Tests of Functional Form and Heteroscedasticity

Z. L. Yang and Y. K. Tse

School of Economics and Social Sciences
Singapore Management University
469 Bukit Timah Road
Singapore 259756
Email addresses: zlyang@smu.edu.sg, yktse@smu.ed.sg

November 2003

Abstract: This paper considers tests of misspecification in a heteroscedastic transformation model. We derive Lagrange multiplier (LM) statistics for (i) testing functional form and heteroscedasticity jointly, (ii) testing functional form in the presence of heteroscedasticity, and (iii) testing heteroscedasticity in the presence of data transformation. We present LM statistics based on the expected information matrix. For cases (i) and (ii), this is done assuming the Box-Cox transformation. For case (iii), the test does not depend on whether the functional form is estimated or pre-specified. Small-sample properties of the tests are assessed by Monte Carlo simulation, and comparisons are made with the likelihood ratio test and other versions of LM test. The results show that the expected-information based LM test has the most appropriate finite-sample empirical size.

Key Words: Functional Form, Heteroscedasticity, Lagrange Multiplier Test.

JEL Classification: C1, C5

1 Introduction

Functional form specification and heteroscedasticity estimation are two problems often faced by econometricians. There is a large literature on estimating and testing heteroscedasticity,[1] as well as a sizable literature on estimating and testing functional form.[2] In this paper we address these two issues together. Although a single transformation on the response may result in a correct functional relationship between the response and the explanatory variables, the residuals may be heteroscedastic. When the issue of functional form is of primary concern, a serious consequence of assuming homoscedastic errors is that the estimate of the response transformation is often biased toward the direction of stabilizing the error variance (Zarembka, 1974).

Lahiri and Egy (1981) suggested that the estimation of functional form should be separated from the issue of stabilizing the error variance and that the tests for functional form and heteroscedasticity should be jointly considered. They proposed a test based on the likelihood ratio (LR) statistic with the Box-Cox transformation and a simple heteroscedasticity structure. Seaks and Layson (1983) derived a procedure for the joint estimation of functional form and heteroscedasticity. Tse (1984) proposed an LM test for jointly testing log-linear versus linear transformation with homoscedastic errors.

In this paper we consider the problem of testing functional form and heteroscedasticity, both separately and jointly, in a heteroscedastic transformation model. We derive Lagrange multiplier (LM) statistics for (i) testing jointly functional form and heteroscedasticity, (ii) testing functional form in the presence of heteroscedasticity, and (iii) testing heteroscedasticity in the presence of data transformation. These tests are either generalizations or alternatives of the tests available in the literature, such as the tests by Lahiri and Egy (1981) and Tse (1984) for case (i), Lawrance (1987) for case (ii) and Breusch and Pagan (1979) for case (iii). The LM statistics are derived based on the

[1]See, for example, Goldfeld and Quant (1965), Glejser (1969), Harvey (1976), Amemiya (1977), Breusch and Pagan (1979), Koenker (1981), Ali and Giaccotto (1984), Griffiths and Surekha (1986), Farebrother (1987), Maekawa (1987), Evans and King (1988), Kalirajan (1989), Evans (1992), and Wallentin and Agren (2002).

[2]See, for example, Box and Cox (1964), Godfrey and Wickens (1981), Davidson and MacKinnon (1985), Lawrance (1987), Baltagi (1997), and Yang and Abeysinghe (2003).

expected information matrix. As argued by Bera and McKenzie (1986), this version of the LM statistic is likely to have the best small-sample performance.

For cases (i) and (ii), formulae for the computation of the expected-information based LM statistics are derived assuming the Box-Cox transformation. For case (iii), the asymptotic distribution of the LM statistic does not depend on whether the functional-form parameter is estimated or pre-specified. Indeed, the result applies to any smooth monotonic functional-form specification. Furthermore, when the null hypothesis corresponds to a homoscedastic model under the alternatives specified by Breusch and Pagan (1979), the test does not depend on the postulated alternative heteroscedasticity structure. Monte Carlo simulations are performed to examine the small-sample performance of the tests against other versions of LM statistics, as well as the LR statistic. The results show that the expected-information based LM test has the most appropriate empirical finite-sample size. It outperforms the LR test and other versions of LM test, including the one based on the double-length regression .

The proposed expected-information based LM statistic has closed-form expressions and is computationally much simpler than the LR statistic. Furthermore, it applies to null hypotheses with an arbitrary transformation parameter. While the existing literature focuses on specific transformation parameters corresponding to linear versus log-linear regression (Godfrey and Wickens, 1981, and Godfrey, McAleer and McKenzie, 1988, among others) and simple difference versus percentage change (Layson and Seaks, 1984, and Baltagi and Li, 2000, among others), our results extend to any hypothesized value within the given range. Also, we accommodate the null hypothesis of heteroscedastic errors, which extends beyond testing only for homoscedasticity.

The rest of the paper is organized as follows. Section 2 describes the model and the maximum likelihood estimation of the model parameters. Section 3 develops the LM statistics. Section 4 evaluates the small-sample properties of the LM tests by means of Monte Carlo simulation. Section 5 concludes.

2 The Model

Let $h(.)$ be a monotonic increasing transformation dependent on a parameter vector λ with p elements. Suppose that the transformed dependent observation $h(y_i, \lambda)$ follows a linear regression model with heteroscedastic normal errors given by:

$$h(y_i, \lambda) = h'(x_{1i}, \lambda)\beta_1 + x'_{2i}\beta_2 + \sigma\,\omega(v_i, \gamma)\,e_i, \qquad i = 1, \cdots, n, \tag{1}$$

where β_1 and β_2 are $k_1 \times 1$ and $k_2 \times 1$ vectors of regression coefficients, x_{1i} and x_{2i} are $k_1 \times 1$ and $k_2 \times 1$ vectors of independent variables, $\omega(v_i, \gamma) \equiv \omega_i(\gamma)$ is the weight function, v_i is a set of q weighting variables, γ is a $q \times 1$ vector of weighting parameters, and σ is a constant. Note that among the regressors the functional transformation $h(\cdot)$ is applied to x_{1i} but not x_{2i}.[3]

It is convenient to re-write equation (1) as

$$h(y_i, \lambda) = x'_i(\lambda)\beta + \sigma\,\omega(v_i, \gamma)\,e_i, \qquad i = 1, \cdots, n, \tag{2}$$

where $x_i(\lambda) = (h'(x_{1i}, \lambda), x'_{2i})'$, and $\beta = (\beta'_1, \beta'_2)'$ has $k = k_1 + k_2$ elements. The weighting variables v_i may include some regressors in x_{1i} and x_{2i}. We assume $\omega(v_i, 0) = 1$ so that $\gamma = 0$ represents a model with homoscedastic errors.

Let $\psi = \{\beta', \sigma^2, \gamma', \lambda'\}'$ be the parameter vector of the model. We now discuss the estimation of the model, followed by the tests of functional form and heteroscedasticity in the next section.

2.1 Estimation

Let $\Omega(\gamma) = \text{diag}\{\omega_1^2(\gamma), \cdots, \omega_n^2(\gamma)\}$. We denote the $n \times k$ regression matrix by $\mathbf{X}(\lambda)$ and the $n \times 1$ vector of (untransformed) dependent variable by $\mathbf{Y}$. The log-likelihood function of model (1), ignoring the constant, is

$$\begin{aligned}\ell(\psi) \;\; &= \;\; -\frac{n}{2}\log\sigma^2 - \sum_{i=1}^{n}\log\omega_i(\gamma) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[\frac{h(y_i, \lambda) - x'_i(\lambda)\beta}{\omega_i(\gamma)}\right]^2 \\ &\quad + \sum_{i=1}^{n}\log h_y(y_i, \lambda), \end{aligned} \tag{3}$$

[3]Throughout this paper we assume that y_i and all elements of x_{1i} are positive.

where $h_y(y,\lambda) = \partial h(y,\lambda)/\partial y$.[4] For given γ and λ, the constrained MLE of β and σ^2 are

$$\hat{\beta}(\gamma,\lambda) = [\mathbf{X}'(\lambda)\Omega^{-1}(\gamma)\mathbf{X}(\lambda)]^{-1}\mathbf{X}'(\lambda)\Omega^{-1}(\gamma)h(\mathbf{Y},\lambda), \tag{4}$$

$$\hat{\sigma}^2(\gamma,\lambda) = \frac{1}{n}h'(\mathbf{Y},\lambda)\mathbf{M}(\gamma,\lambda)\Omega^{-1}(\gamma)\mathbf{M}(\gamma,\lambda)h(\mathbf{Y},\lambda), \tag{5}$$

where $\mathbf{M}(\gamma,\lambda) = I_n - \mathbf{X}(\lambda)[\mathbf{X}'(\lambda)\Omega^{-1}(\gamma)\mathbf{X}(\lambda)]^{-1}\mathbf{X}'(\lambda)\Omega^{-1}(\gamma)$. Substituting these expressions into (3) gives the following concentrated (profile) log-likelihood of γ and λ,

$$\ell_p(\gamma,\lambda) = n\log\left[\dot{J}(\lambda)/\dot{\omega}(\gamma)\right] - \frac{n}{2}\log\hat{\sigma}^2(\gamma,\lambda), \tag{6}$$

where $\dot{\omega}(\gamma)$ and $\dot{J}(\lambda)$ are the geometric means of $\omega_i(\gamma)$ and $J_i(\lambda) = h_y(y_i,\lambda)$, respectively.

Maximizing (6) over γ gives the constrained MLE of γ given λ, denoted as $\hat{\gamma}_c$; maximizing (6) over λ gives the constrained MLE of λ given γ, denoted as $\hat{\lambda}_c$, and maximizing (6) jointly over γ and λ gives the unconstrained MLE of γ and λ, denoted as $\hat{\gamma}$ and $\hat{\lambda}$, respectively. Substituting $\hat{\gamma}_c$ into equations (4) and (5) gives the MLE of β and σ^2 with λ constrained. Likewise, substituting $\hat{\lambda}_c$ into equations (4) and (5) gives the MLE of β and σ^2 with γ constrained. The unconstrained MLE of β and σ^2 are obtained, however, when $\hat{\gamma}$ and $\hat{\lambda}$ are substituted into equations (4) and (5). To facilitate the construction of various test statistics, the Appendix provides the first- and second-order partial derivatives of the log-likelihood function.

If the transformation and weight parameters are given, equation (2) can be converted to the standard linear regression equation by pre-multiplying $\Omega^{-\frac{1}{2}}(\gamma)$ on each side of equation (2), where $\Omega^{-\frac{1}{2}}(\gamma) = \text{diag}\{\omega_1^{-1}(\gamma),\cdots,\omega_n^{-1}(\gamma)\}$. Thus, the standard linear regression theory applies after replacing $h(\mathbf{Y},\lambda)$ by $\Omega^{-\frac{1}{2}}(\gamma)h(\mathbf{Y},\lambda)$ and $\mathbf{X}(\lambda)$ by $\Omega^{-\frac{1}{2}}(\gamma)\mathbf{X}(\lambda)$.

3 Lagrange Multiplier Tests

Let $S(\psi)$ be the score vector, $H(\psi)$ be the Hessian matrix, and $I(\psi) = -\text{E}[H(\psi)]$ be the information matrix. If $\hat{\psi}_0$ is the constrained MLE of ψ under the constraints imposed by the null hypothesis, the LM statistic takes the following general form

$$\text{LM}_\text{E} = S'(\hat{\psi}_0)I^{-1}(\hat{\psi}_0)S(\hat{\psi}_0). \tag{7}$$

[4]This notation for partial derivatives will be used for $\omega_i(\gamma)$ and $x_i(\lambda)$ as well.

See, for example, Godfrey (1988). As $I(\psi)$ may not be easily obtainable, alternative ways of estimating the information matrix have been proposed. In particular, $I(\psi)$ may be replaced by $-H(\psi)$ or the outer product of the gradient (OPG) $G(\psi)'G(\psi)$, with $G(\psi) = \{\partial \ell_i(\psi)/\partial \psi'\}$, where ℓ_i is the element of the log likelihood ℓ corresponding to the ith observation. Hence, the Hessian form and the OPG form of the LM statistic, denoted by $\mathrm{LM_H}$ and $\mathrm{LM_G}$, respectively, can be calculated as follows:

$$\begin{aligned} \mathrm{LM_H} &= -S'(\hat{\psi}_0)H^{-1}(\hat{\psi}_0)S(\hat{\psi}_0), && (8) \\ \mathrm{LM_G} &= \mathbf{1_n}' G(\hat{\psi}_0)[G'(\hat{\psi}_0)G(\hat{\psi}_0)]^{-1}G'(\hat{\psi}_0)\mathbf{1_n}, && (9) \end{aligned}$$

where $\mathbf{1_n}$ is an n-dimensional vector of unity. We shall denote $D(\hat{\psi}_0) = G'(\hat{\psi}_0)G(\hat{\psi}_0)$.

In addition, the LM statistic can also be calculated from the double-length artificial regression proposed by Davidson and MacKinnon (1984). We denote this version of the LM statistic by $\mathrm{LM_D}$. Defining $e_i(\psi) = [h(y_i, \lambda) - x_i'(\lambda)\beta]/[\sigma\omega_i(\gamma)]$, $\mathrm{LM_D}$ is the explained sum of squares of the following regression (with $2n$ observations and $k + p + q + 1$ regressors):

$$\begin{pmatrix} e(\hat{\psi}_0) \\ \mathbf{1_n} \end{pmatrix} = \begin{pmatrix} -\partial e(\hat{\psi}_0)/\partial \psi' \\ \partial \left(\log |\partial e(\hat{\psi}_0)/\partial y|\right)/\partial \psi' \end{pmatrix} \theta + \varepsilon \tag{10}$$

where $e(\psi) = \{e_i(\psi)\}$, and θ and ε are, respectively, the regression parameter and residual. The $\mathrm{LM_D}$ statistic has been found to outperform the $\mathrm{LM_H}$ and $\mathrm{LM_G}$ statistics in finite-sample performance (Davidson and MacKinnon, 1993), and has been applied by many authors in different situations (see Tse, 1984, and Baltagi and Li, 2000, among others).

Let ψ be divided into two components denoted by ψ_1 and ψ_2, with similar notations for the score functions, the Hessian matrix and the information matrix. If the null hypothesis concerns the parameters in ψ_2 only, then $\mathrm{LM_E}$ can be written as

$$\mathrm{LM_E} = S_2'(\hat{\psi}_0)I^{22}(\hat{\psi}_0)S_2(\hat{\psi}_0),$$

where $I^{22}(\psi)$ denotes the relevant block of $I^{-1}(\psi)$.[5] Similar expressions for $\mathrm{LM_H}$ and $\mathrm{LM_G}$ can be obtained.

[5]Similar notations will be used for $H(\psi)$ and $D(\psi)$.

Although the four forms of LM statistic are asymptotically equivalent with the same limiting chi-squared distribution under the null, $\mathrm{LM_E}$ is expected to give the best finite-sample performance.[6] This will be verified empirically in our present context using Monte Carlo experiment.

We now consider the LM tests for functional form and heteroscedasticity, both jointly and conditionally.

3.1 Tests for functional form and heteroscedasticity

Equations (8), (9) and (10) can be used as generic formulae for the Hessian-based, gradient-based and double-length regression based forms of the LM statistic, respectively, for various tests (joint and conditional) of functional form and heteroscedasticity. The derivatives given in the Appendix facilitate the calculation of these statistics, and they are applicable for arbitrary transformations and weight functions. In what follows we focus on the $\mathrm{LM_E}$ statistic, which requires the evaluation of the expectation of the Hessian matrix.

While equation (1) describes the general model, the following assumptions give rise to simplified calculations in specific cases.

Assumption 1. The transformation is given by the Box-Cox power transformation (Box and Cox, 1964):

$$h(y, \lambda) = \begin{cases} (y^\lambda - 1)/\lambda, & \text{if } \lambda \neq 0, \\ \log y, & \text{if } \lambda = 0. \end{cases} \tag{11}$$

with the condition $\mathrm{Max}\{|\theta_i|\} \ll 1$, where $\theta_i = \lambda\sigma\,\omega_i(\gamma)/(1 + \lambda\eta_i)$ with $\eta_i = x_i'(\lambda)\beta$.

Commonly adopted functional forms under the Box-Cox transformation are $\lambda = 1$ (linear) and $\lambda = 0$ (log-linear).[7] While tests of linear versus log-linear functional forms have been the focus of many papers in the literature, we shall generalize the results to

[6] Bera and MacKenzie (1986) has argued for the superior small-sample performance of $\mathrm{LM_E}$ over $\mathrm{LM_H}$ and $\mathrm{LM_G}$, which has been found to be empirically supported. Also, the superior performance of $\mathrm{LM_D}$ over $\mathrm{LM_H}$ and $\mathrm{LM_G}$ in small samples has been shown in many empirical studies (see Davidson and MacKinnon, 1983, 1984). We shall show below, however, that $\mathrm{LM_D}$ is dominated by $\mathrm{LM_E}$ in tests of functional form and heteroscedasticity.

[7] See Godfrey and Wickens (1981), Tse (1984), Davidson and MacKinnon (1985) and Baltagi (1997).

testing for arbitrary values of λ.

It is well known that the Box-Cox power transformation works only for nonnegative observations and it is bounded below or above when $\lambda \neq 0$. An alternative transformation to circumvent this truncation problem is given by the following:

$$h(y,\lambda) = \begin{cases} (y^{\lambda} - y^{-\lambda})/(2\lambda), & \text{if } \lambda \neq 0, \\ \log y, & \text{if } \lambda = 0, \end{cases} \tag{12}$$

which is called the *Dual-Power Transformation* (Yang, 2002). Other proposals can be found in Bickel and Doksum (1981), Burbidge, Magee and Robb (1988), MacKinnon and Magee (1990), and Yeo and Johnson (2000).

Under the Box-Cox transformation, requiring $\text{Max}\{|\theta_i|\}$ to be small is equivalent to requiring the truncation effect to be small. This is seen as follows. Since $(y_i^{\lambda} - 1)/\lambda = x_i'(\lambda)\beta + \sigma\omega_i(\gamma)e_i$, we have $y_i^{\lambda} = 1 + \lambda x_i'(\lambda)\beta + \lambda\sigma\omega_i(\gamma)e_i$. As $y_i > 0$ implies $y_i^{\lambda} > 0$, this in turn implies $|\lambda\sigma\omega_i(\gamma)| \ll 1 + \lambda x_i'(\lambda)\beta$ for the truncation on e_i to be negligible.

The assumption below refers to the heteroscedastic structure, and has been adopted by Breusch and Pagan (1979), among others.

Assumption 2. The weighting function satisfies $\omega(v_i, \gamma) = \omega(v_i'\gamma)$.

A consequence of this assumption is that $\omega_{i\gamma}(\gamma) = \partial\omega(v_i, \gamma)/\partial\gamma$ evaluated at $\gamma = 0$ (homoscedasticity assumption) is equal to a constant multiple of v_i. This assumption incorporates many commonly used heteroscedastic models, including the multiplicative model of Harvey (1976).

We now consider testing for functional form and heteroscedasticity jointly, with the null hypothesis given by

$$H_0 : \gamma = \gamma_0, \lambda = \lambda_0,$$

where γ_0 and λ_0 are arbitrary hypothesized values. The following theorem provides the details for the computation of LM_E.

Theorem 1. Under model (1) and Assumption 1, the $\mathrm{LM_E}$ statistic for testing $H_0 : \gamma = \gamma_0, \lambda = \lambda_0$, is given by

$$\mathrm{LM_E}(\gamma_0, \lambda_0) = S_2'(\hat{\psi}_0) \begin{pmatrix} 2D'AD, & -2D'A\hat{u} \\ -2\hat{u}'AD, & \hat{\xi}_1'\mathbf{M}(\gamma_0, \lambda_0)\hat{\xi}_2 + 2\hat{\phi}'A\hat{\phi} + \frac{3}{2}\hat{\delta}'\hat{\delta} - \kappa \end{pmatrix}^{-1} S_2(\hat{\psi}_0), \tag{13}$$

where $D = \{\omega_{i\gamma}'(\gamma_0)/\omega_i(\gamma_0)\}_{n\times q}$, $A = I_n - (\mathbf{1_n 1_n}')/n$, I_n is an $n \times n$ identity matrix, and $\kappa = 4\phi' A\theta^2 + 2(\mathbf{1_n}'\theta^2)^2/n$ with $\theta^2 = \{\theta_i^2\}$. For $\lambda_0 = 0$, $\kappa = 0$ and various other quantities in equation (13) can be obtained by substituting the constrained MLE into the following $n \times 1$ vectors:

$$\begin{aligned} \delta &= \{\sigma\omega_i(\gamma_0)\} \\ \phi &= \{\eta_i\} \\ u &= \{\eta_i\} \\ \xi_1 &= \left\{\frac{1}{2\sigma\omega_i^2(\gamma_0)}\left[\eta_i^2 - 2x_{i\lambda}'(0)\beta\right]\right\} \\ \xi_2 &= \left\{\frac{1}{2\sigma}\left[\eta_i^2 - 2x_{i\lambda}'(0)\beta\right]\right\}. \end{aligned}$$

For $\lambda_0 \neq 0$, the relevant quantities are:

$$\begin{aligned} \delta &= \{\theta_i/\lambda_0\}, \\ \phi &= \{\lambda_0^{-1}\log(1 + \lambda_0\eta_i)\} \\ u &= \left\{\phi_i - \lambda_0\delta_i^2\right\} \\ \xi_1 &= \left\{\frac{1}{\omega_i(\gamma_0)}\left[\frac{\delta_i}{2} + \frac{\phi_i}{\theta_i} + \frac{\theta_i^3}{2\lambda_0} - \frac{\eta_i}{\lambda_0\sigma\omega(\gamma_0)} - \frac{x_{i\lambda}'(\lambda_0)\beta}{\sigma\omega_i(\gamma_0)}\right]\right\} \\ \xi_2 &= \left\{\omega_i(\gamma_0)\left[\frac{\delta_i}{2} + \frac{\phi_i}{\theta_i} + \frac{\theta_i^3}{2\lambda_0} - \frac{\eta_i}{\lambda_0\sigma\omega(\gamma_0)} - \frac{x_{i\lambda}'(\lambda_0)\beta}{\sigma\omega_i(\gamma_0)}\right]\right\}. \end{aligned}$$

Under H_0, $\mathrm{LM_E}(\gamma_0, \lambda_0) \xrightarrow{D} \chi^2_{p+q}$.

Proof. From the expressions of the elements of the Hessian matrix given in the Appendix, we can show that

$$I^{22}(\psi) = \begin{pmatrix} 2D'AD, & -2D'Au \\ -2u'AD, & \xi_1'\mathbf{M}(\gamma, \lambda)\xi_2 + \mathbf{1_n}'v - \frac{2}{n}(\mathbf{1_n}'u)^2 \end{pmatrix}^{-1},$$

where u, v, ξ_1 and ξ_2 are $n \times 1$ vectors with the ith element being

$$u_i = \mathrm{E}[e_i(\psi)e_{i\lambda}(\psi)],$$

$$
\begin{aligned}
v_i &= \mathrm{Var}[e_{i\lambda}(\psi)] + \mathrm{E}[e_i(\psi)e_{i\lambda\lambda}(\psi)] + \mathrm{E}[\partial^2 \log h_y(y_i,\lambda)/\partial\lambda^2], \\
\xi_{1i} &= \omega_i^{-1}(\gamma)\mathrm{E}[e_{i\lambda}(\psi)], \text{and} \\
\xi_{2i} &= \omega_i(\gamma)\mathrm{E}[e_{i\lambda}(\psi)],
\end{aligned}
$$

where $e_i(\psi) = [h(y_i,\lambda) - x_i'(\lambda)\beta]/[\sigma\omega_i(\gamma)]$, and $e_{i\lambda}$ and $e_{i\lambda\lambda}$ are the first- and second-order partial derivatives of $e_i(\psi)$ with respect to λ. Under the Box-Cox transformation, the case of $\lambda = 0$ follows directly from calculations using $\log y_i = \eta_i + \sigma\omega(\gamma)e_i$. The case of $\lambda \neq 0$ needs approximation. Note that $\log y_i$ has the following expansion

$$
\lambda \log y_i = \log(1+\lambda\eta_i) + \theta_i e_i - \frac{1}{2}\theta_i^2 e_i^2 + \cdots + \frac{(-1)^{k+1}}{k!}\theta_i^k e_i^k + \cdots. \tag{14}
$$

We use the third-order approximation ($k = 3$) in our derivation. The assumption of the theorem ensures its accuracy. With the help of Mathematica, we obtain

$$
\begin{aligned}
\mathrm{E}[e_{i\lambda}(\psi)] &= \left(\frac{\delta_i}{2} + \frac{\phi_i}{\theta_i} + \frac{\theta_i^3}{2\lambda_0} - \frac{\eta_i}{\lambda_0\sigma\omega(\gamma_0)} - \frac{x_{i\lambda}'(\lambda_0)\beta}{\sigma\omega_i(\gamma_0)}\right) + O(\theta_i^4), \\
\mathrm{Var}[e_{i\lambda}(\psi)] &= \frac{1}{\lambda^2}\left(\frac{1}{2}\theta_i^2 - 2\phi\theta_i^2 + \phi_i^2\right) + O(\theta_i^4), \\
\mathrm{E}[e_i(\psi)e_{i\lambda}(\psi)] &= \frac{1}{\lambda^2}\left(\phi_i - \theta_i^2\right) + O(\theta_i^4), \\
\mathrm{E}[e_i(\psi)e_{i\lambda\lambda}(\psi)] &= \frac{1}{\lambda^2}\left(\theta_i^2 - 2\phi\theta_i^2 + \phi_i^2\right) + O(\theta_i^4).
\end{aligned}
$$

Putting the above together and simplifying, we obtain the results of the theorem. The corresponding quantities for the case of $\lambda = 0$ can also be obtained by taking the limits of the above quantities by letting λ go to zero. #

Note that the $\mathrm{LM_E}$ statistic generally depends on the weighting function $\omega(\cdot)$. However, under Assumption 2 and with $\gamma_0 = 0$ (test of homoscedasticity), $\mathrm{LM_E}$ does not depend on the weighting function.[8] This result has been pointed out by Tse (1982) for the cases of testing for linear versus log-linear regressions with homoscedasticity. Theorem 1 extends the result to the general case of a Box-Cox transformation with an arbitrary transformation parameter and an arbitrary weighting function. However, when $\gamma_0 \neq 0$, $\mathrm{LM_E}$ depends on the specification of the weighting function.

[8]This is due to the fact that $\omega_{i\gamma}(0)$ is a constant multiple of v_i.

For testing functional form only, the following theorem is applicable.

Theorem 2. Under Model (1) and Assumption 1, the $\mathrm{LM_E}$ statistic for testing $H_0 : \lambda = \lambda_0$ takes the form

$$\mathrm{LM_E}(\lambda_0) = S'_\lambda(\hat{\psi}_0)\left[\hat{\xi}_1'\mathbf{M}(\hat{\gamma}_c, \lambda_0)\hat{\xi}_2 + 2\hat{\phi}'A\hat{\phi} + \frac{3}{2}\hat{\delta}'\hat{\delta} - \kappa - 2\hat{u}A\hat{D}(\hat{D}'A\hat{D})^{-1}\hat{D}'A\hat{u}\right]^{-1} S_\lambda(\hat{\psi}_0), \tag{15}$$

where the ith row of $\hat{D}$ is $\omega'_{i\gamma}(\hat{\gamma}_c)/\omega_i(\hat{\gamma}_c)$. The quantities $\kappa, \delta, \phi, u, \xi_1$ and ξ_2 are as given in Theorem 1, but evaluated at $\hat{\psi}_0 = \{\hat{\beta}'(\hat{\gamma}_c, \lambda_0), \hat{\sigma}^2(\hat{\gamma}_c, \lambda_0), \hat{\gamma}_c', \lambda_0\}'$. Under H_0, $\mathrm{LM_E}(\gamma_0) \xrightarrow{D} \chi^2_p$.

Proof. This follows directly from the proof of Theorem 1.

Note that $\mathrm{LM_E}$ in Theorem 2 depends on the weighting function $\omega(\cdot)$.

For testing heteroscedasticity only, the following theorem applies.

Theorem 3. Under Model (1), when λ is known, the $\mathrm{LM_E}$ statistic for testing $H_0 : \gamma = \gamma_0$ is[9]

$$\mathrm{LM_E}(\gamma_0|\lambda) = \frac{1}{2}g'(\gamma_0, \lambda)D_1(D_1'D_1)^{-1}D_1'g(\gamma_0, \lambda), \tag{16}$$

where $D_1 = (\mathbf{1_n}, D)$ and the ith element of $g(\gamma_0, \lambda)$ is

$$g_i(\gamma_0, \lambda) = \left(\frac{h(y_i, \lambda) - x_i'(\lambda)\hat{\beta}(\gamma_0, \lambda)}{\omega_i(\gamma_0)\hat{\sigma}(\gamma_0, \lambda)}\right)^2 - 1. \tag{17}$$

Under H_0, $\mathrm{LM_E}(\gamma_0|\lambda) \xrightarrow{D} \chi^2_q$. When λ is unknown and is replaced by $\hat{\lambda}_c$, we have

$$\mathrm{LM_E}(\gamma_0|\hat{\lambda}_c) \overset{d}{=} \mathrm{LM_E}(\gamma_0|\lambda),$$

i.e., the two statistics have the same asymptotic distribution. Furthermore, if Assumption 2 holds and $\gamma_0 = 0$,

$$\mathrm{LM_E}(\gamma_0|\hat{\lambda}_c) = \frac{1}{2}g'(\gamma_0, \hat{\lambda}_c)V_1(V_1'V_1)^{-1}V_1'g(\gamma_0, \hat{\lambda}_c), \tag{18}$$

[9]The notation $\mathrm{LM_E}(\gamma_0|\lambda)$ emphasizes that the statistic is evaluated at λ, which is the true but unknown transformation parameter. Note that λ is neither estimated nor tested. Thus, $\mathrm{LM_E}(\gamma_0|\lambda)$ is not an operational test statistic. In contrast, $\mathrm{LM_E}(\gamma_0|\hat{\lambda}_c)$ is the LM statistic for the heteroscedasticity parameter γ_0 evaluated at $\hat{\lambda}_c$. Note that the transformation parameter λ, however, is not tested.

where $V_1 = (\mathbf{1_n}, V)$ and V is the $n \times q$ matrix with the ith row being v_i', in which case $\mathrm{LM}_E(\gamma_0|\hat{\lambda}_c)$ does not depend on the weighting function.

Proof. When λ is known, the evaluation of the expectation of $H(\psi)$ is straightforward, and the derivation of equation (16) parallels that in Breusch and Pagan (1979). However, the statistic applies to a general value of γ_0, and is generally dependent on the weighting function. To prove $\mathrm{LM_E}(\gamma_0|\hat{\lambda}_c)$ and $\mathrm{LM_E}(\gamma_0|\lambda)$ have the same asymptotic distribution it is sufficient to show that $D_1'(g(\gamma_0, \hat{\lambda}_c) - g(\gamma_0, \lambda))/\sqrt{n} \xrightarrow{p} 0$. Using Taylor expansion, we have

$$\frac{1}{\sqrt{n}} D_1'(g(\gamma_0, \hat{\lambda}_c) - g(\gamma_0, \lambda)) \stackrel{d}{=} \frac{1}{\sqrt{n}} D_1' g_\lambda(\gamma_0, \lambda)(\hat{\lambda}_c - \lambda).$$

As $\hat{\lambda}_c - \lambda \xrightarrow{p} 0$, the result follows provided $D_1' g_\lambda(\gamma_0, \lambda)/\sqrt{n}$ converges to a finite random variable. Finally, under Assumption (2), $\omega_{i\gamma}(0) = cv_i$ for a constant c, which leads to equation (18). Without loss of generality, $\omega_i(0)$ can be set equal to 1 so that $\mathrm{LM_E}(0|\hat{\lambda}_c)$ does not depend on the weighting function. #

The theorem above provides a conditional test for heteroscedasticity with a general weighting function (not necessarily satisfying Assumption 2). It also shows that the Breusch-Pagan test holds when the transformation (not necessarily the Box-Cox transformation) parameters are estimated using MLE (or any consistent estimator).

4 Some Monte Carlo results

In this section we present some Monte Carlo results for the finite-sample performance of the tests discussed in the last section. We shall focus on the empirical size of the tests in small samples. Power comparison can only be meaningful when the tests are corrected for size. As finite-sample critical values are not available unless extensive simulation is conducted, we shall not examine the power of the size-corrected tests.

Joint test of heteroscedasticity and functional form. We consider the following data generation process (DGP) when the data transformation is only applied to the response variable:

$$h(y_i, \lambda) = \beta_0 + \beta_1 x_{1i} + \sigma \exp(\gamma x_{1i}) e_i, \qquad i = 1, \cdots, n, \tag{19}$$

with $x_{1i} \sim U(0, 25)$ (i.e., uniformly distributed in the interval 0 to 25), and $(\beta_0, \beta_1) = (25, 10)$. We examine the performance of $\mathrm{LM_E}$ given in Theorem 1 under the assumption of the Box-Cox transformation. The following parameters are considered: $\lambda_0 = 0,\ 0.2,\ 0.5,\ 0.8,\ 1$, $\gamma_0 = 0,\ 0.1,\ 0.2,\ 0.3$, and $\sigma = 0.1,\ 0.5$. Using 10,000 Monte Carlo sample runs, we calculate the empirical relative rejection frequency (empirical size) for the test of $H_0 : \lambda = \lambda_0, \gamma = \gamma_0$.

For the case when the data transformation is applied to some regressors as well, we consider the following DGP:

$$h(y_i, \lambda) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}(\lambda) + \sigma \exp(\gamma_1 x_{1i} + \gamma_2 x_{2i}) e_i, \qquad i = 1, \cdots, n, \tag{20}$$

with x_{1i} and x_{2i} both generated from $U(0, 25)$. The parameters are given by: $(\beta_0, \beta_1, \beta_2) = (25, 10, 10)$, $\lambda_0 = 0,\ 0.2,\ 0.5,\ 0.8,\ 1$, $\gamma_{10} = 0,\ 0.1,\ 0.2, 0.3$, $\gamma_{20} = 0,\ 0.1,\ 0.2, 0.3$, and $\sigma = 0.1,\ 0.5$.

Table 1 summarizes the empirical size of the tests for a nominal size of five percent. The $\mathrm{LM_E}$, $\mathrm{LM_D}$, LR, $\mathrm{LM_H}$ and $\mathrm{LM_G}$ statistics are considered. $\mathrm{LM_E}$ is computed using equation (13). Panel A provides the results when there is no regressor transformation, while Panel B summarizes the cases with regressor transformation. As the relative performance of the test statistics is qualitatively similar for all parameter values considered, only a subset of the results are presented.[10] We observe that the $\mathrm{LM_E}$ statistic provides appropriate empirical size even for small sample size of $n = 30$. For the $\mathrm{LM_H}$ and $\mathrm{LM_G}$ statistics, the over-rejection of the null hypothesis is very serious, although the performance improves for sample size $n = 80$. The $\mathrm{LM_G}$ version appears to be worse than the $\mathrm{LM_H}$ version. The LR statistic performs better than both $\mathrm{LM_H}$ and $\mathrm{LM_G}$, but is inferior to $\mathrm{LM_E}$. The $\mathrm{LM_D}$ statistic outperforms LR, $\mathrm{LM_H}$ and $\mathrm{LM_G}$. Its performance is comparable to that of $\mathrm{LM_E}$, except for the case with regressor transformation with $n = 30$ and large γ_{20}, where there is slight over-rejection.

[10] Other results are available from the authors on request.

Test of functional form. We now evaluate the finite-sample properties of the proposed LM tests for functional form. The DGP in equations (19) and (20) are used for the cases of without and with regressor transformation, respectively. Theorem 2 is used for the computation of LM_E, and the parameter values and Monte Carlo sample size are the same as in the case of the joint test.

Table 2 summarizes the empirical sizes of the tests of functional form. Again, the LM_E statistic performs the best, followed by LM_D, LR, LM_H and LM_G. While LM_D performs quite well for $n = 80$, over-rejection can be serious for $n = 30$, reaching empirical size of almost 10 percent in some cases. In contrast, while the LM_E statistic shows slight over-rejection for $n = 30$ with regressor transformation, the over-rejection appears to be mild.

Test of heteroscedasticity. Again the same parametric setup and DGPs are used in the Monte Carlo experiment. Two data transformations are considered: the Box-Cox transformation and the dual-power transformation. We compare the performance of $\text{LM}_\text{E}(\gamma_0|\lambda)$, $\text{LM}_\text{E}(\gamma_0|\hat{\lambda}_c)$ and LR. As the expected information matrix can be easily computed, and Theorem 3 shows that $\text{LM}_\text{E}(\gamma_0|\lambda)$ and $\text{LM}_\text{E}(\gamma_0|\hat{\lambda}_c)$ are asymptotically equivalent, we do not consider other versions of LM statistic. As . We are in particular interested in the finite-sample performance of $\text{LM}_\text{E}(\gamma_0|\hat{\lambda}_c)$, as the transformation parameter is unknown in practice.

Table 3 summarizes the results when there is no regressor transformation, while Table 4 presents the results when there is regressor transformation. The asymptotic equivalence of $\text{LM}_\text{E}(\gamma_0|\lambda)$ and $\text{LM}_\text{E}(\gamma_0|\hat{\lambda}_c)$ appears to hold well in small sample, irrespective of whether there is regressor transformation. Also, the empirical size of these tests approximates well the nominal level. In comparison, the LR statistic performs rather poorly for $n = 30$ when there is regressor transformation.

For each test, whether of joint hypotheses or separate hypotheses, a small scale Monte Carlo experiment has been conducted to examine the power of the tests. For the LR, LM_D, LM_H, and LM_G statistics, empirical critical values are estimated by simulation. On a size-corrected basis, we find that the empirical powers of the tests are comparable.

The $\mathrm{LM_E}$ statistic, however, offers an advantage that its nominal size is reliable in small samples, and size correction by simulation is not required.

5 Conclusions

We have presented LM statistics for testing functional form and heteroscedasticity, both jointly and separately, based on the expected information matrix. In the joint test and the test of functional form, analytic formulae are derived assuming the Box-Cox transformation. The tests apply to a specified functional form and heteroscedasticity in general, and are not restricted to the cases of linear versus log-linear (or simple difference versus percentage change) regressions with homoscedasticity. When testing for heteroscedasticity, formulae are obtained conditional on the functional form. Replacing the parameters of the transformation by their constrained MLE, however, does not alter the asymptotic distribution of the test statistic. Our formulae generalize the results of testing for homoscedasticity in the literature.

Our Monte Carlo study shows that the finite-sample performance of the expected-information version is superior to the versions based on the double-length regression, the observed-information and the OPG. It also performs well against the LR statistic, providing empirical size closer to the nominal level. Due to its comparative simplicity in computation versus the LR statistic and its superior small-sample performance, the $\mathrm{LM_E}$ statistic is recommended. It provides convenient diagnostics for functional form and heteroscedasticity with desirable small-sample properties.

APPENDIX: Scores and Observed Information

For the model with a general transformation and a general weighting function, the score function $S(\psi)$ has the following elements:

$$
\begin{aligned}
S_\beta &= \frac{1}{\sigma^2}\sum_{i=1}^n \frac{[h(y_i,\lambda)-x_i'(\lambda)\beta]x_i(\lambda)}{\omega_i^2(\gamma)} \\
S_{\sigma^2} &= \frac{1}{2\sigma^4}\sum_{i=1}^n \frac{[h(y_i,\lambda)-x_i'(\lambda)\beta]^2}{\omega_i^2(\gamma)} - \frac{n}{2\sigma^2} \\
S_\gamma &= \frac{1}{\sigma^2}\sum_{i=1}^n \frac{\omega_{i\gamma}(\gamma)}{\omega_i^3(\gamma)}[h(y_i,\lambda)-x_i'(\lambda)\beta]^2 - \sum_{i=1}^n \frac{\omega_{i\gamma}(\gamma)}{\omega_i(\gamma)} \\
S_\lambda &= \sum_{i=1}^n \frac{h_{y\lambda}(y_i,\lambda)}{h_y(y_i,\lambda)} - \frac{1}{\sigma^2}\sum_{i=1}^n \frac{[h(y_i,\lambda)-x_i'(\lambda)\beta][h_\lambda(y_i,\lambda)-x_{i\lambda}'(\lambda)\beta]}{\omega_i^2(\gamma)}
\end{aligned}
$$

from which the gradient matrix can be easily formulated. The elements of the Hessian matrix $H(\psi) = \partial S(\psi)/\partial\psi'$ are given by:

$$
\begin{aligned}
H_{\beta\beta'} &= -\sigma^{-2}\sum_{i=1}^n x_i(\lambda)x_i'(\lambda)\omega_i^{-2}(\gamma) \\
H_{\sigma^2\sigma^2} &= -\sigma^{-4}\sum_{i=1}^n e_i^2(\psi) + \frac{n}{2\sigma^4} \\
H_{\gamma\gamma'} &= -\sum_{i=1}^n \left[\frac{\omega_{i\gamma\gamma'}(\gamma)}{\omega_i(\gamma)} - \frac{\omega_{i\gamma}(\gamma)\omega_{i\gamma}'(\gamma)}{\omega_i^2(\gamma)}\right] + \sum_{i=1}^n e_i^2(\psi)\left[\frac{\omega_{i\gamma\gamma'}(\gamma)}{\omega_i(\gamma)} - \frac{3\omega_{i\gamma}(\gamma)\omega_{i\gamma}'(\gamma)}{\omega_i^2(\gamma)}\right] \\
H_{\lambda\lambda} &= -\sum_{i=1}^n \left[e_{i\lambda}^2(\psi) + e_i(\psi)e_{i\lambda\lambda}(\psi)\right] + \sum_{i=1}^n \left[\partial^2 \log h_y(y_i,\lambda)/(\partial\lambda^2)\right] \\
H_{\beta\sigma^2} &= -\sigma^{-3}\sum_{i=1}^n e_i(\psi)x_i(\lambda)\omega_i^{-1}(\gamma) \\
H_{\beta\gamma'} &= -2\sigma^{-1}\sum_{i=1}^n e_i(\psi)x_i(\lambda)\omega_{i\gamma}'(\gamma)\omega_i^{-2}(\gamma) \\
H_{\beta\lambda} &= \sigma^{-1}\sum_{i=1}^n \left[e_{i\lambda}(\psi)x_i(\lambda) + e_i(\psi)x_{i\lambda}(\lambda)\right]\omega_i^{-1}(\gamma) \\
H_{\sigma^2\gamma} &= -\sigma^{-2}\sum_{i=1}^n e_i^2(\psi)\omega_{i\gamma}(\gamma)\omega_i^{-1}(\gamma) \\
H_{\sigma^2\lambda} &= \sigma^{-2}\sum_{i=1}^n e_i(\psi)e_{i\lambda}(\psi) \\
H_{\gamma\lambda} &= 2\sum_{i=1}^n e_i(\psi)e_{i\lambda}(\psi)\omega_{i\gamma}(\gamma)\omega_i^{-1}(\gamma).
\end{aligned}
$$

Now, for the Box-Cox transformation, we have $h_y(y,\lambda) = y^{\lambda-1}$, $h_{y\lambda}(y,\lambda) = y^{\lambda-1}\log y$, $h_{y\lambda\lambda}(y,\lambda) = y^{\lambda-1}(\log y)^2$, and

$$
h_\lambda(y,\lambda) = \begin{cases} \frac{1}{\lambda}[1+\lambda h(y,\lambda)]\log y - \frac{1}{\lambda}h(y,\lambda), & \lambda \neq 0, \\ \frac{1}{2}(\log y)^2, & \lambda = 0, \end{cases}
$$

$$h_{\lambda\lambda}(y,\lambda) = \begin{cases} h_\lambda(y,\lambda)(\log y - \frac{1}{\lambda}) + \frac{1}{\lambda^2}[h(y,\lambda) - \log y], & \lambda \neq 0, \\ \frac{1}{3}(\log y)^3, & \lambda = 0. \end{cases}$$

For the dual-power transformation of Yang (2002), we have $h_y(y,\lambda) = \frac{1}{2}[y^{\lambda-1} + y^{-\lambda-1}]$, $h_{y\lambda}(y,\lambda) = \frac{1}{2}(y^{\lambda-1} - y^{-\lambda-1})\log y$, $h_{y\lambda\lambda}(y,\lambda) = \frac{1}{2}(y^{\lambda-1} + y^{-\lambda-1})(\log y)^2$, and

$$h_\lambda(y,\lambda) = \begin{cases} \frac{1}{2\lambda}(y^\lambda + y^{-\lambda})\log y - \frac{1}{\lambda}h(y,\lambda), & \lambda \neq 0, \\ 0, & \lambda = 0, \end{cases}$$

$$h_{\lambda\lambda}(y,\lambda) = \begin{cases} h(y,\lambda)(\log y)^2 - \frac{2}{\lambda}h_\lambda(y,\lambda) & \lambda \neq 0, \\ \frac{1}{3}(\log y)^3, & \lambda = 0. \end{cases}$$

These partial derivatives are also available for other transformations such as Yeo and Johnson (2000),

References

Ali, M. M. and Giaccotto, C. (1984). A study of several new and existing tests for heteroscedasticity in the general linear model. *Journal of Econometrics*, **26**, 355-373.

Amemiya, T. (1977). A note on a heteroscedastic model. *Journal of Econometrics*, **6,** 365-370.

Baltagi, B. H. (1997). Testing linear and loglinear error components regressions against Box-Cox alternatives. *Statistics and Probability Letters*, **33**, 63-68.

Baltagi, B. H. and Li, D. (2000). Double-length regressions for the Box-Cox difference model with heteroscedasticity or autocorrelation. *Economics Letters*, **69**, 9-14.

Bera, A. K. and MacKenzie, C. (1986). Alternative forms and properties of the score test. *Journal of Applied Statistics*, **13**, 13-25.

Bickel, P. J. and Doksum, K. A. (1981). An analysis of transformations revisited. *Journal of the American Statistical Association*, **76**, 296–311.

Box, G.E.P. and Cox, D.R. (1964). An analysis of transformation (with discussion). *Journal of the Royal Statistical Society B*, **26**, 211-252.

Breusch, T. S. and Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, **47**, 1287-1293.

Burbidge, J. B., Magee, L. and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, **83**, 123-127.

Davidson, R. and MacKinnon, J. G. (1983). Small sample properties of alternative forms of the Lagrange multiplier test. *Economics Letters*, **12**, 269-275.

Davidson, R. and MacKinnon, J. G. (1984). Model specification tests based on artificial linear regressions. *International Economic Review*, **25**, 485-502.

Davidson, R. and MacKinnon, J. G. (1985). Testing linear and loglinear regression against Box-Cox alternatives. *Canadian Journal of Economics*, **18**, 499-517.

Davidson, R. and MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.

Evans, M. A. and King, M. L. (1988). A further class of tests for heteroscedasticity. *Journal of Econometrics*, **37**, 265-276.

Farebrother, R. W. (1987). The statistical foundations of a class of parametric tests for heteroscedasticity. *Journal of Econometrics*, **36**, 359-368.

Glejser, H. (1969). A new test for heteroscedasticity. *Journal of the American Statistical Association*, **64**, 316-323.

Godfrey, L. G. (1988). *Misspecification Tests in Econometrics.* Cambridge University Press.

Godfrey, L. G., McAleer, M. and McKenzie, C.R. (1988). Variable addition and Lagrange Multiplier tests for linear and logarithmic regression models. *Review of Economics and Statistics*, **70**, 492-503.

Godfrey, L. G. and Wickens, M. R. (1981). Testing linear and log-linear regressions for functional form. *Review of Economic Studies*, **48**, 487-496

Goldfeld, S. M. and Quandt, R. E. (1965). Some tests for homoscedasticity. *Journal of the American Statistical Association*, **60**, 539-547.

Griffiths, W. E. and Surekha, K. (1986). A Monte Carlo evaluation of the power of some tests for heteroscedasticity. *Journal of Econometrics*, **31**, 219-231.

Harvey, A. C. (1976). Estimating regression models with multiplicative heteroscedasticity. *Econometrica*, **44**, 461-465.

Kalirajan, K. P. (1989). A test for heteroscedasticity and non-normality of regression residuals. *Economics Letters*, **30**, 133-136.

Koenker, R. (1981). A note on studentizing a test for heteroscedasticity. *Journal of Econometrics*, **17**. 107-112.

Lahiri, K. and Egy, D. (1981). Joint estimation and testing for functional form and heteroscedasticity. *Journal of Econometrics*, **15**, 299-307.

Lawrance, A. J. (1987). The score statistic for regression transformation. *Biometrika*, **74**, 275-279.

Layson, S. K. and Seaks, T. G. (1984). Estimation and testing for functional form in first difference models. *Review of Economics and Statistics*, **66**, 338-343.

MacKinnon, J.G. and Magee, L. (1990). Transforming the dependent variable in regression models. *International Economic Review*, **31**, 315-339.

Maekawa, K. (1988). Comparing the Wald, LR and LM tests for heteroscedasticity in a linear regression model. *Economics Letters*, **26**, 37-41.

Seaks, T. G. and Layson, S. K. (1983). Box-Cox estimation with standard econometric problems. *Review of Economics and Statistics*, **65**, 160-164.

Tse, Y. K. (1984). Testing for linear and log-linear regression with heteroscedasticity. *Economics Letters*, **16**, 63-69.

Wallentin, B. and Agren, A. (2002). Test of heteroscedasticity in a regression model in the presence of measurement errors. *Economics Letters*, **76**, 205-211.

Yang, Z. L. (2002). A Modified Family of Power Transformations. *Working Paper #12-2002.* School of Economics and Social Sciences, Singapore Management University.

Yang, Z. L. and Abeysinghe, T. (2003). A score test for Box-Cox functional form. *Economics Letters*, **79**, 107-115.

Yeo, I. K. and Johnson, R. A. (2000). A new family of power transformation to improve normality or symmetry. *Biometrika*, **87**, 954-959.

Zarembka, P. (1974). Transformation of variables in econometrics, in: Paul Zarembka, ed., *Frontiers in Econometrics*, Academic Press, New York.

Table 1. Empirical size (%) for joint tests of functional form and heteroscedasticity, H_0: $\gamma = \gamma_0, \lambda = \lambda_0$, at $\alpha = 5\%$ with $\sigma = 0.1$

			$n = 30$					$n = 80$				
λ_0	γ_{10}	γ_{20}	LM_E	LM_D	LR	LM_H	LM_G	LM_E	LM_D	LR	LM_H	LM_G
A: Without regressor transformation												
0.0	0.1		4.41	4.84	6.87	11.76	17.22	5.01	5.32	5.80	7.44	11.08
0.2	0.1		5.12	5.75	7.67	12.66	17.65	4.68	5.02	5.68	7.21	10.97
0.5	0.1		4.85	5.68	7.74	12.33	18.16	4.89	5.19	5.77	7.13	10.43
0.8	0.1		4.82	5.65	7.53	12.49	17.72	4.35	4.74	5.19	6.74	10.72
1.0	0.1		5.00	5.86	8.01	12.81	17.88	5.01	5.27	5.84	7.45	11.02
0.0	0.2		4.93	5.48	7.27	12.29	18.01	5.17	5.43	5.86	7.37	10.87
0.2	0.2		5.00	5.73	7.63	12.32	17.35	4.96	5.32	5.90	7.37	10.78
0.5	0.2		4.77	5.47	7.36	12.58	17.95	4.94	5.30	5.89	7.31	11.06
0.8	0.2		5.01	5.77	7.53	12.24	17.96	4.34	4.76	5.40	6.76	10.27
1.0	0.2		4.84	5.41	7.34	12.32	17.58	4.87	5.32	5.83	7.37	11.04
B: With regressor transformation												
0.0	0.0	0.1	5.07	6.39	11.87	22.73	26.52	5.02	5.53	6.92	10.07	16.16
0.2	0.0	0.1	4.76	5.87	11.55	21.39	26.23	4.81	5.43	6.59	9.61	15.92
0.5	0.0	0.1	4.39	5.46	10.76	20.94	24.27	4.75	5.41	6.50	9.00	15.33
0.8	0.0	0.1	4.92	5.49	11.82	21.17	24.53	5.01	5.21	6.69	9.32	15.28
1.0	0.0	0.1	4.61	5.27	11.50	20.67	24.45	5.07	5.13	6.68	8.70	15.12
0.0	0.0	0.2	5.47	6.65	11.44	22.74	27.66	5.22	6.02	7.21	10.60	17.24
0.2	0.0	0.2	5.06	6.40	11.85	21.83	27.40	4.96	5.74	7.14	10.23	16.12
0.5	0.0	0.2	4.94	6.19	12.17	22.33	26.58	4.91	5.21	6.52	9.43	15.30
0.8	0.0	0.2	4.54	5.97	12.42	22.13	24.62	4.59	5.05	6.50	9.11	15.33
1.0	0.0	0.2	4.98	5.70	11.99	21.87	24.51	4.84	5.19	6.62	9.09	15.66
0.0	0.0	0.3	5.13	6.72	11.99	23.06	28.32	5.53	6.48	7.82	11.56	18.08
0.2	0.0	0.3	5.20	7.11	11.74	22.47	28.60	4.79	5.53	6.93	10.59	16.63
0.5	0.0	0.3	4.93	6.51	11.99	22.73	27.18	5.31	6.11	7.45	10.75	17.05
0.8	0.0	0.3	5.08	6.39	12.03	22.83	25.39	5.01	5.44	7.19	9.72	15.61
1.0	0.0	0.3	5.10	6.24	12.08	22.10	24.29	5.16	5.58	6.96	9.48	15.95
0.0	0.1	0.1	5.35	6.13	9.88	21.20	25.59	4.86	5.01	6.57	9.25	15.55
0.2	0.1	0.1	4.86	5.65	10.91	20.61	25.95	4.82	5.25	6.27	8.95	15.09
0.5	0.1	0.1	4.88	4.97	10.29	19.50	23.04	4.92	5.31	6.63	9.04	14.90
0.8	0.1	0.1	5.02	4.93	10.34	19.27	23.28	5.20	5.50	6.78	9.14	15.29
1.0	0.1	0.1	5.32	5.60	10.93	19.29	23.95	5.23	5.67	6.76	8.63	14.86
0.0	0.1	0.2	5.23	6.25	10.11	21.29	25.73	5.45	5.73	6.93	9.74	16.22
0.2	0.1	0.2	5.19	6.27	11.32	22.12	27.06	4.70	5.10	6.40	8.92	14.90
0.5	0.1	0.2	5.40	6.02	11.19	20.86	24.08	4.91	5.12	6.71	9.32	14.91
0.8	0.1	0.2	5.40	5.45	10.92	20.62	22.99	5.42	5.48	6.89	9.53	15.45
1.0	0.1	0.2	5.30	5.34	11.20	20.05	22.98	5.10	5.09	6.83	9.18	15.32
0.0	0.1	0.3	5.77	6.91	11.67	23.00	27.99	5.21	5.56	6.98	10.46	16.62
0.2	0.1	0.3	5.95	6.96	11.92	23.18	29.08	5.52	6.12	7.62	11.16	17.46
0.5	0.1	0.3	5.30	6.28	11.47	22.26	26.01	4.96	5.31	6.93	10.21	16.41
0.8	0.1	0.3	5.46	6.20	10.93	20.56	24.02	5.19	5.48	6.76	9.64	15.95
1.0	0.1	0.3	5.31	6.18	11.37	21.02	23.82	5.03	4.98	6.55	9.28	15.70

Table 2. Empirical size (%) for tests of functional form, H_0: $\lambda = \lambda_0$, at $\alpha = 5\%$ with $\sigma = 0.1$

			$n = 30$					$n = 80$				
λ_0	γ_1	γ_2	LM_E	LM_D	LR	LM_H	LM_G	LM_E	LM_D	LR	LM_H	LM_G
A: Without regressor transformation												
0.0	0.1		5.65	6.75	7.45	10.43	12.74	4.61	4.96	5.11	6.38	7.72
0.2	0.1		5.67	6.90	7.80	11.24	13.12	4.78	5.19	5.39	6.53	7.76
0.5	0.1		5.05	6.36	7.17	10.67	12.51	5.15	5.63	5.85	6.86	8.32
0.8	0.1		5.36	6.81	7.65	11.31	12.80	5.10	5.48	5.68	6.68	8.01
1.0	0.1		5.94	7.17	8.04	11.39	13.26	4.87	5.31	5.49	6.47	8.10
0.0	0.2		5.48	6.33	7.00	9.75	12.71	5.19	5.57	5.77	6.77	7.99
0.2	0.2		5.38	6.77	7.64	11.15	13.65	5.56	5.93	6.14	7.32	8.94
0.5	0.2		5.68	6.81	7.75	11.23	13.61	5.37	5.75	5.88	6.86	8.47
0.8	0.2		5.60	6.92	7.69	11.26	13.70	5.23	5.56	5.71	6.82	8.25
1.0	0.2		5.48	6.78	7.79	11.19	13.17	5.01	5.47	5.63	6.50	8.08
B: With regressor transformation												
0.0	0.0	0.1	6.09	8.30	9.04	15.57	21.19	5.54	6.29	6.52	8.41	11.95
0.2	0.0	0.1	5.52	8.04	9.64	15.88	20.11	5.29	6.14	6.33	8.35	11.38
0.5	0.0	0.1	5.81	8.20	9.91	15.73	19.08	5.07	5.53	6.16	7.76	9.99
0.8	0.0	0.1	5.85	7.33	8.99	13.34	15.58	5.19	5.89	6.21	7.55	9.57
1.0	0.0	0.1	5.74	6.94	8.98	12.53	14.61	5.30	5.72	6.18	7.29	9.35
0.0	0.0	0.2	6.28	9.04	9.70	16.12	25.15	5.52	6.20	6.40	8.81	13.10
0.2	0.0	0.2	6.05	8.70	10.33	16.85	25.08	5.10	6.12	6.28	8.39	11.94
0.5	0.0	0.2	5.81	8.28	10.13	15.89	22.96	5.82	6.07	6.88	8.61	11.19
0.8	0.0	0.2	6.03	8.07	9.96	15.05	19.16	5.24	5.63	6.50	8.11	10.51
1.0	0.0	0.2	5.91	7.93	9.60	13.74	16.45	5.50	5.63	6.65	7.99	10.01
0.0	0.0	0.3	6.37	9.32	10.50	17.45	29.16	5.82	6.66	6.78	9.81	14.83
0.2	0.0	0.3	7.03	9.86	10.97	17.22	28.61	5.42	6.44	6.79	9.37	14.01
0.5	0.0	0.3	6.58	9.46	11.00	17.50	28.58	5.40	6.23	6.75	9.23	12.39
0.8	0.0	0.3	6.54	8.90	10.81	16.47	24.34	5.67	6.11	6.91	9.00	11.27
1.0	0.0	0.3	6.27	8.15	10.57	15.51	20.76	5.37	6.21	6.66	8.41	10.32
0.0	0.1	0.1	5.82	7.91	9.04	15.22	21.93	5.49	6.18	6.31	8.21	11.63
0.2	0.1	0.1	5.50	7.39	9.03	14.71	21.11	5.15	5.79	6.43	8.09	11.34
0.5	0.1	0.1	6.07	8.00	9.55	14.56	19.38	5.35	5.97	6.47	8.15	10.52
0.8	0.1	0.1	5.63	7.26	8.80	12.71	15.54	5.15	5.38	6.29	7.57	10.06
1.0	0.1	0.1	5.69	7.57	9.19	12.96	15.37	5.32	5.27	6.58	7.96	9.62
0.0	0.1	0.2	6.23	8.43	9.33	15.85	24.62	5.21	6.13	6.33	8.57	13.09
0.2	0.1	0.2	6.59	8.90	10.56	16.29	25.62	5.35	6.22	6.53	8.65	12.53
0.5	0.1	0.2	5.92	7.81	9.78	15.43	22.64	5.71	6.33	7.08	9.29	11.49
0.8	0.1	0.2	5.92	8.43	9.87	13.91	17.27	5.55	6.49	6.78	8.38	10.89
1.0	0.1	0.2	6.06	7.67	9.22	12.71	15.45	5.39	6.13	6.53	8.07	10.53
0.0	0.1	0.3	6.13	8.91	9.70	16.25	28.07	5.93	6.83	6.83	10.04	15.04
0.2	0.1	0.3	6.56	9.13	10.43	16.70	28.76	5.44	6.54	6.98	9.79	13.93
0.5	0.1	0.3	6.18	8.39	10.35	16.50	26.65	5.58	6.45	7.16	9.62	12.90
0.8	0.1	0.3	6.42	8.52	10.73	15.89	21.94	5.58	6.45	7.14	9.22	11.81
1.0	0.1	0.3	6.84	8.71	10.98	15.24	18.94	5.27	6.36	6.60	8.55	10.72

Table 3. Empirical size (%) for tests of heteroscedasticity, $H_0 : \gamma = \gamma_0$, at $\alpha = 5\%$, without regressor transformation, with $\sigma = 0.1$

		$n = 30$			$n = 80$		
λ	γ_0	$\mathrm{LM_E}(\lambda)$	$\mathrm{LM_E}(\hat{\lambda})$	LR	$\mathrm{LM_E}(\lambda)$	$\mathrm{LM_E}(\hat{\lambda})$	LR
A: Box-Cox transformation							
0.0	0.0	4.00	3.89	8.18	4.76	4.58	6.87
0.2	0.0	3.87	3.51	6.83	4.66	4.81	5.71
0.5	0.0	3.92	3.77	7.34	4.51	4.48	5.36
0.8	0.0	4.43	4.29	8.14	4.81	4.76	5.58
1.0	0.0	4.28	4.17	7.59	4.48	4.47	5.39
0.0	0.1	4.07	3.98	8.08	4.88	4.94	7.40
0.2	0.1	4.25	4.18	7.46	4.77	4.75	5.68
0.5	0.1	4.21	4.04	7.08	4.44	4.48	5.26
0.8	0.1	3.95	3.92	7.33	4.40	4.40	5.04
1.0	0.1	4.37	4.09	7.50	4.59	4.54	5.49
0.0	0.2	4.31	4.13	8.06	4.61	4.63	6.50
0.2	0.2	4.35	4.00	7.27	4.94	4.91	5.64
0.5	0.2	4.16	3.77	7.01	4.76	4.96	5.67
0.8	0.2	4.35	4.21	7.27	4.42	4.46	5.44
1.0	0.2	3.79	3.63	7.03	4.64	4.97	5.65
B: Dual-power transformation							
0.0	0.0	4.00	4.06	6.66	4.76	4.64	5.45
0.2	0.0	3.87	3.51	6.84	4.66	4.83	5.68
0.5	0.0	3.92	3.78	7.35	4.51	4.49	5.37
0.8	0.0	4.43	4.29	8.16	4.81	4.76	5.58
1.0	0.0	4.28	4.17	7.59	4.48	4.47	5.40
0.0	0.1	4.04	3.85	7.18	4.83	4.83	5.88
0.2	0.1	4.41	4.27	7.14	4.75	4.87	5.61
0.5	0.1	4.40	4.36	7.15	4.61	4.48	5.37
0.8	0.1	4.04	4.14	7.10	4.40	4.49	5.13
1.0	0.1	4.43	4.33	7.36	4.63	4.65	5.57
0.0	0.2	4.50	4.35	7.02	4.62	4.78	5.51
0.2	0.2	4.69	4.64	7.40	4.91	4.97	5.77
0.5	0.2	4.22	4.46	7.17	4.94	5.09	5.81
0.8	0.2	4.63	4.83	7.42	4.46	4.63	5.43
1.0	0.2	4.12	4.29	6.99	4.79	4.95	5.71

Table 4. Empirical size (%) for tests of heteroscedasticity, $H_0 : \gamma = \gamma_0$, at $\alpha = 5\%$, with regressor transformation, with $\sigma = 0.1$

			$n = 30$			$n = 80$		
λ	γ_{10}	γ_{20}	$\mathrm{LM_E}(\lambda)$	$\mathrm{LM_E}(\hat{\lambda})$	LR	$\mathrm{LM_E}(\lambda)$	$\mathrm{LM_E}(\hat{\lambda})$	LR
A: Box-Cox transformation								
0.0	0.0	0.0	4.09	4.30	10.19	4.60	4.67	6.58
0.2	0.0	0.0	4.40	4.58	11.28	4.27	4.43	6.65
0.5	0.0	0.0	4.42	4.65	11.04	4.37	4.43	6.19
0.8	0.0	0.0	4.22	4.43	10.35	4.47	4.41	6.33
1.0	0.0	0.0	4.02	4.50	10.31	4.57	4.58	6.64
0.0	0.0	0.2	4.31	5.86	11.96	5.45	5.73	8.08
0.2	0.0	0.2	4.87	5.98	12.45	4.91	5.37	7.10
0.5	0.0	0.2	4.53	5.43	12.40	4.85	4.98	6.41
0.8	0.0	0.2	4.72	5.41	12.29	4.79	4.91	6.81
1.0	0.0	0.2	4.13	4.67	11.48	4.63	5.00	7.02
0.0	0.2	0.0	4.65	5.06	10.14	5.16	4.98	7.30
0.2	0.2	0.0	4.34	4.89	11.78	4.96	4.98	7.05
0.5	0.2	0.0	4.42	4.66	11.34	5.34	5.31	7.04
0.8	0.2	0.0	4.11	4.86	11.31	4.65	4.74	6.84
1.0	0.2	0.0	4.58	4.95	11.91	4.62	4.55	6.66
0.0	0.2	0.2	5.57	6.13	12.25	5.13	5.39	7.70
0.2	0.2	0.2	5.21	6.23	13.29	4.81	5.29	7.34
0.5	0.2	0.2	5.35	5.97	13.87	5.01	5.47	7.47
0.8	0.2	0.2	5.06	5.71	13.15	5.10	5.48	7.40
1.0	0.2	0.2	4.90	5.50	12.97	5.12	5.27	7.32
B: Dual-power transformation								
0.0	0.0	0.0	3.60	3.85	10.28	4.75	4.68	5.63
0.2	0.0	0.0	3.96	4.32	11.05	4.50	4.68	6.57
0.5	0.0	0.0	4.37	4.68	12.56	4.76	5.01	6.80
0.8	0.0	0.0	3.77	4.42	12.02	5.02	5.30	6.64
1.0	0.0	0.0	4.04	4.59	11.89	4.91	5.10	6.73
0.0	0.0	0.2	5.00	5.77	11.13	4.63	4.65	6.66
0.2	0.0	0.2	5.10	6.43	12.33	5.05	5.48	7.22
0.5	0.0	0.2	5.81	6.72	14.16	5.22	5.83	7.85
0.8	0.0	0.2	5.90	6.66	14.58	5.40	6.04	7.63
1.0	0.0	0.2	5.59	5.85	13.69	5.23	5.74	7.59
0.0	0.2	0.0	4.14	4.16	9.21	4.92	5.04	6.25
0.2	0.2	0.0	4.83	4.95	10.34	5.01	5.04	6.62
0.5	0.2	0.0	4.71	4.99	10.57	4.79	5.09	6.74
0.8	0.2	0.0	4.20	4.84	10.47	4.92	5.28	6.45
1.0	0.2	0.0	4.35	4.98	10.70	4.87	5.06	6.92
0.0	0.2	0.2	5.52	5.82	9.63	5.21	5.84	6.76
0.2	0.2	0.2	5.84	6.58	11.24	5.46	5.87	7.19
0.5	0.2	0.2	5.64	6.40	11.44	5.38	5.94	7.44
0.8	0.2	0.2	5.53	6.79	10.76	5.38	5.83	7.42
1.0	0.2	0.2	5.37	6.54	10.90	5.60	6.11	7.48