Double-length regressions for the Box–Cox difference model with heteroskedasticity or autocorrelation

Badi H. Baltagi∗, Dong Li

∗Department of Economics, Texas A&M University, College Station, TX 77843-4228, USA

Department of Economics, East Carolina University Greenville, NC 27858-4353, USA

Received 30 June 1999; accepted 14 December 1999

Abstract

This paper derives Lagrange multiplier tests based on artificial double length regressions (DLR) to jointly test for differenced linear or loglinear models with no heteroskedasticity or autocorrelation against a more general differenced Box–Cox model with heteroskedasticity or autocorrelation. These tests are easy to implement and are illustrated using an empirical example. © 2000 Elsevier Science S.A. All rights reserved.

Keywords: Box–Cox difference model; Double length regression; Heteroskedasticity; Autocorrelation

JEL classification: C12

1. Introduction

The familiar Box–Cox model in levels is given by

\[ Y_t^{(A)} = \sum_{k=1}^{K} X_{ik}^{(A)} \beta_k + \sum_{s=1}^{S} Z_s \gamma_s + \epsilon_t, \quad t = 1, \ldots, T \]

where

\[ Y_t^{(A)} = \begin{cases} \frac{Y_t^{\lambda} - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log(Y_t) & \text{if } \lambda = 0 \end{cases} \]

Both \( Y_t \) and \( X_{ik} \) are subject to the Box–Cox transformation and are required to take positive values.
only, while the $Z_j$ variables are not subject to the Box--Cox transformation. The $Z_j$s may include dummy variables like sex, race, region, etc. and the intercept. Note that for $\lambda = 1$, Eq. (1) becomes a linear model whereas for $\lambda = 0$ it becomes a loglinear model. Davidson and MacKinnon (1985) derived a double length regression (DLR) test for linear and log-linear regressions against a more general Box--Cox alternative given by (1). They showed that the DLR test performs better than its OPG counterpart in small samples. These DLR tests have been also generalized to account for heteroskedasticity by Tse (1984) and for AR(1) disturbances by Baltagi (1999).

For time series data, Layson and Seaks (1984) noted that the time derivative of (2) is given by
\[
dY_t^{(A)} = \frac{dY_t^{(A)}}{dt} = \frac{dY_t}{dt} = Y_t^{(A-1)} \frac{dY_t}{dt}
\]
with the discrete approximation of Eq. (3) given by
\[
\Delta Y_t^{(A)} = Y_{t-1}^{(A-1)} \Delta Y_t
\]
and $\Delta Y_t = Y_t - Y_{t-1}$. For $\lambda = 1$, $\Delta Y_t^{(A)}$ in Eq. (4) yields the simple first difference of the variables, while for $\lambda = 0$, $\Delta Y_t^{(A)}$ gives $(Y_t - Y_{t-1})/Y_{t-1}$ which is the percentage change or approximately the first difference of the natural log of $Y_t$. Layson and Seaks (1984) then proposed the Box--Cox difference (BCD) model
\[
\Delta Y_t^{(A)} = \sum_{k=1}^{K} \Delta X_{tk}^{(A)} \beta_k + \sum_{s=1}^{S} Z_s \gamma_s + u_t, \quad t = 2, \ldots, T
\]
where the transformation $\Delta Y_t^{(A)}$ is defined by (4), and $\Delta X_{tk}^{(A)}$ is similarly defined. Using matrix notation, the Box--Cox difference model in (5) can be expressed as
\[
\Delta Y_t^{(A)} = \Delta X_t^{(A)} \beta + Z_t \gamma + u_t,
\]
where $\Delta Y_t^{(A)}$ is $(T-1) \times 1$, $\Delta X_t^{(A)}$ is $(T-1) \times K$, $Z$ is $(T-1) \times S$, and $\beta$ and $\gamma$ are $K \times 1$ and $S \times 1$, respectively. $E(u_t^2) = \sigma_u^2$. Once again, the $Z$ variables are not subject to differencing or the percentage change transformation. The BCD transformation allows one to test for first difference or percentage change models against a more general Box--Cox difference model. Layson and Seaks (1984) suggested a likelihood ratio test for $\lambda = 0$ or $\lambda = 1$ while Park (1991) derived the double length regression (DLR) for testing $\lambda = 0$ or 1 against a more general BCD model.

Note that Layson and Seaks (1984) suggested extensions of the BCD model to the case of heteroskedastic or autocorrelated errors. This paper considers the general BCD model in Eq. (6) with $u_t$ allowed to be heteroskedastic with variance $\sigma_u^2 = h(\alpha + W_\delta)$, where the functional form $h$ need not be known and where $W_\delta$ is of dimension $1 \times L$ and may include $Z_t$ and $X_t$, and $\delta$ is $L \times 1$. Alternatively, the error term $u_t$ can follow a stationary first order autoregressive pattern given by $u_t = \rho u_{t-1} + \varepsilon_t$ where $|\rho| < 1$ and $\varepsilon_t \sim i.i.d(0, \sigma^2_\varepsilon)$. This paper derives DLRs for testing $\lambda = 0$ or 1 without heteroskedasticity or autocorrelation against a more general BCD model with heteroskedastic or autocorrelated errors.
2. BCD with AR(1) error

The model is given by (6) with \( u_t = \rho u_{t-1} + \varepsilon_t \) where \( |\rho| < 1 \) and \( \varepsilon_t \sim i.i.d.(0, \sigma^2) \). Denote the \( t \)th row of \( \Delta Y \), \( \Delta X \), and \( Z \) by \( \Delta Y_t \), \( \Delta X_t \), and \( Z_t \). Conditional on the first observation, the following Cochrane–Orcutt transformation removes the autocorrelation from the model and transforms the remainder disturbance into a standard normal random variable:

\[
\frac{\Delta Y^{(A)}_t}{\sigma_e} = \frac{(\Delta X^{(A)}_t) - \rho(\Delta X^{(A)}_{t-1})}{\sigma_e} \beta + \frac{(Z_t - \rho Z_{t-1})}{\sigma_e} \gamma + \frac{\varepsilon_t}{\sigma_e}
\]  

(7)

Define

\[
f_t(\Delta Y_t, \theta) = \frac{1}{\sigma_e} \left[ (\Delta Y^{(A)}_t) - (\Delta X^{(A)}_t) - \rho(\Delta X^{(A)}_{t-1}) - (Z_t - \rho Z_{t-1}) \right]
\]

and

\[
k_t = \log \left| \frac{\partial f_t(\Delta Y_t, \theta)}{\partial \Delta Y_t} \right| = (\lambda - 1) \log Y_{t-1} - \log \sigma_e
\]

(9)

where \( \theta = (\beta', \gamma', \sigma_e, \lambda, \rho)' \). Let

\[
F_t(\Delta Y_t, \theta) = \frac{\partial f_t(\Delta Y_t, \theta)}{\partial \theta_t} \quad \text{and} \quad K_t(\Delta Y_t, \theta) = \frac{\partial k_t(\Delta Y_t, \theta)}{\partial \theta_t}
\]

(10)

and define \( F(\Delta Y, \theta) \) and \( K(\Delta Y, \theta) \) as the \((T - 2) \times (K + S + 3)\) matrices with typical elements \( F_{it}(\Delta Y_t, \theta) \) and \( K_{it}(\Delta Y_t, \theta) \) for \( t = 3, \ldots, T \) and \( i = 1, \ldots, K + S + 3 \). Similarly, let \( f(y, \theta) \) be the \((T - 1) \times 1\) vector with typical element \( f_t(\Delta Y_t, \theta) \).

The DLR may be written as an artificial regression with \(2(T - 2)\) observations:

\[
\begin{bmatrix} f_t(\Delta Y_t, \theta) \\ y_{t-2} \end{bmatrix} = \begin{bmatrix} -F(\Delta Y, \theta) \\ K(\Delta Y, \theta) \end{bmatrix} b + \text{residuals}
\]

(11)

where \( y_{t-2} \) is a vector of ones of dimension \( T - 2 \). The test statistic is the explained sum of squares of (11) when the latter is evaluated at the restricted MLE under the null hypothesis. This DLR statistic has the same score form as the LM statistic. It is asymptotically distributed under the null as \( \chi^2_{2} \), see Davidson and MacKinnon (1993).

To jointly test \( H_0' : \lambda = 0 \) and \( \rho = 0 \), we first estimate the model under the null hypothesis. This yields OLS on (6) with \( \lambda = 0 \), i.e. with the \( Y \) and the \( X \) variables expressed in percentage form

\[
\frac{\Delta Y_t}{Y_{t-1}} = \sum_{k=1}^{K} \frac{\Delta X_{tk} B_k}{X_{t-1,k}} + \sum_{s=1}^{S} Z_{ts} \gamma_s + \varepsilon_t.
\]

(12)

Denote the resulting estimates by \( \hat{\beta} \), \( \hat{\gamma} \), and \( \hat{\sigma}^2_e \) where \( \hat{\sigma}^2_e = \sum_{t=3}^{T} \hat{\varepsilon}_t^2 / (T - 2) \) with \( \hat{\varepsilon}_t \) denoting the OLS residual from (12). The regressand of the DLR has a typical element \( \hat{\varepsilon}_t / \hat{\sigma}_e \) for the first \( T - 2 \) observations and 1 for the next \( T - 2 \) observations. The typical elements for the first and second \( T - 2 \) observations for the regressors are given by the following:
\[- F(\Delta Y, \theta) \quad K(\Delta Y, \theta) \]

for \( \beta_k \):
\[
\frac{1}{\hat{\sigma}_\epsilon} \Delta X_{ik}^{(0)} = \frac{1}{\hat{\sigma}_\epsilon} X_{i-1,k} \]
and 0;

for \( \gamma \):
\[
\frac{1}{\hat{\sigma}_\epsilon} Z_{is} \]
and 0;

for \( \sigma_\epsilon \):
\[
\hat{\epsilon}_i \quad \hat{\epsilon}_i \]
and \( - \frac{1}{\hat{\sigma}_\epsilon} \);

for \( \lambda \):
\[
- \frac{1}{\hat{\sigma}_\epsilon} \left[ \frac{\Delta Y_{i-1, \log(Y_{i-1})}}{Y_{i-1}} - \sum_{k=1}^K \frac{\Delta X_{ik, \log(X_{i-1,k})} \hat{\beta}_k}{X_{i-1,k}} \right] \]
and \( \log(Y_{i-1}) \);

for \( \rho \):
\[
\frac{1}{\hat{\sigma}_\epsilon} \left[ \frac{\Delta Y_{i-1}}{Y_{i-2}} - \sum_{s=1}^S \frac{\Delta X_{i-1,s} \hat{\gamma}_s}{X_{i-2,s}} - \sum_{s=1}^S Z_{i-1,s} \hat{\gamma}_s \right] \]
and 0.

The test statistic is the explained sum of squares of this regression, or equivalently, it is \( 2(T - 2) \) — the residual sum of squares of this regression. It is asymptotically distributed as \( \chi^2 \) under \( H_0^a \).

Similarly, one can perform the joint test for \( H_0^L; \lambda = 1 \) and \( \rho = 0 \). First, one estimates the model under the null hypothesis. This yields OLS on (6) with \( \lambda = 1 \), i.e. with the \( Y \) and the \( X \) variables expressed in first difference form

\[
\Delta Y_i = \sum_{k=1}^K \Delta X_{ik} \beta_k + \sum_{s=1}^S Z_{is} \gamma + \epsilon_i \quad (13)
\]

and the resulting estimates are used to obtain the DLR statistic as described below (12).

### 3. BCD with heteroskedasticity

The model is given by (6) with \( \sigma_{u_{it}} = h(\alpha + W_i \hat{\delta}) \) where \( h(.) \) is the skedastic function and \( W \) is of dimension \( T \times L \) and may include \( X \) or \( Z \) variables. To implement the tests in this paper, we do not have to know the functional form \( h(.) \). If \( \hat{\delta} = 0 \), then \( \sigma_{u_{it}} = h(\alpha) \) and the model is homoskedastic.

In order to apply the DLR, we transform the model such that the remainder disturbance reduces to a standard normal random variable

\[
\frac{\Delta Y_i^{(A)}}{h(\alpha + W_i \hat{\delta})} = \frac{\Delta X_i^{(A)} \beta}{h(\alpha + W_i \hat{\delta})} + \frac{Z_i \gamma}{h(\alpha + W_i \hat{\delta})} + \frac{u_i}{h(\alpha + W_i \hat{\delta})}. \quad (14)
\]

Define

\[
f_i(\Delta Y_i, \theta) = \frac{1}{h(\alpha + W_i \hat{\delta})} [\Delta Y_i^{(A)} - \Delta X_i^{(A)} \beta - Z_i \gamma] \quad (15)
\]

and

\[
k_i = \log \left| \frac{\partial f_i(\Delta Y_i, \theta)}{\partial \Delta Y_i} \right| = (\lambda - 1) \log Y_{i-1} - \log h(\alpha + W_i \hat{\delta}) \quad (16)
\]
where \( \theta = (\beta', \gamma', \lambda, \alpha, \delta)' \). In this case \( F(\Delta Y, \theta) \) and \( K(\Delta Y, \theta) \) are the \((T - 1) \times (K + S + L + 2)\) matrices with typical elements \( F_{it}(\Delta Y, \theta) \) and \( K_{it}(\Delta Y, \theta) \) for \( t = 2, \ldots, T \) and \( i = 1, \ldots, K + S + L + 2 \) defined in (10).

To jointly test \( H^c_0: \lambda = 0 \) and \( \delta = 0 \), we estimate the model under the null hypothesis. This yields OLS on (6) with \( \lambda = 0 \), i.e. with the \( Y \) and the \( X \) variables expressed in percentage form, see (12). The regressand of the DLR has a typical element \( \hat{c}/s^2 \) for the first \( T - 1 \) observations and 1 for the next \( T - 1 \) observations where \( \hat{c}^2 = \sum_{t=2}^{T} \hat{e}_t^2/(T - 1) \) and \( \hat{e}_t \) is the OLS residual from (12). The typical elements for the first and second \( T - 1 \) observations for the regressors are given by the following:

\[
F(\Delta Y, \theta) = \frac{1}{\hat{\sigma}_e} \Delta X_{ik}^{(0)} = \frac{1}{\hat{\sigma}_e} \frac{\Delta X_{ik}}{X_{t-1,k}} \quad \text{and} \quad 0
\]

\[
K(\Delta Y, \theta) = \frac{1}{\hat{\sigma}_e} Z_{ts} \quad \text{and} \quad 0
\]

\[
\lambda = -\frac{1}{\hat{\sigma}_e} \left[ \frac{\Delta Y_t \log(Y_{t-1})}{Y_{t-1}} - \sum_{i=1}^{K} \Delta X_{ik} \log(X_{t-1,k}) \hat{\beta}_k \right] \quad \text{and} \quad \log(Y_{t-1})
\]

\[
\alpha: \frac{h'(\hat{\alpha})}{\hat{\sigma}_e} \hat{e}_t \quad \text{and} \quad -\frac{h'(\hat{\alpha})}{\hat{\sigma}_e}
\]

\[
\delta: \frac{h'(\hat{\delta})}{\hat{\sigma}_e} W_{tr} \hat{e}_t \quad \text{and} \quad -\frac{h'(\hat{\delta})}{\hat{\sigma}_e} W_{tr}
\]

where \( \hat{\beta} \) and \( \hat{\gamma} \) are the OLS estimates of \( \beta \) and \( \gamma \) from (12). Because \( h'(\hat{\alpha}) \) is a constant, the typical elements for \( \alpha \) and \( \delta \) can be rewritten as

\[
\alpha: \frac{1}{\hat{\sigma}_e} \hat{e}_t \quad \text{and} \quad -\frac{1}{\hat{\sigma}_e}
\]

\[
\delta: \frac{1}{\hat{\sigma}_e} W_{tr} \hat{e}_t \quad \text{and} \quad -\frac{1}{\hat{\sigma}_e} W_{tr}
\]

so that \( h'(\hat{\alpha}) \) drops out and the skedastic function \( h(.) \) need not be known to compute the DLR. The test statistic is the explained sum of squares of this regression. This DLR statistic is asymptotically distributed as \( \chi^2_{L+1} \) under \( H^c_0 \).

Similarly, to jointly test \( H^d_0: \lambda = 1 \) and \( \delta = 0 \), we first estimate the model under the null hypothesis. This yields OLS on (6) with \( \lambda = 1 \), i.e. with the \( Y \) and the \( X \) variables expressed in first difference form, see (13). The resulting estimates are used to obtain the DLR statistic as described below (16).

4. Empirical example

For illustrative purposes, we apply these DLR tests to a simple demand relationship for liquor (Table 1). This is based on the data in Baltagi and Griffin (1995) and covers 43 states over the period 1959–1982. Consumption of liquor is measured in gallons per capita. This is regressed on the price of liquor, which is a Paasche index based on nine leading brands, and the real per capita disposable
Table 1
DLR test results: liquor demand

<table>
<thead>
<tr>
<th></th>
<th>$H^0_{0i}; \lambda = 0$, $\rho = 0$</th>
<th>$H^{1i}_{0i}; \lambda = 1$, $\rho = 0$</th>
<th>$H^{0i}_{0i}; \lambda = 0$, $\delta = 0$</th>
<th>$H^{0i}_{0i}; \lambda = 1$, $\delta = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>4.24 (0.12)</td>
<td>1.15 (0.56)</td>
<td>5.46 (0.07)</td>
<td>1.28 (0.53)</td>
</tr>
<tr>
<td>Arkansas</td>
<td>5.32 (0.07)</td>
<td>8.03 (0.02)</td>
<td>1.54 (0.46)</td>
<td>6.10 (0.04)</td>
</tr>
<tr>
<td>Minnesota</td>
<td>14.90 (0.00)</td>
<td>6.48 (0.04)</td>
<td>16.86 (0.00)</td>
<td>6.97 (0.03)</td>
</tr>
</tbody>
</table>

* Asymptotic $P$-values are given in parentheses.

income. For the state of Alabama, we do not reject any of the hypotheses under consideration. This means that the differenced linear model and the differenced loglinear model with no autocorrelation or no heteroskedasticity are not rejected. Heteroskedasticity was also modeled as a function of real per capita disposable income. For the state of Arkansas, the differenced loglinear models without autocorrelation or heteroskedasticity are not rejected at the 0.05 significance level, while the differenced linear models without autocorrelation or heteroskedasticity are rejected. For the state of Minnesota, all hypotheses considered are rejected at the 0.05 significance level. Therefore, the differenced linear or loglinear models without serial correlation or heteroskedasticity are rejected.

This example is illustrative, but it shows how simple it is to apply the DLR for jointly testing for differenced functional form and serial correlation or heteroskedasticity. The choice between differenced models and percentage change models is frequently encountered in macro models and finance models. The joint DLR tests proposed in this paper should prove useful in guarding against the possible presence of serial correlation or heteroskedasticity.

**Acknowledgements**

The authors acknowledge the financial support from the Private Enterprise Research Center.

**References**