ELSEVIER

Tests of transformation in nonlinear regression

Zhenlin Yang[a],*, Gamai Chen[b]

[a]*School of Economics and Social Sciences, Singapore Management University, Federal Building,
469 Bukit Timah Rd, Singapore 259756, Singapore*
[b]*Department of Mathematics and Statistics, University of Calgary, Canada T2N 1N4*

Abstract

This paper presents three versions of the Lagrange multiplier (LM) tests of transformation in nonlinear regression: (i) LM test based on expected information, (ii) LM test based on Hessian, and (iii) the LM test based on gradient. All three tests can be easily implemented through a nonlinear least squares procedure. Simulation results show that, in terms of finite sample performance, the LM test based on expected information is the best, followed by the LM test based on Hessian and then the LM test based on gradient. The LM test based on gradient can perform rather poorly. An example is given for illustration.
© 2004 Elsevier B.V. All rights reserved.

1. Introduction

In many econometric applications, a theoretical relationship between the endogenous and exogenous variables exists, but the form in which the error enters the model is not clear. In testing the form of error specification in nonlinear regression, Leech (1975) incorporated the Box–Cox transformation (Box and Cox, 1964): $h(y,\lambda)=(y^{\lambda}-1)/\lambda$ if $\lambda \neq 0$; $\log y$ if $\lambda=0$, to both the response and the systematic part of the model to give a transformed nonlinear model:

$$h(y_i, \lambda) = h[f(x_i, \beta), \lambda] + \sigma e_i, i = 1, \ldots, n \tag{1}$$

where f is a known nonlinear function characterizing the economic relationship between the endogenous

---

* Corresponding author. Tel.: +65-6822-0852; fax: +65-6822-0833.
E-mail address: zlyang@smu.edu.sg (Z. Yang).

variable y_i and the exogenous variables x_i ($m \times 1$), β is a $p \times 1$ model parameter vector, σ is the error standard deviation and σe_i are the independent $N(0, \sigma^2)$ errors.[1]

Model (1) includes the model with additive normal errors ($\lambda = 1$) and the model with multiplicative lognormal errors ($\lambda = 0$) as special cases. In identifying the plausible transformations, Leech used the likelihood ratio test. The likelihood ratio test requires calculations of both restricted and unrestricted maximum likelihood estimates (MLE). A closely related test, the Wald test, also needs the unrestricted MLEs. Clearly, for the type of models defined in Eq. (1), calculation of unrestricted MLEs can be difficult and inconvenient. However, when λ is given, model (1) reduces to a regular nonlinear model and standard software can be used. This motivates us to consider the Lagrange multiplier (LM) test.

This paper presents three versions of the LM test, namely, the LM test based on expected information, the LM test based on Hessian, and the LM test based on gradient. Finite sample properties of these three tests are investigated and compared through extensive Monte Carlo simulation. The results show that the LM test based on the expected information performs the best in the sense that it has the empirical size the closest to the nominal level, and hence is recommended. Section 2 presents the three LM tests. Section 3 presents Monte Carlo results. Section 4 presents an example, and Section 5 concludes.

2. Lagrange multiplier tests

Denote the original data by $\{y_i, x_i, i = 1, 2, \ldots, n\}$, and let $\psi = (\lambda, \beta^T, \sigma)^T$. The log likelihood $\ell(\psi)$ is proportional to

$$-n\log(\sigma) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} [h(y_i, \lambda) - h(f(x_i, \beta), \lambda)]^2 + (\lambda - 1) \sum_{i=1}^{n} \log y_i. \tag{2}$$

Maximizing $\ell(\psi)$ gives the MLE $\hat{\psi}$ and maximizing $\ell(\psi)$ under certain constraints on ψ gives the restricted MLE $\hat{\psi}_0$. Let $S(\psi)$, $G(\psi)$, $H(\psi)$, and $I(\psi)$ be, respectively, the score vector, gradient matrix, Hessian matrix and expected information matrix. The three versions of the LM test for testing H_0: $\lambda = \lambda_0$ are defined as

$$\text{LM}_E = S_\lambda^2(\hat{\psi}_0) I^{\lambda\lambda}(\hat{\psi}_0), \tag{3}$$

$$\text{LM}_H = -S_\lambda^2(\hat{\psi}_0) H^{\lambda\lambda}(\hat{\psi}_0), \tag{4}$$

$$\text{LM}_G = S_\lambda^2(\hat{\psi}_0) D^{\lambda\lambda}(\hat{\psi}_0), \tag{5}$$

where $\hat{\psi}_0 = (\lambda_0, \hat{\beta}_0^T, \hat{\sigma}_0)^T$, $S_\lambda(\hat{\psi}_0)$ is the first element of $S(\hat{\psi}_0)$, and $I^{\lambda\lambda}(\hat{\psi}_0)$, $H^{\lambda\lambda}(\hat{\psi}_0)$ and $D^{\lambda\lambda}(\hat{\psi}_0)$ are, respectively, the first diagonal element of $I^{-1}(\hat{\psi}_0)$, $H^{-1}(\hat{\psi}_0)$ and $[G^T(\hat{\psi}_0) G(\hat{\psi}_0)]^{-1}$. All these tests are referred to the chi-squared distribution with 1 degree of freedom.[2]

---

[1] The model is called the *transform-both-side* model in the statistical literature. Its theoretical properties can be found in Carroll and Ruppert (1988). Most of the econometric literature on Box–Cox regressions has concerned models in which the endogenous variable and some of the exogenous variables are each subject to a Box–Cox transformation. A test of transformation for this model is given in Yang and Abeysinghe (2003).

[2] Godfrey (1988, p. 15) pointed out that the LM test can be sensitive to the way in which the expected information is estimated. Davidson and MacKinnon (1983) and Bera and MacKenzie (1986) argued that the estimated information matrix should be as non-stochastic as possible and preferably depends on the data only through the parameter estimators. See Davidson and MacKinnon (1993) for other versions of LM test.

We now present expressions for the various quantities needed for implementing the three LM tests, among which the expression for $I^{\lambda\lambda}(\psi)$ is the most difficult one to derive. It requires certain approximations in order to obtain explicit expressions for the information matrix. Appendix A gives the detail. Let $f_i = f(x_i,\beta)$, $f_{\beta i} = \partial f(x_i,\beta)/\partial\beta$, and h_λ and $h_{\lambda\lambda}$ be the first- and second-order partial derivatives of of h with respect to λ. We have,

$$
S_\lambda(\hat{\psi})_0 = \sum_{i=1}^n \log y_i - \frac{n \sum_{i=1}^n r_i(\lambda_0,\hat{\beta}_0)s_i(\lambda_0,\hat{\beta}_0)}{\sum_{i=1}^n r_i^2(\lambda_0,\hat{\beta}_0)}
\tag{6}
$$

where $r_i(\lambda,\beta) = h(y_i,\lambda) - h(f_i,\lambda)$, $s_i(\lambda,\beta) = h_\lambda(y_i,\lambda) - h_\lambda(f_i,\lambda)$. Using the explicit expressions for the expected information sub-matrices given in Appendix A, we obtain, after some algebra,

$$
I^{\lambda\lambda}(\psi) = \left(2\Delta_1 - \frac{\sigma^2}{4}\Delta_2 + \frac{7}{4}n\sigma^2\right)^{-1}, \qquad \lambda = 0,
\tag{7}
$$

$$
\approx \left(2\Delta_1 - \frac{\sigma^2}{4}\Delta_2 + \frac{2}{\lambda^2}\Delta_3 - \frac{4}{\lambda}\Delta_4 + \frac{7}{4\lambda^2}n\overline{\theta^2}\right)^{-1}, \qquad \lambda \neq 0,
\tag{8}
$$

where $\phi_i = \log f_i$, $\bar{\phi} = (\Sigma\phi_i)/n$, $\theta_i = \lambda\sigma/[1 + \lambda h(f_i,\lambda)] = \lambda\sigma f_i^{-\lambda}$, $\Delta_1 = \Sigma(\phi_i - \bar{\phi})^2$, $\Delta_2 = [\Sigma(1+\theta_i^2)f_i^{-1}f_{\beta i}^T][\Sigma f_i^{2(\lambda-1)}f_{\beta i}f_{\beta i}^T]^{-1}[\Sigma(1+\theta_i^2)f_i^{-1}f_{\beta i}]$, which reduces to $(\Sigma f_i^{-1}f_{\beta i}^T)(\Sigma f_i^{-2}f_{\beta i}f_{\beta i}^T)^{-1}(\Sigma f_i^{-1}f_{\beta i})$ when $\lambda = 0$, $\Delta_3 = \sum(\theta_i^2 - \overline{\theta^2})^2$ with $\overline{\theta^2} = (\sum\theta_i^2)/n$, and $\Delta_4 = \sum(\theta_i^2 - \overline{\theta^2})(\phi_i - \bar{\phi})$.

Direct calculation of the Hessian is complicated. Following an approximation procedure suggested by Carroll and Ruppert (1988, p. 129), we have

$$
H^{\lambda\lambda}(\psi) = \left(\frac{\partial^2\ell_p(\lambda)}{\partial\lambda^2}\right)^{-1} \approx \left(\frac{\ell_p(\lambda+\epsilon) + \ell_p(\lambda-\epsilon) - 2\ell_p(\lambda)}{\epsilon^2}\right)^{-1},
\tag{9}
$$

where $\ell_p(\lambda)$ is the profile likelihood of λ, and ϵ is a small number usually taken to be 0.01[3]. Finally, the $G(\psi)$ matrix has the expression,

$$
G(\psi) = \left\{\log y_i - \frac{1}{\sigma^2}r_i(\lambda,\beta)s_i(\lambda,\beta), \frac{1}{\sigma^2}r_i(\lambda,\beta)f_i^{\lambda-1}f_{\beta i}^T, \frac{1}{\sigma^3}r_i^2(\lambda,\beta) - \frac{1}{\sigma}\right\}_{n\times(p+2)}.
\tag{10}
$$

3. A performance study

In this section, we carry out some Monte Carlo experiments to investigate the finite sample properties of the three LM tests. We focus on the sizes, the null distributions, and the powers of the tests. The following model is used to generate simulation data:

$$
h(y_i,\lambda) = h(\beta_1 + \beta_2 x_i,\lambda) + \sigma e_i, \quad i = 1,2,\cdots,n,
$$

---

[3] This estimate is easy to calculate. However, there is a major drawback of this method: positivity of the estimate is not guaranteed, especially when the null value λ_0 is far from $\hat{\lambda}$, the MLE of λ (Lawrance, 1987).

where the log x_i values are chosen uniformly from (0,6). The simulation process is as below. For a given parameter configuration $\{\beta_1, \beta_2, \lambda, \sigma\}$, generate data $\{y_i, i = 1, \ldots, n\}$ from

$$y_i = \{1 + \lambda[h(\beta_1 + \beta_2 x_i, \lambda) + \sigma e_i]\}^{1/\lambda} = [(\beta_1 + \beta_2 x_i)^\lambda + \lambda \sigma e_i]^{1/\lambda},$$

if $\lambda \neq 0$, and from $y_i = (\beta_1 + \beta_2 x_i)\exp(\sigma e_i)$ if $\lambda = 0$, where $\{e_i, i = 1, \ldots, n\}$ is a simple random sample from the standard normal population. Under a given $H_0: \lambda = \lambda_0$, run a nonlinear least square with $h(y_i, \lambda_0)$

Table 1
Empirical size (%) for the tests of transformation

n	σ	λ	$\alpha = 10\%$			$\alpha = 5\%$		
			LM_E	LM_H	LM_G	LM_E	LM_H	LM_G
25	0.01	−0.50	9.73	14.26	25.19	4.89	8.08	16.73
	0.01	−0.25	9.47	13.66	25.61	4.85	8.36	16.46
	0.01	−0.10	9.73	14.29	25.59	4.66	8.44	16.86
	0.01	0.00	9.82	14.60	26.57	4.73	8.80	17.63
	0.01	0.10	9.32	13.99	25.62	4.38	8.60	17.29
	0.01	0.25	9.44	14.61	26.94	4.57	8.77	17.89
	0.01	0.50	9.06	14.20	26.66	4.11	8.54	17.77
	0.10	−0.25	9.10	13.34	25.03	4.54	7.60	16.13
	0.10	0.00	9.24	14.11	25.86	4.24	8.36	16.92
	0.10	0.25	9.27	13.93	26.52	4.25	8.45	17.42
	0.30	0.00	9.25	13.97	25.85	4.59	8.36	17.13
	0.30	0.25	9.39	14.60	27.15	4.14	8.61	18.10
50	0.01	−0.50	10.13	12.05	18.69	5.30	6.59	12.07
	0.01	−0.25	10.41	12.07	19.49	4.79	6.64	12.02
	0.01	−0.10	9.98	12.34	19.51	4.73	6.68	12.45
	0.01	0.00	10.07	12.08	19.88	4.84	6.67	12.49
	0.01	0.10	9.89	12.24	19.82	4.49	6.47	12.33
	0.01	0.25	9.56	12.12	19.85	4.79	6.86	12.80
	0.01	0.50	9.41	11.66	19.28	4.76	6.63	12.36
	0.10	−0.25	9.61	11.81	19.52	4.82	6.31	12.04
	0.10	0.000	9.80	11.89	19.34	4.85	6.43	12.18
	0.10	0.25	9.99	12.35	20.58	4.64	6.89	12.79
	0.30	0.00	9.47	11.82	19.75	4.80	6.36	12.13
	0.30	0.25	9.78	12.02	19.92	4.57	6.70	12.51
100	0.01	−0.50	10.02	10.96	14.81	4.97	5.65	8.54
	0.01	−0.25	10.14	11.06	15.05	5.18	5.90	8.73
	0.01	−0.10	10.08	10.86	15.55	4.91	5.67	8.85
	0.01	0.00	9.87	11.14	14.96	4.97	5.83	9.05
	0.01	0.10	10.08	11.26	16.07	4.92	5.83	9.60
	0.01	0.25	9.97	11.29	15.72	4.74	5.76	9.38
	0.01	0.50	9.41	10.91	15.31	5.06	5.74	9.31
	0.10	−0.25	9.80	10.79	15.23	4.98	5.72	8.99
	0.10	0.00	9.99	10.93	14.96	4.83	5.84	9.07
	0.10	0.25	10.08	11.22	15.45	5.22	5.95	9.38
	0.30	0.00	9.95	11.06	15.89	5.00	5.86	9.18
	0.30	0.25	9.80	10.74	15.33	4.78	5.74	9.02

as the response and $h(\beta_1 + \beta_2 x_i, \lambda_0)$ as the model to get the restricted MLE's $\hat{\beta}_{10}$, $\hat{\beta}_{20}$, and $\hat{\sigma}_0$. Then, calculate LM_E, LM_H and LM_G. Repeat this process 10,000 times to obtain the sizes, the simulated null distributions, and the powers of the tests. Several combinations of the values of σ and λ are considered, and three different sample sizes are used. The values of β_1 and β_2 are fixed at 8.0 and 2.0. For brevity, we only report a part of the simulation results.

3.1. The sizes of the tests

First, we check the sizes of the three LM tests given in Eqs. (3)–(5). A part of simulation results are summarized in Table 1. The size of the LM_E test is always much closer to the nominal level than the other two tests. The LM_H test performs reasonably when n is not small, whereas the LM_G test performs rather poorly; its size is usually a few times higher than the nominal level.

3.2. The null distributions of the tests

We compare the means, the standard deviations and the quantiles $Q_{0.5}$, $Q_{0.90}$, $Q_{0.95}$ and $Q_{0.99}$ of the simulated distributions for LM_E, LM_H and LM_G with those of χ_1^2, the chi-squared distribution with 1 degree of freedom. Table 2 contains some representative results. It is seen from Table 2 that the null distribution of the LM_E test is much closer to that of χ_1^2 than the other two tests. The null distribution of the LM_G test can be far from that of χ_1^2 even when sample size is as large as 100.

3.3. The powers of the tests

To make a fair comparison, the powers of the tests are simulated using the simulated percentage points given in Table 2 to set up the tests, i.e., the tests are adjusted to have the same size. The data are generated from the model with $\psi=(0, 8, 2, 0.01)$, and the tests of H_0: $\lambda = \lambda_0$, for $\lambda_0 = -0.5$, to 0.5 with an increment of 0.1, are simulated. Our results (not reported for brevity) show that the three size-adjusted tests have comparable powers in general.

Table 2
Summary of simulated null distributions: $\psi=(0, 8, 2, 0.01)$

n	Test	Mean	S.D.	$Q_{0.50}$	$Q_{0.90}$	$Q_{0.95}$	$Q_{0.99}$
	Nominal	1.0000	1.4142	0.4549	2.7055	3.8415	6.6349
25	LM_E	0.9719	1.3596	0.4450	2.6478	3.6111	6.4024
	LM_H	1.3075	2.0428	0.5398	3.4976	5.3514	9.7045
	LM_G	1.9395	2.3676	1.0026	5.3222	6.9796	10.1678
50	LM_E	0.9899	1.4027	0.4536	2.6380	3.7636	6.6621
	LM_H	1.1402	1.7394	0.4979	2.9791	4.4483	8.3776
	LM_G	1.5955	2.2184	0.7303	4.3522	6.2300	10.4329
100	LM_E	0.9948	1.4033	0.4473	2.6968	3.8604	6.5122
	LM_H	1.0620	1.5399	0.4668	2.8744	4.2222	7.2959
	LM_G	1.3335	1.9415	0.5746	3.6500	5.2565	9.3912

4. An example

In formulating the functional form of a production function, one often assumes the constant return to scale and the constant elasticity of substitution (CES). Weitzman (1970) proposed a production function based on these two assumptions and incorporated Hicks neutral technical change over time to give

$$f(\cdot) = \gamma e^{\mu t}[\delta K_t^{-\rho} + (1 - \delta)L_t^{-\rho}]^{-1/\rho}, \tag{11}$$

where K_t is the aggregate capital and L_t the aggregate labor at time t. Weitzman fitted this model assuming a multiplicative and serially independent error term. Leech (1975) relaxed the specification of Eq. (11) to a general Box–Cox form and used the likelihood ratio (LR) test to test whether the error is indeed multiplicative lognormal or additive normal. We now apply the LM tests to the same data set. The results are summarized in Table 3.

From Table 3, we see that at 5% level, all the tests except the LM_E test reject $\lambda_0 = 0.5$. All the tests reject $\lambda_0 = -1.0$, the reciprocal transformation, and all do not reject $\lambda_0 = -0.5$. The LR and LM_E tests do not reject the logarithmic transformation, but the LM_H and LM_G tests do. The huge values of LM_H at $\lambda_0 = 0.5$ and -1.5 are clear indications for the failure of the LM_H test, resulted from the fact that the null value is too far from the MLE. Therefore, the final decision would be that transformation with $\lambda_0 = -0.5$ gives an appropriate functional form.

5. Discussions

Three versions of the Lagrange multiplier tests have been developed for the purpose of testing a transformation in nonlinear regression. These tests are all very easy to implement: standard nonlinear least squares plus some simple arithmetic calculations. Monte Carlo simulation shows that for finite samples, the LM test based on expected information is the most reliable one. The LM test using Hessian can fail when the null value is too far from the MLE, and the distribution of the LM test based on gradient tends to be too far away from the limiting χ_1^2 distribution, rendering it the least useful test. The three tests are implemented using SAS/NLIN and SAS/IML procedures. A Fortran code is also available from the authors.

Table 3
LM tests for Leech's example

λ_0	LR	LM_E	LM_H	LM_G
0.5	9.1074	3.4195	857.2100	8.7726
0.0	2.9040	2.2146	4.5315	7.7870
−0.5	0.0082	0.0103	0.0081	0.0254
−1.0	4.5798	5.8256	5.2785	12.4024
−1.5	15.2220	12.9249	39.9593	14.5788

$n = 20$, $\hat{\lambda} = -0.4773$, $\ell(\hat{\psi}) \propto -10.9621$.

Appendix A. The expected information

The components of the expected Fisher information matrix are:

$$I_{\lambda\lambda} = \frac{1}{\sigma^2} \sum_{i=1}^n E[s_i^2(\lambda, \beta)] + \frac{1}{\sigma} \sum_{i=1}^n E[e_i h_{\lambda\lambda}(y_i, \lambda)],$$

$$I_{\beta\lambda} = -\frac{1}{\sigma^2} \sum_{i=1}^n \{E[s_i(\lambda, \beta)]f_i^{\lambda-1}f_{\beta i}\},$$

$$I_{\beta\beta} = \frac{1}{\sigma^2} \sum_{i=1}^n [f_i^{2(\lambda-1)}f_{\beta i}f_{\beta i}^T],$$

$$I_{\beta\sigma} = 0, \ I_{\sigma\beta}^T = 0, \ I_{\sigma\lambda} = -\frac{2}{\sigma^2} \sum_{i=1}^n E[e_i h_\lambda(y_i, \lambda)], \ I_{\sigma\sigma} = \frac{2n}{\sigma^2}.$$

Clearly, explicit expressions for the expected information are obtainable only when h_λ and $h_{\lambda\lambda}$ can be expressed explicitly in terms of the error elements e_is. We have

$$h_\lambda(y, \lambda) = \frac{1}{\lambda}[1 + \lambda h(y, \lambda)]\log y - \frac{1}{\lambda}h(y, \lambda),$$

and

$$h_{\lambda\lambda}(y, \lambda) = h_\lambda(y, \lambda)\left[\log y - \frac{1}{\lambda}\right] - \frac{1}{\lambda^2}[\log y - h(y, \lambda)],$$

which converge, as $\lambda \to 0$, to $(1/2)(\log y)^2$ and $(1/3)(\log y)^3$, respectively. Thus, the problem reduces to expressing $\log y_i$ explicitly in terms of e_is. When $\lambda = 0$, $\log y_i = \log f_i + \sigma e$, and some simple algebra leads immediately to

$$I_{\lambda\lambda} = \frac{7}{4}n\sigma^2 + 2\sum_{i=1}^n \phi_i^2, \ I_{\beta\lambda} = -\frac{1}{2}\sum_{i=1}^n f_i^{-1}f_{\beta i}, \ I_{\sigma\lambda} = -\frac{2}{\sigma}\sum_{i=1}^n \phi_i,$$

where $\phi_i = \log f_i$. When $\lambda \neq 0$, $\log y_i = (1/\lambda)\log[1 + \lambda(h(f_i, \lambda) + \sigma e_i)]$, which is a nonlinear function of e_i. Thus, an approximation is necessary. Let $\theta_i = \lambda\sigma/[1 + \lambda h(f_i, \lambda)] = \lambda\sigma f_i^{-\lambda}$. A Taylor series expansion gives

$$\log y_i = \phi_i + \frac{1}{\lambda}\left[\theta_i e_i - \frac{1}{2}(\theta_i e_i)^2 + \cdots + \frac{1}{k!}(-1)^{k-1}(\theta_i e_i)^k + \cdots\right].$$

Since y_is are nonnegative, it has to be that $P(f_i^\lambda + \lambda\sigma e_i < 0)$ be negligible, which is equivalent to $\theta_i \ll 1$. Hence, it is sufficient to keep the terms in the above expansion up to the third-order. Some tedious but straightforward algebra leads to

$$I_{\lambda\lambda} \approx \sum_{i=1}^n \left(\frac{7}{4\lambda^2}\theta_i^2 - \frac{4}{\lambda}\theta_i^2\phi_i + 2\phi_i^2\right), \ I_{\beta\lambda} \approx -\frac{1}{2}\sum_{i=1}^n (1 + \theta_i)\frac{f_{\beta i}}{f_i}, \ I_{\sigma\lambda} \approx \frac{2}{\lambda\sigma}\sum_{i=1}^n (\theta_i^2 - \lambda\phi_i).$$

References

Bera, A.K., MacKenzie, C.R., 1986. Alternative forms and properties of the Lagrange multiplier test. Journal of Applied Statistics 13, 13–25.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations (with discussion). Journal of the Royal Statistical Society. Series B. 26, 211–252.

Carroll, R.L., Ruppert, D., 1988. Transformation and Weighting in Regression. Chapman & Hall, New York.

Davidson, R., MacKinnon, J.G., 1983. Small sample properties of alternative forms of the Lagrange multiplier test. Economics Letters 12, 269–275.

Davidson, R., MacKinnon, J.G., 1993. Estimation and Inference in Econometrics. Oxford Univ. Press, New York.

Godfrey, L.G., 1988. Misspecification Tests in Econometrics. Cambridge Univ. Press, Cambridge.

Lawrance, A.J., 1987. The score statistic for regression transformation. Biometrika 74, 275–279.

Leech, D., 1975. Testing the error specification in nonlinear regression. Econometrika 43, 719–725.

Weitzman, M.L., 1970. Soviet postwar economic growth and capital-labor substitution. American Economic Review 60, 676–692.

Yang, Z.L., Abeysinghe, T., 2003. A score test for Box–Cox functional form. Economics Letters 79, 107–115.