A robust test for linear and log-linear models against Box-Cox alternatives

David Vincent

September 2023

2023 UK Stata Conference

Contents

Motivation

- In linear regression analysis, it is common to consider whether the dependent variable should be levels or in logarithms
- Except for some obvious cases, such as negative values, both models might seem to be appropriate, so it is usual for practitioners to make decisions based on statistical procedures
- Plots of the residuals vs fitted values and specification tests such as RESET can be helpful in detecting non-linearities, but these methods are not designed to test against any specific alternative and can lack power
- Another approach that is often implemented, is to estimate a model that allows for more general non-linearities, but which nests the linear and log-linear models as special cases

Motivation

- The Box-Cox transformation, first introduced by Box and Cox (1964), is the obvious choice and is estimated by maximum likelihood assuming normally distributed and homoskedastic errors. This can be implemented in Stata for cross sectional analysis using the `boxcox` command
- Wald or likelihood-ratio tests of the null that the regression is linear or log-linear, against the alternative of a Box-Cox regression, can then be carried out by testing if the transformation parameter, denoted λ is one or zero
- Lagrange Multiplier (LM) tests can also be computed and are generally easier to implement, as they only require estimating the linear or log-linear model under the null hypothesis

Motivation

- Godfrey and Wickens (1981) show that the test-statistic can be calculated using artificial regressions and Davidson and MacKinnon (1985) consider an extension of the LM-test that has better small sample properties
- Despite the popularity of these methods, they are not robust to departures from the distributional assumptions of the errors, especially heteroskedasticity
- Zarembka (1974) shows that the estimator is not just restricted to the search for non-linearity, but also to one that leads to more normal errors, with constant variance
- This reflects the original intention of Box and Cox (1964) which was to find a transformation of the data to make it closer to linearity, normality and homoskedasticity

Motivation

- This makes the maximum likelihood based tests less useful for functional form testing, as the results could favour log-linearity when the true model is linear with non-normal or heteroskedastic errors
- The purpose of this presentation, is to present an LM-test `xtloglin` for linear and log-linear models against Box-Cox alternatives which is robust to arbitrary error-distributions
- The test can be implemented after `regress` and `xtreg` and is based on a GMM-estimator of the Box-Cox model first proposed by Amemiya and Powell (1981)
- This extends the analysis by Savin and Wurtz (2005), to additionally allow for heteroskedasticity, clustering and longitudinal datasets

Box Cox Model

- Let $y_i > 0$ denoted the i-th value of a non-negative dependent variable and x_i a K-vector of exogenous explanatory variables
- The Box-Cox model assumes that for some value of the parameter λ, the transformed dependent variable $T(y_i, \lambda)$ is a linear function of the explanatory variables:

$$T(y_i, \lambda) = x_i^{\prime}\beta + \epsilon_i \tag{1}$$

where β is a vector of coefficients and ϵ_i is the error term and the Box-Cox transformation is defined as:

$$T(y_i, \lambda) = \begin{cases} (y_i^{\lambda} - 1)/\lambda & \text{if } \lambda \neq 0 \text{ Box-Cox} \\ y_i - 1 & \text{if } \lambda = 1 \text{ Linear} \\ \log y_i & \text{if } \lambda = 0 \text{ Log-linear.} \end{cases} \tag{2}$$

GMM estimation

- Let z_i denote a $R > K$ set of instrumental variables, that yield the moment conditions $E[z_i v_i] = E[g_i(\theta)] = 0$
- In matrix form with N observations, the model in (1) is $T(y, \lambda) = X\beta + v$ and the the sample moment conditions are:

$$\bar{g}(\theta) = \frac{1}{N} Z' v(\theta) \tag{3}$$

- As setting $\lambda = -\infty$ and $\beta = 0$ sets the residuals to zero when $y_i > 1$, the sample moments are rescaled:

$$\bar{g}^*(\theta) = \frac{1}{N} Z' v(\theta) \dot{y}^{-\lambda} \tag{4}$$

- This is suggested by Powell (1996) where $\dot{y}$ is the geometric mean of the absolute values of the dependent variable and helps to ensure the estimator is well behaved

GMM estimation

- Let W_N denote a positive definite weighting matrix, the generalized method of moments (GMM) estimator finds the parameters $\theta = (\beta', \lambda)'$ that minimise the objective function:

$$Q_N = \bar{g}^{*'}(\theta) \; W_N \; \bar{g}^{*}(\theta) \tag{5}$$

- Differentiating the objective function Q_N with respect to θ yields the first order conditions to be solved:

$$\frac{\partial Q_N}{\partial \theta} = \left(\frac{\partial \bar{g}^{*}(\theta)}{\partial \theta'} \right)' W_N \bar{g}^{*}(\theta) = 0 \tag{6}$$

GMM estimation

- Ignoring irrelevant constants, the first order conditions are:

$$\frac{\partial Q_N}{\partial \beta} = X' Z W_N Z' \left[T(y,\lambda) - X\beta \right] = 0 \quad (7)$$

$$\frac{\partial Q_N}{\partial \lambda} = T_\lambda^{*\prime} Z W_N Z' \left[T(y,\lambda) - X\beta \right] = 0 \quad (8)$$

- The term T_λ^* is the derivative of the scaled Box-Cox transformation with respect to λ, which for observation i is:

$$T_{\lambda,i}^* = T_{\lambda,i} - v_i \log \dot{y} \quad (9)$$

and where

$$T_{\lambda,i} = \begin{cases} \frac{\left(y_i^\lambda [\lambda \log y_i - 1] + 1\right)}{\lambda^2} & \text{if } \lambda \neq 0 \\ \frac{(\log y_i)^2}{2} & \text{if } \lambda = 0. \end{cases} \quad (10)$$

GMM estimation

- The estimates for a given sample can be found by first solving the first order condition (7) for β:

$$\widehat{\beta}(\lambda) = (X'ZWZ'X)^{-1}X'ZWZ'T(y,\lambda) \qquad (11)$$

and then concentrating these out of the first order condition for λ. The search is then over a single parameter $\widehat{\lambda}$ that solves:

$$T^{*'}_{\widehat{\lambda}} ZW_N Z' \left[T(y,\widehat{\lambda}) - X\widehat{\beta}(\widehat{\lambda}) \right] = 0 \qquad (12)$$

- Note that the scaling factor cancels in the estimator of β and only impacts λ via its effect on the derivative T^{*}_{λ}

LM-test procedure

- We want to test the null $H_0 : \lambda = \lambda_r$ against the alternative $H_1 : \lambda \neq \lambda_r$, where $\lambda_r = 1$ is for testing the suitability of the linear model and $\lambda_r = 0$ for the log-linear specification
- Consider estimating the model imposing the null restriction. From the above, this simply requires setting $\lambda = \lambda_r$ and computing the restricted GMM estimates of β from (11) denoted $\tilde{\beta}_r(\lambda_r)$
- If this restriction is true, $\widehat{\lambda} \approx \lambda_r$ and the gradient $\partial Q_N / \partial \lambda$ in (12) evaluated at the restricted estimates $\tilde{\theta}_r = (\tilde{\beta}_r, \lambda_r)$ will be close to zero. This is the basis of the LM test

LM-test procedure

- For the the LM-test to be useful after least squares, the restricted estimates $\tilde{\beta}_r(\lambda_r)$ must be the same as the OLS estimates for the model being investigated
- In other words, if one has estimated the log-linear model $\log y = x_i^{'} + v_i$ say using `regress`, the GMM-estimates must be identical to the OLS estimates when $\lambda_r = 1$
- This can be achieved by setting the weighting matrix to the usual IV-matrix $W_N = (Z^{'}ZN^{-1})^{-1}$ as then:

$$\tilde{\beta}_r = (X^{'}P_Z X)^{-1} X^{'} P_Z^{'} T(y, \lambda_r) \tag{13}$$

where P_z is the projector matrix. Because X is part of the instrument set $P_z X = X$ and (13) is the OLS estimator

LM-test procedure

- The the gradient $\partial Q_N / \partial \lambda$ in (12) evaluated at $\tilde{\theta}_r = (\tilde{\beta}_r, \lambda_r)$, now simplifies to:

$$\frac{\partial Q_N}{\partial \lambda_r} = \frac{1}{N} \tilde{v}_r' P_z T^*_{\lambda_r} \tag{14}$$

where $\tilde{v}_r$ are the OLS residuals from the null model and $P_z T^*_{\lambda_r}$ are the predicted values from an regression of the derivative of the scaled transformation $T^*_{\lambda_r,i}$ on z_i,

- The LM test is therefore a test of no correlation between the OLS residuals and the predicted values of the Box-Cox transformation under the null
- Note that although one estimates a model using y rather than $y - 1$ which is the correct transformation when $\lambda = 1$, this only shifts the constant and has no impact of the residuals

Instruments

- The above makes it clear that suitable instruments for λ must uncorrelated with the errors but correlated with the derivative of the Box-Cox transformation $T^*_{\lambda_r,i}$
- Following the logic of the RESET test, I include the second, third and fourth powers of fitted OLS values $\tilde{\mu}_i = x_i' \tilde{\beta}_r$ from the null model being tested as additional instruments in Z
- The predicted values of the derivative of the scaled Box-Cox transformation $T^*_{\lambda_r,i}$ from an OLS regression on z_i are then:

$$\widehat{T}^*_{\lambda_r,i} = x_i' \widehat{\gamma} + \widehat{\phi}_2 \tilde{\mu}_i^2 + \widehat{\phi}_3 \tilde{\mu}_i^3 + \widehat{\phi}_4 \tilde{\mu}_i^4 \qquad (15)$$

- Thus testing a linear or log-linear model against the alternative $\lambda \neq \lambda_r$ is then a test of no correlation between the OLS residuals and a single linear combination of all three powers

Comparisons with the RESET test

- Before deriving the LM test statistic, it is informative to note that this approach will generally be more powerful than the RESET test, as it takes into account the functional form of the transformation being tested
- To illustrate, let the errors be distributed as $v_i \sim iidN(0, \sigma^2)$ and consider a test of the log-linear model, where $\lambda_r = 0$
- Under this null, the expected value of the derivative of the scaled Box-Cox transformation $T^*_{\lambda_r=0,i}$ is

$$E[T^*_{\lambda_r=0,i}|x_i] = \frac{1}{2}\left[(x_i'\beta)^2 + \sigma_v^2\right] - x_i'\beta \log \dot{y} \qquad (16)$$

- As the expectation depends linearly on $(x_i'\beta)^2$, the true coefficients in the above regression are $\phi_2 = 1/2$, and $\phi_3 = \phi_4 = 0$ as the third and fourth powers have no effect

Comparisons with the RESET test

- Thus under the null, the plim of the gradient is:

$$\operatorname*{plim}_{N\to\infty} \frac{\partial Q_N(\tilde{\theta}_r)}{\partial \lambda_r} = \operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} v_i (x_i' \beta)^2 = 0$$

- By contrast, the RESET test using the same three-powers of the fitted values tests three separate constraints:

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} v_i (x_i' \beta)^m = 0, \quad \text{where } m = 2, .., 4 \qquad (17)$$

- Thus when testing the null of log-linearity, and when $v_i \sim iidN(0, \sigma^2)$ including the third and fourth powers of the fitted values in the RESET test will reduce its asymptotic power when the data are generated by a Box-Cox model

LM test-statistic

- To compute the LM test-statistic requires the limiting distribution of the gradient $\partial Q_N(\tilde{\theta}_r)/\partial\lambda_r$ in (14). Under the null $H_0 : \lambda = \lambda_r$, it can be shown that this is:

$$\frac{1}{\sqrt{N}}\tilde{v}_r' P_z T^*_{\lambda_r} \xrightarrow{d} N(0, V) \tag{18}$$

where the limiting variance is:

$$V = \lim_{N\to\infty}\frac{1}{N}E\left[T^{*'}_{\lambda_r}(P_z - P_x)\Omega(P_z - P_x)T^*_{\lambda_r}\right] \tag{19}$$

- The terms P_z and P_x are the projector matrices and $\Omega = E[vv^{'}|Z]$ is the covariance matrix of error terms

LM test-statistic

- From the above, the LM test-statistic is given by:

$$LM = N^{-1}(\tilde{v}_r' P_z T^*_{\lambda_r})\tilde{V}_r^{-1}(T'^*_{\lambda_r} P_z \tilde{v}_r) \xrightarrow{d} \chi^2_1 \qquad (20)$$

where $\tilde{V}_r$ is a consistent estimator of V under the null:

$$\tilde{V}_r = \frac{1}{N}\left[T^{*\prime}_{\lambda_r}(P_z - P_x)\tilde{\Omega}(P_z - P_x)T^*_{\lambda_r}\right] \qquad (21)$$

- Three estimators of Ω are available depending on the assumptions made about the error terms:

$$\tilde{\Omega} = \begin{cases} \frac{1}{N}\tilde{v}_r'\tilde{v}_r \times I_N & \text{if homoskedastic} \\ \text{diag}\{\tilde{v}_{r1}...\tilde{v}_{rN}\} & \text{if heteroskedastic} \\ \text{diag}\{\tilde{\Omega}_1...\tilde{\Omega}_C\} & \text{if clustering } \tilde{\Omega}_c = \tilde{v}_c\tilde{v}_c'. \end{cases} \qquad (22)$$

Savin and Wurtz (2005)

- It is therefore possible to make this version of the LM-test robust to arbitrary heteroskedasticity and clustering of the error covariance matrix
- This is not the case in Savin and Wurtz (2005) which use the formula developed by Newey and West (1987) for the optimal GMM-estimator which sets the weighting matrix $W = S^{-1}$, where $S = \lim N^{-1}E[Z^{'}\Omega Z]$
- But as Savin and Wurtz (2005) set $W = S^{-1} = (Z^{'}ZN^{-1})^{-1}$ they are restricting the errors to be homoskedastic $\Omega = \sigma_v^2 I_N$
- Although their test remains robust to non-normality, it is not robust to heteroskedasticity, which is more important cause of inconsistency in maximum likelihood based methods

Auxiliary regressions

- Matrix operations are required to compute the test statistic that allows for clustering, but for the homoskedastic and heteroskedastic versions, the test statistics can also be computed using auxiliary regressions.
- To see this, note that (20) can be written more compactly as:

$$LM = \tilde{v}_r' D(D'\tilde{\Omega}X)^{-1} D' \tilde{v}_r \tag{23}$$

where

$$D = (P_z - P_x) T^*_{\lambda_r} = \widehat{T}^{*z}_{\lambda_r} - \widehat{T}^{*x}_{\lambda_r i}$$

which is is the difference in the predicted values of $T^*_{\lambda_r}$ from two separate OLS regressions on Z and X

Auxiliary regressions: homoskedastic errors

- For the homoskedastic case $\tilde{\Omega} = \tilde{v}_r' \tilde{v}_r / N$ the LM-statistic is:

$$LM = \frac{N \tilde{v}_r' D (D' X)^{-1} D' \tilde{v}_r}{\tilde{v}_r' \tilde{v}_r} = N \times R^2 \tag{24}$$

which is $N \times R^2$ from a regression of the residuals $\tilde{v}_{ri}$ on D_i:

$$\tilde{v}_r = \delta(\widehat{T}^{*z}_{\lambda_r i} - \widehat{T}^{*x}_{\lambda_r i}) + e_i \tag{25}$$

- As this is the same as regressing the residuals on x_i and $\widehat{T}^{*z}_{\lambda_r i}$, the above is the usual LM-test that $\widehat{T}^{*z}_{\lambda_r i}$ should be excluded as an extra variable in the linear or log-linear model after OLS
- But as $\widehat{T}^{*z}_{\lambda_r i}$ is the predicted value of the Box-Cox derivative which is endogenous, the test is for an omitted endogenous variable where the included variables are exogenous

Auxiliary regressions: homoskedastic errors

Homoskedastic errors

1. Regress the null model y_i or $\log y_i$, on x_i and save the residuals $\tilde{v}_r$ and the 2^{nd},3^{rd} and 4^{th} powers of the fitted values
2. Regress the derivative of the scaled Box-Cox transformation $T^{*}_{\lambda_r i}$ on x_i and the above powers of the fitted values and save the predicted values $\widehat{T}^{*z}_{\lambda_r i}$ from this regression
3. Compute the LM test statistic as $N \times R^2$ from a regression of $\tilde{v}_{ri}$ on x_i and $\widehat{T}^{*z}_{\lambda_r i}$
4. Compare the test statistic to the relevant percentiles of the chi-square distribution with 1.d.f.

Auxiliary regressions: heteroskedastic errors

- For the heteroskedastic case let $D^{\dagger} = D \circ \tilde{v}_r$, with i-th row $D_i^{\dagger} = D_i \tilde{v}_{ri}$ and let e denote an $N \times 1$ vector of ones. The LM-statistic becomes:

$$LM = e^{'} D^{\dagger} (D^{\dagger'} D^{\dagger})^{-1} D^{\dagger} e \tag{26}$$

which is $N \times R^2$ from a regression of 1 on $D_i^{\dagger}$

$$1 = \delta(\widehat{T}^{*z}_{\lambda_r i} - \widehat{T}^{*x}_{\lambda_r i})\tilde{v}_{ri} + e_i \tag{27}$$

- This test is robust to arbitrary heteroskedasticity, whereas the test used by Savin and Wurtz (2005) is identical to the homoskedastic version described earlier, which will suffer from size distortions when the errors are heteroskedastic

Auxiliary regressions: heteroskedastic errors

Heteroskedastic errors

1. Regress the null model y_i or $\log y_i$, on x_i and save the residuals $\tilde{v}_r$ and the 2^{nd},3^{rd} and 4^{th} powers of the fitted values
2. Regress the derivative of the scaled Box-Cox transformation $T^*_{\lambda_r i}$ on z_i and the above powers of the fitted values and save the predicted values $\widehat{T}^{*z}_{\lambda_r i}$ from this regression
3. Repeat the above regressing $T^*_{\lambda_r i}$ on x_i only and save the predicted values $\widehat{T}^{*x}_{\lambda_r i}$ from this regression
4. Compute the LM test statistic as $N \times R^2$ from a regression of 1 on $(\widehat{T}^{*z}_{\lambda_r i} - \widehat{T}^{*x}_{\lambda_r i})\tilde{v}_{ri}$, without a constant
5. Compare the test statistic to the relevant percentiles of the chi-square distribution with 1.d.f.

Panel data settings

- The same procedure outlined above can be applied to test the linear and log-linear models in panel-settings where $i = 1, .., N$ subjects are observed over multiple t-periods
- This requires working with the panel-transformed model, which is the model that is to be estimated by least squares
- Letting Q_i denote the relevant transformation matrix for subject i, which can be the within, random-effects and between-effects transformations, the model to be estimated is:

$$Q_i T(\lambda, y_i) = Q_i X_i \beta + Q_i v_i \tag{28}$$

- As the above makes clear, the transformation is applied to the Box-Cox transformed data. Thus when testing the log-linear model, the transformation is applied to $\log y$

Panel data settings

- The LM-test requires computing the derivative of the Box-Cox transformed dependent variable $Q_i T(\lambda, y_i)$ with respect to λ and evaluating these under the null
- Although these might seem difficult to derive, they are simply the Q_i transformed derivatives using the untransformed model
- This simplicity occurs because the panel-estimators apply a linear transformation of the data and because Q_i does not depend on the parameter λ, hence:

$$\frac{\partial Q_i T(\lambda, y_i)}{\partial \lambda} = Q_i \frac{\partial T(\lambda, y_i)}{\partial \lambda} \tag{29}$$

- The LM-statistic assuming homoskedastic errors after the Q_i transformation can be computed using auxiliary regression assuming. The cluster-robust version requires the main formula set out earlier

Auxiliary regressions: panel data estimators

Homoskedastic errors (panel transformed model)

1. Regress the transformed null model y_i or $\log y_i$, on x_i and save the <u>transformed</u> residuals $\tilde{v}_r$ and the <u>untransformed</u> 2^{nd},3^{rd} and 4^{th} powers of the fitted values
2. Regress the transformed derivative of the scaled Box-Cox transformation $T^*_{\lambda_r it}$ on the transformed x_{it} and transformed powers and save the predictions $\widehat{T}^{*z}_{\lambda_r i}$
3. Repeat the above regressing $T^*_{\lambda_r it}$ on the transformed x_{it} only and save the predictions $\widehat{T}^{*z}_{\lambda_r i}$
4. Compute the LM test statistic as $N \times R^2$ from a regression of the residuals $\tilde{v}_{ri}$ on the transformed x_i and $\widehat{T}^{*z}_{\lambda_r i}$
5. Compare the test statistic to the relevant percentiles of the chi-square distribution with 1.d.f.

xtloglin

Robust LM-test of linear and log-linear models against Box-Cox alternatives after regress and xtreg

`xtloglin, null(model) [ negative notrobust robust cluster(varname) ]`

`null(linear|log)` specifies the null model to be tested

`negative` specified when the dependent variable contains negative values of the dependent variable. Otherwise aborts

`notrobust` computes the LM-test assuming homoskedastic errors

`robust` reports the test allowing for arbitrary heteroskedasticity

`cluster(`*varname*`)` allows for intra-group correlation of the errors

Example after regress

```
. sysuse auto,clear

. *(1) estimate linear model
. qui reg mpg weight length displacement, robust

  *test null of linearity
. xtloglin, null(linear)

Robust LM test of linear and log-linear functional forms
H0: linear functional form
_________________________________________________________

    LM-chi2(1) =      5.308
    Prob > LM  =      0.0212
_________________________________________________________

Error Variance: robust


. *(2) estimate log-linear model
. qui reg l_mpg weight length  displacement, robust

. *test null of log-linearity
. xtloglin, null(log)

Robust LM test of linear and log-linear functional forms
H0: log functional form
_________________________________________________________

    LM-chi2(1) =      0.454
    Prob > LM  =      0.5004
_________________________________________________________

Error Variance: robust
```

Example after xtregress

```
. webuse nlswork
. xtset idcode
. *(1) estimate log-linear model (random effects, CRSE)
. qui xtreg ln_w grade age tenure i.race  south, re robust

. *test null of log-linearity
. xtloglin, null(log)
Robust LM test of linear and log-linear functional forms
H0: log functional form
---------------------------------------------------------

    LM-chi2(1) =      3.408
    Prob > LM  =      0.0649
---------------------------------------------------------
Error Variance: cluster
Cluster Varible: idcode


. *(2) estimate linear model (random effects,CSRE)
. qui xtreg w grade age tenure i.race  south, re robust

. *test null of linearity
. xtloglin, null(linear)
Robust LM test of linear and log-linear functional forms
H0: linear functional form
---------------------------------------------------------

    LM-chi2(1) =     41.652
    Prob > LM  =      0.0000
---------------------------------------------------------
Error Variance: cluster
Cluster Varible: idcode
```

Monte Carlo Experiments

- The size and power of the LM test is investigated by drawing data from the following linear model:

$$\begin{aligned} y_i &= 10 + x_{1i} + x_{2i} + v_i \\ v_i &\sim TN(-2, \infty, 0, 0.5^2) \\ x_i &\sim TN(-2, \infty, 0, 1) \end{aligned}$$

- $N = 100$ and the rejection probabilities are computed using asymptotic critical values for 1,000 replications
- To investigate the power, I use the same process and log-transform y to test the log-linear model
- The above is repeated setting $\sigma_i^2 \propto exp(0.5x_{1i} + 0.5x_{2i})$ to investigate the performance under heteroskedasticity

Empirical rejection probabilities: Size

True Model: Linear - Null Model: Linear

	Error Variance					
	Homoskedastic			Heteroskedastic		
Nominal Rate	1%	5%	10%	1%	5%	10%
Box-Cox GMM (LM)	.007	.047	.101	.022	.077	.152
Box-Cox ML (LR)	.009	.057	.103	.666	.791	.845
RESET (Wald)	.148	.255	.328	.288	.368	.441
RESET (LM)	.001	.031	.079	.005	.042	.093

*All tests except LR are the robust-versions

Empirical rejection probabilities: Power

True Model: Linear - Null Model: Log-linear

	Error Variance					
	Homoskedastic			Heteroskedastic		
Nominal Rate	1%	5%	10%	1%	5%	10%
Box-Cox GMM (LM)	.547	.881	.956	.354	.61	.725
Box-Cox ML (LR)	.94	.983	.993	.168	.308	.394
RESET (Wald)	.881	.95	.969	.952	.981	.99
RESET (LM)	.348	.687	.822	.29	.619	.774

*All tests except LR are the robust-versions

Conclusion

- This presentation has set out a simple LM-test of linear and log-linear models against Box-Cox alternatives based on the GMM framework
- Unlike the maximum likelihood bases tests, it is robust to heteroskedasticity and non-normality and should outperform the RESET test when data is generated by a Box-Cox transformation
- The test is implemented in the Stata command `xtloglin` and is available for cross sectional and panel data models fitted using `regress` and `xtreg`

References

Amemiya, T., and J. L. Powell. 1981. A comparison of the Box-Cox maximum likelihood estimator and the non-linear two-stage least squares estimator. *Journal of Econometrics* 7(3): 351–381.

Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society, Series B* 26: 211–252.

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

Godfrey, L. G., and M. R. Wickens. 1981. Testing linear and log-linear regressions for functional form. *The Review of Economic Studies* 48(3): 487–496.

Newey, W. K., and K. D. West. 1987. Hypothesis testing with efficient method of moments estimation. *International Economic Review* 28: 777–787.

Zarembka, P. 1974. *Frontiers in econometrics*.