.7 | 57.5 | 57.8 | 58.3 | 58.4 | 68.2 | 68.3 | 82.8 | 82.9 |
| -0.20 | 86.7 | 86.8 | 84.9 | 85.0 | 87.3 | 87.6 | 92.9 | 92.9 | 98.0 | 98.0 |
| -0.25 | 96.7 | 96.7 | 95.7 | 95.8 | 97.6 | 97.6 | 99.1 | 99.1 | 99.9 | 99.8 |

**Table 10 - Percent Rejection Rates (%) at α = 5%, Sample Size 1000**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.25 | 99.7 | 99.5 | 99.8 | 99.8 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 0.20 | 98.6 | 93.7 | 97.9 | 97.9 | 97.6 | 97.6 | 99.4 | 99.4 | 100.0 | 100.0 |
| 0.15 | 86.6 | 67.5 | 84.2 | 84.3 | 84.5 | 84.6 | 90.3 | 90.3 | 96.3 | 96.2 |
| 0.10 | 51.6 | 26.8 | 48.7 | 48.7 | 51.9 | 51.9 | 54.2 | 54.1 | 67.3 | 67.3 |
| 0.05 | 16.1 | 4.5 | 18.0 | 18.0 | 14.6 | 14.6 | 18.3 | 18.3 | 20.4 | 20.3 |
| 0 | 6.2 | 1.9* | 6.3 | 6.3 | 6.2 | 6.2 | 5.6 | 5.6 | 6.4 | 6.4 |
| -0.05 | 19.3 | 8.1 | 19.1 | 19.1 | 17.5 | 17.5 | 19.5 | 19.6 | 22.7 | 22.7 |
| -0.10 | 49.4 | 27.9 | 46.8 | 46.8 | 47.0 | 47.0 | 56.4 | 56.4 | 67.2 | 67.1 |
| -0.15 | 81.0 | 59.7 | 78.4 | 78.4 | 79.9 | 79.9 | 86.9 | 86.9 | 93.6 | 93.6 |
| -0.20 | 96.3 | 86.8 | 94.7 | 94.7 | 96.5 | 96.5 | 98.2 | 98.2 | 99.7 | 99.7 |
| -0.25 | 99.1 | 96.7 | 99.1 | 99.1 | 99.8 | 99.8 | 99.8 | 99.8 | 100.0 | 99.9 |

***this number has been double checked**

**Table 11 - Percent Rejection Rates (%) at α = 10%, Sample Size 1000**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.25 | 99.9 | 99.9 | 99.9 | 99.9 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| 0.20 | 99.6 | 99.6 | 99.1 | 99.1 | 99.0 | 99.0 | 99.6 | 99.6 | 100.0 | 100.0 |
| 0.15 | 92.3 | 92.3 | 90.5 | 90.5 | 91.3 | 91.4 | 94.5 | 94.5 | 98.4 | 98.3 |
| 0.10 | 64.7 | 64.8 | 61.8 | 61.8 | 62.4 | 62.4 | 68.2 | 68.1 | 78.4 | 78.4 |
| 0.05 | 25.0 | 25.0 | 27.7 | 27.7 | 23.9 | 23.9 | 26.7 | 26.8 | 31.9 | 31.8 |
| 0 | 10.8 | 10.8 | 11.9 | 11.9 | 10.8 | 10.8 | 12.1 | 12.1 | 12.2 | 12.2 |
| -0.05 | 30.7 | 30.5 | 29.4 | 29.4 | 27.9 | 27.9 | 29.7 | 29.7 | 33.3 | 33.3 |
| -0.10 | 60.5 | 60.5 | 60.3 | 60.3 | 60.1 | 60.1 | 67.4 | 67.4 | 77.4 | 77.3 |
| -0.15 | 88.6 | 88.6 | 85.6 | 85.6 | 88.9 | 88.9 | 92.2 | 92.2 | 96.6 | 96.6 |
| -0.20 | 98.6 | 98.6 | 97.0 | 97.0 | 98.6 | 98.6 | 99.6 | 99.6 | 99.9 | 99.9 |
| -0.25 | 99.6 | 99.6 | 99.7 | 99.7 | 99.9 | 99.9 | 100.0 | 100.0 | 100.0 | 99.9 |

**Table 12 - Percent Rejection Rates (%) at α = 1%, Sample Size 500**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.35 | 97.6 | 97.6 | 97.1 | 97.1 | 98.0 | 98.1 | 99.5 | 99.5 | 100.0 | 100.0 |
| 0.30 | 93.6 | 93.6 | 93.0 | 93.1 | 93.6 | 93.7 | 96.6 | 96.6 | 99.5 | 99.5 |
| 0.25 | 80.9 | 81.2 | 74.9 | 75.2 | 78.2 | 78.6 | 85.2 | 85.3 | 95.6 | 95.6 |
| 0.20 | 58.6 | 58.8 | 49.2 | 49.7 | 56.2 | 56.7 | 62.5 | 63.1 | 78.5 | 79.3 |
| 0.15 | 27.9 | 28.0 | 25.5 | 25.8 | 27.6 | 27.6 | 33.3 | 33.8 | 47.9 | 48.0 |
| 0.10 | 13.0 | 13.3 | 10.4 | 10.5 | 9.6 | 9.7 | 12.3 | 12.5 | 15.9 | 16.1 |
| 0.05 | 3.1 | 3.2 | 3.0 | 3.1 | 3.0 | 3.0 | 2.6 | 2.6 | 3.0 | 3.0 |
| 0 | 1.9 | 2.0 | 1.5 | 1.5 | 1.6 | 1.6 | 2.0 | 2.0 | 1.6 | 1.6 |
| -0.05 | 5.7 | 5.9 | 6.1 | 6.3 | 4.4 | 4.4 | 5.8 | 5.9 | 5.5 | 5.5 |
| -0.10 | 15.7 | 16.0 | 14.8 | 15.0 | 14.7 | 14.9 | 16.0 | 16.2 | 22.0 | 22.4 |
| -0.15 | 32.7 | 32.8 | 27.3 | 27.7 | 31.3 | 31.5 | 35.7 | 36.2 | 49.0 | 49.4 |
| -0.20 | 48.5 | 48.8 | 49.1 | 49.3 | 52.2 | 52.8 | 59.6 | 59.9 | 73.4 | 73.4 |
| -0.25 | 70.8 | 81.0 | 70.5 | 70.8 | 72.9 | 73.0 | 81.5 | 81.7 | 90.9 | 90.9 |
| -0.30 | 84.5 | 84.9 | 84.6 | 84.6 | 87.8 | 88.0 | 92.8 | 93.0 | 97.6 | 97.6 |
| -0.35 | 91.6 | 91.6 | 91.4 | 91.5 | 95.7 | 95.7 | 97.7 | 97.7 | 100.0 | 100.0 |

**Table 13 - Percent Rejection Rates (%) at α = 5%, Sample Size 500**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.35 | 99.5 | 99.5 | 96.6 | 99.6 | 99.7 | 99.7 | 100.0 | 100.0 | 100.0 | 100.0 |
| 0.30 | 99.1 | 99.1 | 99.2 | 99.2 | 98.7 | 98.7 | 99.6 | 99.6 | 100.0 | 100.0 |
| 0.25 | 94.8 | 94.8 | 92.1 | 92.2 | 92.6 | 92.6 | 96.7 | 96.7 | 99.0 | 99.0 |
| 0.20 | 82.1 | 82.1 | 76.5 | 76.5 | 80.6 | 80.6 | 84.9 | 84.9 | 94.2 | 94.3 |
| 0.15 | 57.0 | 57.0 | 49.8 | 49.8 | 54.5 | 54.7 | 60.1 | 60.1 | 72.2 | 72.3 |
| 0.10 | 30.5 | 30.6 | 26.2 | 26.3 | 24.1 | 24.2 | 30.3 | 30.2 | 36.8 | 36.9 |
| 0.05 | 11.4 | 11.5 | 9.9 | 9.8 | 11.3 | 11.3 | 11.6 | 11.6 | 14.1 | 14.1 |
| 0 | 7.4 | 7.5 | 7.2 | 7.3 | 7.0 | 7.0 | 7.9 | 8.1 | 6.4 | 6.4 |
| -0.05 | 16.6 | 16.7 | 16.0 | 16.1 | 14.7 | 14.8 | 16.4 | 16.5 | 15.9 | 15.8 |
| -0.10 | 35.2 | 35.2 | 30.7 | 30.7 | 31.4 | 31.4 | 34.0 | 34.1 | 42.7 | 42.8 |
| -0.15 | 56.3 | 56.3 | 50.2 | 50.1 | 53.0 | 53.0 | 59.5 | 59.6 | 71.6 | 71.7 |
| -0.20 | 73.5 | 73.5 | 73.4 | 73.5 | 74.9 | 75.0 | 79.7 | 79.8 | 89.1 | 89.1 |
| -0.25 | 86.9 | 86.9 | 87.3 | 87.3 | 89.5 | 89.5 | 95.1 | 95.2 | 97.9 | 97.9 |
| -0.30 | 95.0 | 94.9 | 95.9 | 95.9 | 95.8 | 95.9 | 98.2 | 98.2 | 99.4 | 99.4 |
| -0.35 | 97.6 | 97.6 | 97.7 | 97.7 | 99.1 | 99.1 | 99.7 | 99.7 | 100.0 | 100.0 |

**Table 14 - Percent Rejection Rates (%) at α = 10%, Sample Size 500**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.35 | 99.7 | 99.7 | 100.0 | 100.0 | 99.9 | 99.9 | 100.0 | 100.0 | 100.0 | 100.0 |
| 0.30 | 99.5 | 99.5 | 99.4 | 99.4 | 99.5 | 99.5 | 99.8 | 99.8 | 100.0 | 100.0 |
| 0.25 | 98.0 | 98.0 | 95.6 | 95.6 | 96.0 | 96.0 | 98.6 | 98.6 | 99.7 | 99.7 |
| 0.20 | 89.3 | 89.3 | 85.8 | 85.8 | 87.9 | 88.0 | 92.0 | 92.0 | 97.1 | 97.1 |
| 0.15 | 69.1 | 69.3 | 63.3 | 63.4 | 67.9 | 68.0 | 74.0 | 74.1 | 84.2 | 84.2 |
| 0.10 | 43.6 | 43.7 | 37.8 | 37.8 | 35.2 | 35.2 | 44.4 | 44.3 | 52.2 | 52.3 |
| 0.05 | 18.4 | 18.4 | 18.1 | 18.1 | 18.0 | 18.0 | 21.0 | 21.0 | 23.2 | 23.2 |
| 0 | 14.9 | 14.9 | 13.1 | 13.2 | 13.4 | 13.4 | 14.4 | 14.4 | 11.9 | 11.9 |
| -0.05 | 24.8 | 24.8 | 26.2 | 26.2 | 23.8 | 23.9 | 24.6 | 24.6 | 24.1 | 24.0 |
| -0.10 | 48.2 | 48.2 | 40.9 | 41.0 | 42.2 | 42.2 | 47.1 | 47.1 | 54.0 | 54.0 |
| -0.15 | 67.0 | 67.0 | 63.0 | 62.9 | 64.7 | 64.7 | 70.6 | 70.7 | 82.1 | 82.1 |
| -0.20 | 81.9 | 81.9 | 82.5 | 82.5 | 84.7 | 84.7 | 87.9 | 87.9 | 93.8 | 93.7 |
| -0.25 | 93.3 | 93.3 | 92.9 | 92.9 | 94.1 | 94.1 | 97.4 | 97.4 | 98.9 | 98.9 |
| -0.30 | 97.6 | 97.5 | 98.0 | 98.0 | 97.7 | 97.7 | 99.0 | 99.0 | 99.9 | 99.9 |
| -0.35 | 98.8 | 98.8 | 99.2 | 99.2 | 99.7 | 99.7 | 99.8 | 99.8 | 100.0 | 100.0 |

**Table 15 - Percent Rejection Rates (%) at α = 1%, Sample Size 250**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.6 | 92.3 | 92.7 | 95.1 | 95.4 | 97.3 | 97.5 | 99.9 | 99.9 | 100.0 | 100.0 |
| 0.5 | 82.2 | 83.2 | 84.6 | 85.0 | 91.5 | 91.8 | 97.5 | 97.6 | 99.1 | 99.1 |
| 0.4 | 63.9 | 65.1 | 66.7 | 67.5 | 74.1 | 74.5 | 87.1 | 87.2 | 94.0 | 94.3 |
| 0.3 | 43.3 | 44.1 | 39.7 | 40.2 | 48.5 | 49.1 | 53.4 | 54.5 | 67.8 | 68.4 |
| 0.2 | 15.7 | 16.4 | 16.9 | 17.4 | 17.0 | 17.3 | 20.3 | 20.6 | 25.1 | 25.8 |
| 0.1 | 3.4 | 3.9 | 3.5 | 3.5 | 4.9 | 5.1 | 4.1 | 4.3 | 5.4 | 5.5 |
| 0 | 2.7 | 2.7 | 4.7 | 4.8 | 4.4 | 4.4 | 3.1 | 3.2 | 3.5 | 3.7 |
| -0.1 | 9.4 | 9.8 | 9.2 | 9.4 | 10.8 | 11.0 | 11.4 | 11.8 | 14.7 | 14.8 |
| -0.2 | 30.3 | 30.5 | 28.2 | 28.8 | 30.4 | 30.6 | 35.8 | 36.0 | 46.0 | 46.4 |
| -0.3 | 52.0 | 52.4 | 52.4 | 52.8 | 57.5 | 58.3 | 64.7 | 65.0 | 76.0 | 76.2 |
| -0.4 | 72.3 | 72.7 | 73.1 | 73.3 | 76.4 | 76.7 | 84.7 | 85.2 | 92.5 | 92.7 |
| -0.5 | 83.9 | 83.9 | 86.1 | 86.4 | 89.1 | 89.4 | 94.7 | 94.9 | 98.9 | 99.0 |
| -0.6 | 92.4 | 92.7 | 93.4 | 93.6 | 96.2 | 96.3 | 99.1 | 99.1 | 99.5 | 99.6 |

**Table 16 - Percent Rejection Rates (%) at α = 5%, Sample Size 250**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.6 | 98.4 | 98.4 | 99.0 | 99.0 | 99.6 | 99.6 | 100.0 | 100.0 | 100.0 | 100.0 |
| 0.5 | 95.8 | 95.9 | 96.0 | 96.0 | 98.5 | 98.5 | 99.8 | 99.8 | 99.8 | 99.8 |
| 0.4 | 87.0 | 87.1 | 88.4 | 88.5 | 91.9 | 91.9 | 97.0 | 97.1 | 99.2 | 99.2 |
| 0.3 | 71.4 | 71.8 | 68.9 | 69.4 | 75.3 | 75.4 | 79.7 | 80.1 | 86.0 | 86.2 |
| 0.2 | 38.7 | 38.8 | 37.2 | 37.6 | 39.4 | 39.5 | 43.4 | 43.7 | 54.7 | 54.8 |
| 0.1 | 13.9 | 14.0 | 14.1 | 14.3 | 14.2 | 14.3 | 13.4 | 13.6 | 17.2 | 17.5 |
| 0 | 9.8 | 9.9 | 11.5 | 11.6 | 11.7 | 11.8 | 10.5 | 10.5 | 10.5 | 10.6 |
| -0.1 | 22.8 | 23.0 | 23.8 | 24.0 | 22.5 | 22.7 | 24.2 | 24.7 | 30.5 | 30.5 |
| -0.2 | 49.7 | 49.7 | 47.1 | 47.1 | 51.1 | 51.2 | 56.6 | 56.9 | 64.8 | 65.1 |
| -0.3 | 70.4 | 70.4 | 69.9 | 70.1 | 75.1 | 75.3 | 80.0 | 80.1 | 88.4 | 88.4 |
| -0.4 | 86.1 | 86.3 | 86.4 | 86.5 | 90.4 | 90.4 | 94.2 | 94.3 | 96.8 | 96.8 |
| -0.5 | 92.8 | 92.7 | 94.5 | 94.6 | 96.4 | 96.5 | 98.7 | 98.7 | 99.7 | 99.7 |
| -0.6 | 97.8 | 97.8 | 98.3 | 98.4 | 99.5 | 99.5 | 99.8 | 99.8 | 99.9 | 99.9 |

**Table 17 - Percent Rejection Rates (%) at α = 10%, Sample Size 250**

| | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|
| γ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.6 | 99.4 | 99.4 | 99.7 | 99.7 | 99.8 | 99.8 | 100.0 | 100.0 | 100.0 | 100.0 |
| 0.5 | 98.7 | 98.7 | 98.9 | 98.9 | 99.6 | 99.6 | 100.0 | 100.0 | 99.9 | 99.9 |
| 0.4 | 92.6 | 92.6 | 94.5 | 94.5 | 96.6 | 96.6 | 98.3 | 98.3 | 99.7 | 99.7 |
| 0.3 | 82.4 | 82.5 | 79.6 | 79.6 | 84.6 | 84.6 | 87.8 | 87.9 | 93.0 | 93.0 |
| 0.2 | 52.5 | 52.6 | 50.2 | 50.3 | 54.7 | 54.7 | 57.6 | 57.7 | 65.8 | 65.8 |
| 0.1 | 22.2 | 22.2 | 21.9 | 21.9 | 23.3 | 23.3 | 23.8 | 23.9 | 27.6 | 27.8 |
| 0 | 17.4 | 17.4 | 18.6 | 18.8 | 17.6 | 17.7 | 17.9 | 17.9 | 18.8 | 18.8 |
| -0.1 | 32.0 | 32.0 | 34.1 | 34.1 | 31.3 | 31.4 | 33.7 | 33.7 | 38.6 | 38.5 |
| -0.2 | 61.4 | 61.3 | 57.7 | 57.7 | 60.5 | 60.5 | 67.2 | 67.3 | 74.1 | 74.1 |
| -0.3 | 78.4 | 78.4 | 79.1 | 79.2 | 82.7 | 82.7 | 86.7 | 86.7 | 93.1 | 93.1 |
| -0.4 | 91.5 | 91.5 | 91.7 | 91.7 | 94.7 | 94.7 | 96.2 | 96.2 | 98.5 | 98.5 |
| -0.5 | 96.3 | 96.2 | 97.4 | 97.4 | 98.0 | 98.0 | 99.3 | 99.3 | 100.0 | 100.0 |
| -0.6 | 99.1 | 99.4 | 99.0 | 99.0 | 99.9 | 99.9 | 100.0 | 100.0 | 100.0 | 100.0 |

**Table 18 - Mixed γ Percent Rejection Rates (%) at α = 5%, Sample Size 500**

| | | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| γ1 | γ2 | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.1 | -0.1 | 28.3 | 28.4 | 27.0 | 27.1 | 31.4 | 31.5 | 24.7 | 24.7 | 21.6 | 21.6 |
| 0.2 | 0.1 | 65.1 | 65.3 | 63.9 | 64.0 | 62.7 | 62.7 | 66.3 | 66.4 | 80.1 | 80.1 |
| -0.2 | 0.1 | 59.7 | 59.8 | 59.1 | 59.3 | 61.7 | 61.7 | 51.6 | 51.6 | 52.0 | 52.1 |
| -0.2 | 0.2 | 74.7 | 74.7 | 76.8 | 76.9 | 76.7 | 76.8 | 66.5 | 66.6 | 57.7 | 57.7 |
| -0.2 | -0.1 | 33.8 | 33.9 | 59.4 | 59.4 | 60.4 | 60.4 | 68.6 | 68.7 | 78.1 | 78.3 |
| -0.3 | 0.1 | 65.1 | 65.3 | 87.2 | 87.1 | 84.4 | 84.4 | 81.9 | 82.0 | 82.7 | 82.8 |
| -0.3 | 0.2 | 80.7 | 80.7 | 92.9 | 92.9 | 92.3 | 92.3 | 86.6 | 86.7 | 82.1 | 82.1 |
| -0.3 | -0.1 | 64.5 | 64.5 | 85.1 | 85.1 | 85.5 | 85.6 | 90.1 | 90.1 | 96.3 | 96.3 |
| -0.3 | -0.2 | 89.9 | 89.9 | 89.3 | 89.3 | 90.5 | 90.6 | 96.2 | 96.3 | 99.0 | 99.0 |

**Table 19 - Mixed γ Percent Rejection Rates (%) at α = 5%, Sample Size 1000**

| | | ρ = 0.6 | | ρ = 0.3 | | ρ = 0 | | ρ = -0.3 | | ρ = -0.6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| γ1 | γ2 | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ | $LM_1$ | $F_1$ |
| 0.1 | -0.1 | 49.6 | 49.6 | 51.7 | 51.8 | 51.0 | 51.0 | 43.4 | 43.4 | 33.7 | 33.8 |
| 0.2 | 0.1 | 92.3 | 92.4 | 90.5 | 90.6 | 90.3 | 90.3 | 93.8 | 93.9 | 98.7 | 98.7 |
| -0.2 | 0.1 | 87.3 | 87.3 | 88.5 | 88.5 | 85.5 | 85.5 | 82.6 | 82.6 | 80.2 | 80.3 |
| -0.2 | 0.2 | 96.8 | 96.8 | 97.9 | 97.9 | 97.4 | 97.4 | 93.2 | 93.3 | 85.9 | 85.9 |
| -0.2 | -0.1 | 89.2 | 89.2 | 86.5 | 86.6 | 85.7 | 85.7 | 90.2 | 90.2 | 97.6 | 97.6 |