Economics Bulletin

Volume 35, Issue 2

Robust estimation based on the third-moment restriction of the error terms for the Box-Cox transformation model: An estimator consistent under heteroscedasticity

Kazumitsu Nawata
University of Tokyo

Abstract

The Box-Cox (1964) transformation model is widely used in various fields of econometrics and statistics. Generally, the maximum likelihood estimator under the normality assumption (BC MLE) is used. However, the BC MLE is not consistent under heteroscedasticity, even if the "small sigma" assumption is satisfied. Here I propose a new robust estimator of the Box-Cox transformation model. The estimator is based on only the first- and third-moment restrictions of the error terms, and it is consistent even under heteroscedasticity. Moreover, it can be easily calculated by the least-squares and scanning methods. The asymptotic distribution of the proposed estimator was obtained, and the results of Monte Carlo experiments are presented.

This study was supported by a Grant-in-Aid for Scientific Research "Analyses of Japanese Medical Information and Policy Using Large-scale Individual-level Survey Data" (No. 20243016 and No. 24330067) of the Japan Society of Science. The author wishes to thank an anonymous referee for helpful comments and suggestions. The author also would like to thank the participants of the International Congress on Modelling and Simulation 2013 (MODSIM 2013) in Adelaide, Australia for their helpful comments.
Citation: Kazumitsu Nawata, (2015) "Robust estimation based on the third-moment restriction of the error terms for the Box-Cox transformation model: An estimator consistent under heteroscedasticity", *Economics Bulletin*, Volume 35, Issue 2, pages 1056-1064
Contact: Kazumitsu Nawata - nawata@tmi.t.u-tokyo.ac.jp.
Submitted: July 04, 2014. **Published:** April 22, 2015.

1. Introduction

The Box-Cox (1964) transformation model (hereafter, the BC model) is widely used in various fields of econometrics and statistics. Generally, the maximum likelihood estimator under the normality assumption (BC MLE) is used. The BC MLE is consistent if the error terms are independent and identically distributed (i.i.d.) random variables, and the "small σ" assumption, described in Bickel and Doksum (1981) and Nawata and Kawabuchi (2014), is satisfied. However, the BC MLE is not consistent under heteroscedasticity even if the "small σ" assumption is satisfied. Although Foster, Tain, and Wei (2001) [1)] and Nawata (2013) proposed semiparametric estimators, these estimators are also not consistent under heteroscedasticity. Powell (1996) proposed a semiparametric estimator based on the moment restriction. Powell's estimator is consistent under heteroscedasticity, however, it has the following problems: (i) to identify the transformation parameter, it is necessary to introduce one or more instrumental variables, w_t, which satisfy $E(w_t \cdot u_t) = 0$, where u_t is the error term, and the result of the estimation depends on the selection of instrumental variables; (ii) a rather arbitrary rescaling of the dependent variable is necessary, and the result also depends on a rescaling method; and (iii) its finite-sample properties are not good and the estimator often performs poorly, as demonstrated in Monte Carlo experiments.

Here I propose a new robust estimator of the Box-Cox transformation model. The estimator is based on only the first- and third-moment restrictions of the error terms, and does not require the assumption of a specific distribution. The estimator is consistent even under heteroscedasticity. In the present study, its asymptotic distribution was obtained, and the results of Monte Carlo experiments are presented.

2. Model

We consider the BC model,

$$z_t = x_t'\beta + u_t, \qquad y_t^{\lambda} > 0 \tag{1}$$

$$z_t = \begin{cases} (y_t^{\lambda} - 1)/\lambda, & \text{if } \lambda \neq 0, \\ \log(y_t), & \text{if } \lambda = 0, \end{cases} \qquad t = 1,2,\ldots,T,$$

where x_t and β are the k-th dimensional vectors of the explanatory variables and the coefficients, respectively, and λ is the transformation parameter. $\{x_t\}$ and $\{u_t\}$ do not have to be i.i.d. random variables, and heteroscedasticity can be assumed. The following assumptions are made:

Assumption 1. $\{(x_t, u_t)\}$ are independent but not necessarily identically distributed. The distribution of u_t may depend on x_t.

Assumption 2. u_t follows distributions in which the supports are bounded from below; that is, $f_t(u) = 0$ if $u \le -a$ for some $a > 0$, where $f_t(u)$ is the probability (density) function. For any t, the following moment conditions are satisfied: (i) $E(u_t \mid x_t) = 0$, (ii) $E(u_t^3 \mid x_t) = 0$, and (iii) $\delta_1 < E(u_t^6 \mid x_t) < \delta_2$ for some $0 < \delta_1 < \delta_2 < \infty$.

Assumption 3. $\{x_t\}$ are independent, and its fourth moments are finite. The distributions of $\{x_t\}$ and the parameter space of β are restricted so that $\inf_x (x'\beta_0) > a$ and $\inf_{x,\beta}(x'\beta) > c$ for some $c > 0$ in the neighborhood of β_0 where β_0 is the true parameter value of β.

Here, we use the first- and third-moment restrictions and consider the roots of the equations

$$G_T(\theta) = \sum_t g_t(\theta) = 0, \quad g_t(\theta) \equiv g(\theta, x_t, y_t) = (z_t - x_t'\beta)^3 \text{ and } \sum_t x_t(z_t - x_t'\beta) = 0, \quad (2)$$

where $\theta' = (\lambda, \beta)$. Although Nawata (2013) considered more complicated moment restrictions for $G_T(\theta)$ obtained by the modification of the likelihood function of the BC model, a simple third-order moment restriction is used in this paper. Note that the last equation in (2) gives the least-squares estimator when the value of λ is given. Let $\theta_0' = (\lambda_0, \beta_0)$ be the true parameter value of θ. Since $E[G(\theta_0)] = 0$, we obtain the following proposition:

Proposition 1

Among the roots of (2), there exists a consistent root.

The proof is given in Appendix A. Let $\hat{\theta}' = (\hat{\lambda}, \hat{\beta}')$ be the consistent root. The asymptotic distribution of $\hat{\theta}$ is obtained by the following proposition.

Proposition 2

Let $\xi_t(\theta) = x_t(z_t - x_t'\beta)$ and $\ell_t(\theta)' = [g_t(\theta), \xi_t(\theta)']$. Suppose that $\frac{1}{T}\sum_t \frac{\partial \ell_t(\theta)}{\partial \theta'}|_{\theta_0}$ converges to a nonsingular matrix *A* in probability, and that $\frac{1}{T}\sum_t E[\ell_t(\theta_0)\ell_t(\theta_o)']$ converges to a nonsingular matrix *B*. Then the asymptotic distribution of $\hat{\theta}$ is given by

$$\sqrt{T}(\hat{\theta} - \theta_0) \to N[0, A^{-1}B(A')^{-1}], \quad (3)$$

where $A = p\lim_{T\to\infty} \sum_t \frac{\partial \ell_t(\theta)}{\partial \theta'}|_{\theta_0}$, and $B = \lim_{T\to\infty} \frac{1}{T}\sum_t E[\ell_t(\theta_0)\ell_t(\theta_o)']$.

[Proof]

Let

$$\ell(\theta) = \sum_t \ell_t(\theta) = \begin{bmatrix} G_T(\theta) \\ \sum_t \xi_t(\theta) \end{bmatrix}. \quad (4)$$

Then,

$$\sqrt{T}(\hat{\theta}-\theta_0) = -[\frac{1}{T}\frac{\partial \ell}{\partial \theta'}|_{\theta^*}]^{-1}\frac{1}{\sqrt{T}}\ell(\theta_0), \tag{5}$$

where θ^* is some value between $\hat{\theta}$ and θ_0. Here,

$$\ell_t(\theta_0) = \begin{bmatrix} u_t^3 \\ x_t u_t \end{bmatrix}. \tag{6}$$

Therefore, $E[\ell_t(\theta_0)] = 0$. Since the variables $\{\ell_t(\theta_0)\}$ are independent and the Lindeberg condition is satisfied under Assumptions 2 and 3, we obtain

$$\frac{1}{\sqrt{T}}\ell(\theta_0) \to N(0, B), \tag{7}$$

from Theorem 3.3.6 in Amemiya (1985, p. 92).

Since $\partial \ell / \partial \theta = \begin{bmatrix} \frac{3}{\lambda}\sum_t (z_t - x_t'\beta)^2\{z_t \log(y_t) - z_t\} & -3\sum_t (z_t - x_t'\beta)^2 x_t' \\ \frac{1}{\lambda}\sum_t x_t\{z_t \log(y_t) - z_t\} & -\sum_t x_t x_t' \end{bmatrix}$,

$$\frac{1}{T}\frac{\partial \ell(\theta)}{\partial \theta'}|_{\theta^*} \xrightarrow{P} A, \tag{8}$$

from Theorem 4.1.4 in Amemiya (1985, pp. 112-113). From Theorem 4.1.3 in Amemiya (1985, p. 111), the asymptotic distribution of $\hat{\theta}$ is given by Equation (3). Since $\lim_{\lambda_0 \to 0}(y_t^{\lambda_0} - 1)/\lambda_0 = \log y_t$, we can get the asymptotic distribution given by the same formula even when $\lambda_0 = 0$. Let $\hat{\theta}_{BC}' = (\hat{\lambda}_{BC}, \hat{\beta}_{BC}')$ be the BC MLE. Using $\hat{\lambda}$ and $\hat{\lambda}_{BC}$, we can test both the i.i.d. and "small σ" assumptions; that is, we can test whether we can successfully use the BC MLE by the Hausman test. However, since the rank of the asymptotic variance-covariance matrix of $[\sqrt{T}(\hat{\lambda}_{BC} - \hat{\lambda}), \sqrt{T}(\hat{\beta}_{BC} - \hat{\beta})']$ becomes one, we cannot use any element of β in the Hausman test (Nawata and McAleeer, 2014).

3. Monte Carlo Study

In this section, Monte Carlo results are presented for the BC MLE, the newly proposed estimator, and Powell's estimator. The behavior of the estimators under both homoscedasticity and heteroscedasticity is studied. The basic model is:

$$z_t = \beta_1 + \beta_2 x_t + u_t, \quad z_t = (y_t^{\lambda_0} - 1)/\lambda_0, \quad y_t \geq 0, \quad t = 1,2,...,T. \tag{9}$$

The BC MLE and the proposed estimator are calculated using the same scanning method used by Nawata (2013) over $\lambda \in [0.01, 2.0]$. For the proposed estimator, there are two possible problems. They are:

(2) has multiple solutions, and (2) does not have a solution.

However, all trials had just one solution, and the above problems did not occur in the Monte Carlo study.

Since Powell (1996) suggested a function of x_t as the instrument variable w_t, $w_i = 30 - (x_t - 5.0)^2$, is used. Although other types of functions[2] of x_t have been used for w_t, the conclusion of this study does not change; that is, Powell's estimator performs quite poorly. Since heteroscedasticity is also considered, the generalized method of moment (GMM) type estimator is not used. Powell's estimator is obtained by: i) rescaling $\{y_t\}$ by dividing by their geometric mean; ii) minimizing,

$$S = \{\sum_t w_i (z_t^* - \hat{\beta}_0 - \hat{\beta}_1 x_t)\}^2, \tag{10}$$

where z_t^* is the value of the BC transformation using the rescaled value, and $\hat{\beta}_0$ and $\hat{\beta}_1$ are least-squares estimators based on $\{z_t^*\}$; and iii) recalculating the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ to adjust the effects of rescaling. Powell's estimator is also calculated by the scanning method over $\lambda \in [0.01, 2.0]$. There are two possible problems with Powell's estimator. They are:

i) S is not minimized in $\lambda \in (0.01, 2.0)$, and S is minimized on the boundary (i.e., $\lambda = 0.01$ or $\lambda = 2.0$).

ii) S becomes 0 for multiple values of θ.

Unlike with the proposed estimator, these problems were observed in many trials. Since we were not able to obtain accurate values of the estimator in these trials, the results of Powell's estimator were calculated for trials without these problems.

3.1 Under homoscedasticity

In this section, the behavior of the estimators under homoscedasticity is analyzed. The values of 0.2, 0.5 and 0.8 are considered for λ_0. The sample size was 200 in all cases. $\{x_t\}$ represents i.i.d. random variables distributed uniformly on (0, 10). The case in which $\lambda_0 u_t$ is distributed uniformly on (-5, 5) is considered. (Although other types of distributions were also considered, the results were similar to those for this distribution.) The true parameter values are:

$$\beta_1^* \equiv 1 + \lambda_0 \beta_1 = 5.0 \quad \text{and} \quad \beta_2^* \equiv \lambda_0 \beta_2 = 0.1. \tag{11}$$

The number of trials was 1,000 for all cases. The results are presented in Table I. The BC MLE underestimated λ and had a fairly large bias for all cases. Although the standard deviations of the proposed estimator were about 1.7 times larger than those of the BC MLE, the biases of the proposed estimator were much smaller. In terms of the MSE, the proposed estimator was better than the BC MLE. Powell's estimator performed poorly in many trials; especially when λ_0=0.2, we were not able to obtain accurate values with the estimator because of the problems mentioned earlier. Moreover, the standard deviations were much larger than those of the proposed estimator even for trials without problems.

3.2 Under heteroscedasticity

In this section, the effect of heteroscedasticity is analyzed. The values of x_t are chosen as in the previous section. The true parameter values are:

$$\beta_1^* \equiv 1 + \lambda_0 \beta_1 = 2.5 \quad \text{and} \quad \beta_2^* = \lambda_0 \beta_2 = 0.25. \tag{12}$$

The error terms are given by

$$\lambda_0 u_t = (1 + 0.1 \cdot x_t) \times \varepsilon_t, \tag{13}$$

where $\{\varepsilon_t\}$ represents the i.i.d. random variables distributed uniformly on (-2.5, 2.5). As before, the values of 0.2, 0.5 and 0.8 are considered for λ_0. The sample size was 200 in all cases. The results are presented in Table II.

The BC MLE underestimated λ and the biases of the BC MLE were larger than those under homoscedasticity for all cases. These findings coincided with those of a previous study (Showalter, 1994), in which large biases of the BC MLE under heteroscedasticity were reported. The standard deviations of the proposed estimator were about 2.5 times larger than those of the BC MLE. However, the biases of the proposed estimator were much smaller than those of the BC MLE. As a result, in terms of the MSE, the proposed estimator was better than the BC MLE. Again, Powell's estimator performed poorly. In many trials, we were not able to obtain accurate values with the estimator. The standard deviations were much larger than those of the newly proposed estimator even for trials without problems, again especially when λ_0=0.2.

4. Conclusion

Although the BC model is widely used in various fields, BC MLE is not consistent under heteroscedasticity even if the "small σ" assumption is satisfied. In this paper, a new robust estimator of the BC model was proposed. The estimator was based on the first- and third-moment restrictions of the error terms. The estimator was consistent even under heteroscedasticity, and its asymptotic distribution was also obtained. Moreover, the estimator was easily calculated using the least-squares and scanning methods. The results of the Monte Carlo experiments showed the superiority of the proposed estimator over the BC MLE and Powell's estimator. However, the performance of each of these estimators may depend on the model, and the findings may differ in other models. Further investigation is thus necessary to determine the conditions under which the proposed estimator shows superiority. Wooldridge (1992) proposed an alternative model considering $E(y_t \mid x_t)$. Although his model does not belong to the same category as the BC model, it can be used as an alternative to the BC model in an empirical study. A comparison of these models is another important subject for future research.

Appendix A: Proof of Proposition 1

The proof of the consistency of the estimator is given using a modification of Nawata (2013). When λ is given, β is uniquely estimated by the least-squares method. Let $\hat{\beta}(\lambda)$ be the estimator, and

$$h_T(\lambda) = \frac{1}{T} G_T\{\lambda, \hat{\beta}(\lambda)\} = \frac{1}{T}\sum_t \{z_t - x_t'(\sum_s x_s'x_s)^{-1}(\sum_s x_s z_s)\}^3. \tag{14}$$

The following assumptions were made: that i) $\sum x_t'x_t/T$ converges to a nonsingular matrix in probability; that ii) $\sum x_t z_t/T, \sum x_t z_t^2/T, z_t^3/T$, $\sum x_t z_t \log(y_t)/T$, $\sum z_t^2 \log(y_t)/T$ $\sum \log(y_t) x_t x_t'/T$ and $z_t x_t x_t'/T$ converge to (vectors or a matrix of) continuous functions in probability in the neighborhood of λ_0; and that iii) these functions are differentiable and their first derivatives are continuous in the neighborhood of λ_0. Under these assumptions, $h_T(\lambda)$ and $h_T'(\lambda)$ converge to $h(\lambda)$ and $h'(\lambda)$ in the neighborhood of λ_0. When $\lambda = \lambda_0$, the model becomes an ordinary regression model, and $\hat{\beta}(\lambda_0)$ is consistent. Therefore,

$$h(\lambda_0) = p\lim_{T\to\infty} \frac{1}{T} G_T(\theta_0) = \lim_{T\to\infty} \frac{1}{T} \sum_t u_t^3. \tag{15}$$

Since $E(u_t^3) = 0$, we get

$$h(\lambda_0) = 0\,, \tag{16}$$

by Theorem 3.3.1 of Amemiya (1985, p. 90). Since

$$h_T'(\lambda) = \frac{1}{T}\left[\frac{\partial G_T\{\lambda, \hat{\beta}(\lambda)\}}{\partial \lambda} + \frac{\partial G_T\{\lambda, \hat{\beta}(\lambda)\}}{\partial \hat{\beta}(\lambda)} \frac{\partial \hat{\beta}(\lambda)}{\partial \lambda}\right] \tag{17}$$

$$= \frac{3}{T\cdot\lambda} \sum_t \{z_t - x_t'\hat{\beta}(\lambda)\}^2 \left[\{y_t^\lambda \log(y_t) - z_t\} - x_t'(\sum_s x_s x_s')^{-1} \sum_s \{x_s \log(y_s) y_s^\lambda - z_s\}\right], \text{ if } \lambda \neq 0\text{, and}$$

$$h_T'(\lambda) = \lim_{\lambda\to 0} h_T'(\lambda) = \frac{3}{2\cdot T} \sum_t \{z_t - x_t'\hat{\beta}(\lambda)\}^2 \left[\{\log(y_t)\}^2 - x_t'(\sum_s x_s x_s')^{-1} \sum_s \{\log(y_s)\}^2 x_s\right] \text{ if } \lambda = 0,$$

$h'(\lambda)$ is continuous in the neighborhood of λ_0 and $h'(\lambda_0)$ does not become zero except in very special cases. (Since $h(\theta)$ and $h'(\theta)$ are continuous functions of θ at $\lambda = 0$, we can treat the $\lambda_0 = 0$ case the same as the $\lambda_0 \neq 0$ case.) Therefore, we can assume that $h'(\lambda_0) \neq 0$, and that there exists $\delta > 0$ such that $sign\{h'(\lambda)\} = sign\{h'(\lambda_0)\}$ and $|h'(\lambda)| \geq \gamma \equiv \frac{1}{2}|h'(\lambda_0)| > 0$ if $\lambda \in [\lambda_0 - \delta, \lambda_0 + \delta]$. By the mean value theorem, for any $\varepsilon \in (0, \delta)$,

$$h(\lambda_0 + \varepsilon) = h(\lambda_0 + \varepsilon) - h(\lambda_0) = h'(\lambda^*)\varepsilon \quad \text{and} \quad h(\lambda_0 - \varepsilon) = h(\lambda_0 - \varepsilon) - h(\lambda_0) = -h'(\lambda^{**})\varepsilon \tag{18}$$

where λ^* and λ^{**} are values in $[\lambda_0 - \varepsilon, \lambda_0 + \varepsilon]$. Therefore,

$$sign\{h(\lambda_0 - \varepsilon)\} \neq sign\{h(\lambda_0 + \varepsilon)\}, \qquad |h(\lambda_0 - \varepsilon)| > \gamma\varepsilon, \text{ and } |h(\lambda_0 + \varepsilon)| > \gamma\varepsilon. \tag{19}$$

Since $h_T(\lambda_0 - \varepsilon) \xrightarrow{P} h(\lambda_0 - \varepsilon)$ and $h_T(\lambda_0 + \varepsilon) \xrightarrow{P} h(\lambda_0 + \varepsilon)$,

$$P\left[\ sign\{h_T(\lambda_0 - \varepsilon)\} \neq sign\{h_T(\lambda_0 + \varepsilon)\}, |h_T(\lambda_0 - \varepsilon)| > 0, \text{ and } |h_T(\lambda_0 + \varepsilon)| > 0\ \right] \to 1. \tag{20}$$

From the intermediate value theorem, $h_T(\lambda) = 0$ for some $\lambda \in [\lambda_0 - \varepsilon, \lambda_0 + \varepsilon]$ if $sign\{h_T(\lambda_0 - \varepsilon)\} \neq sign\{h_T(\lambda_0 + \varepsilon)\}$, $|h_T(\lambda_0 - \varepsilon)| > 0$, and $|h_T(\lambda_0 + \varepsilon)| > 0$. Therefore,

$$P[\text{There exists } \hat{\lambda} \text{ such that } h_T(\hat{\lambda}) = 0 \text{ and } \hat{\lambda} \in [\lambda_0 - \varepsilon, \lambda_0 + \varepsilon]\ ] \to 1. \tag{21}$$

Since (21) holds for any $\varepsilon \in (0, \delta)$, $h_T(\lambda) = 0$ has a consistent root of λ_0. Since $\hat{\beta}(\hat{\lambda})$ is obtained by the least-squares method, it is a consistent estimator when $\hat{\lambda} \xrightarrow{P} \lambda_0$. Hence, there exists a consistent root among the roots of (2).

Notes:

1) Their method requires a weight function to calculate the estimator. However, the values of the estimator depend heavily on a weight function. (The details of their estimator are available upon request through the author.)
2) I have also considered the cases where $w_i = x_i^2$, $w_i = (x_i - 5)^2$, $w_i = 25 - (x_i - 5)^2$ and $w_i = 35 - (x_i - 5)^2$. However, the performance of Powell's estimator was still quite poor, even worse than the results presented in this paper.

References

Amemiya, T. (1985) *Advanced Econometrics*, Harvard University Press, Cambridge, MA.

Bickel, P. J. and K. A. Doksum (1981) "An Analysis of Transformations Revisited," *Journal of the American Statistical Association* **76**, 296-311.

Box, G. E. P., and D. R. Cox (1964) "An Analysis of Transformation," *Journal of the Royal Statistical Society B* **26**, 211-252.

Foster, A. M., L. Tain, and L. J. Wei (2001) "Estimation for the Box-Cox Transformation Model without Assuming Parametric Error Distribution," *Journal of the American Statistical Association* **96**, 1097-1101.

Nawata K. (2013) "A new estimator of the Box-Cox transformation model using moment conditions," *Economics Bulletin* **33**, 2287-2297.

Nawata, K. and K. Kawabuchi (2014) "A new test for the Box-Cox transformation model: An analysis of length of hospital stay for diabetes patients in Japan," *Economics Bulletin* **34**, 324-332.

Nawata. K. and M. McAleer (2014) "The maximum number of parameters for the Hausman test when the estimators are from different sets of equations," *Economics Letters* **123**, 291–294.

Powell, J. L. (1996) "Rescaled Method-of-Moments Estimation for the Box–Cox Regression Model," *Economics Letters* **51**, 259–265.

Showalter, M. H. (1994) "A Monte Carlo Investigation of the Box-Cox Model and a Nonlinear Least Squares Alternative," *Review of Economics and Statistics* **76**, 560-570.

Wooldridge, J. (1992) "Some alternatives to the Box-Cox regression model," *International Economic Review* **33**, 935-955.

Table I. Summaries of the estimators under homoscedasticity, sample size= 200

		Mean	STD	Q1	Median	Q3	MSE
BC MLE							
$\lambda_0 = 0.2$	λ	0.1467	0.0161	0.1353	0.1461	0.1577	0.0557
	β_1^*	3.1731	0.5655	2.7799	3.0954	3.4898	1.9124
	β_2^*	0.0517	0.0388	0.0245	0.0502	0.0728	0.0620
$\lambda_0 = 0.5$	λ	0.3707	0.0397	0.3439	0.371	0.3968	0.1353
	β_1^*	3.222	0.5651	2.8256	3.165	3.5623	1.8656
	β_2^*	0.0527	0.0381	0.0258	0.05	0.0767	0.0607
$\lambda_0 = 0.8$	λ	0.5885	0.0658	0.5427	0.5865	0.6295	0.2215
	β_1^*	3.1876	0.6022	2.7621	3.0934	3.5234	1.9098
	β_2^*	0.0523	0.0379	0.0267	0.0493	0.0765	0.0609
	Proposed Estimator						
$\lambda_0 = 0.2$	λ	0.2013	0.0269	0.1828	0.2007	0.2179	0.0269
	β_1^*	5.2786	1.6795	4.1651	5.0064	6.0025	1.7025
	β_2^*	0.1113	0.0876	0.0535	0.0997	0.1576	0.0883
$\lambda_0 = 0.5$	λ	0.5052	0.0682	0.4585	0.503	0.5474	0.0684
	β_1^*	5.3176	1.6269	4.202	5.0359	6.1069	1.6576
	β_2^*	0.1117	0.0945	0.0475	0.0968	0.1593	0.0952
$\lambda_0 = 0.8$	λ	0.7984	0.1064	0.7269	0.7938	0.8614	0.1064
	β_1^*	5.1929	1.614	4.1189	4.8425	5.9076	1.6255
	β_2^*	0.1054	0.0837	0.0493	0.0958	0.1516	0.0839
Powell's Estimator							
$\lambda_0 = 0.2$	λ	0.4272	0.4600	0.1367	0.2293	0.5319	0.5131
(N1=197, N2=0,	β_1^*	6.733E6	5.082E7	2.8199	6.6926	133.16	5.126.E7
N3=36)	β_2^*	1.024E6	7.801E6	0.02508	0.11368	3.5747	7.868E6
$\lambda_0 = 0.5$	λ	0.5975	0.4285	0.3019	0.5015	0.7754	0.5845
(N1=231, N2=29,	β_1^*	58.140	211.72	2.4770	5.0465	14.241	218.28
N3=36)	β_2^*	3.5264	18.674	0.0167	0.0751	0.3557	18.985
$\lambda_0 = 0.8$	λ	0.8007	0.4880	0.4227	0.7396	1.1076	0.7740
(N1=235, N2=79,	β_1^*	10.802	16.457	2.2468	4.3619	10.057	17.451
N3=9)	β_2^*	0.2863	0.7182	0.0106	0.0517	0.1903	0.7420

True values: β_1^*=5.0 and β_2^*=0.1.

Following notations are used: STD, standard deviation; Ql, first quartile; Q3, third quartile; and MSE, mean squared error. For Powell's estimator, following notations are also used: N1, number of trials where S is minimized at $\lambda = 0.01$; N2, number of trials where S is minimized at $\lambda = 2.0$; and N3, number of trials where $S = 0$ becomes 0 at multiple values of θ.

Table II. Summaries of the estimators under heteroscedasticity, sample size= 200

		Mean	STD	Q1	Median	Q3	MSE
BC MLE							
$\lambda_0 = 0.2$	λ	0.1286	0.0108	0.1210	0.1282	0.1353	0.0722
	β_1^*	1.7427	0.1753	1.6247	1.7426	1.8500	0.7773
	β_2^*	0.0972	0.0293	0.0749	0.0951	0.1163	0.1556
$\lambda_0 = 0.5$	λ	0.3206	0.0271	0.3034	0.3207	0.3377	0.1815
	β_1^*	1.7370	0.1832	1.6066	1.7300	1.8417	0.7846
	β_2^*	0.0975	0.0287	0.0778	0.0958	0.1162	0.1551
$\lambda_0 = 0.8$	λ	0.5104	0.0434	0.4791	0.5087	0.5400	0.2928
	β_1^*	1.7307	0.1741	1.6111	1.7168	1.8356	0.7888
	β_2^*	0.0958	0.0299	0.0741	0.0937	0.1136	0.1571
Proposed Estimator							
λ_0 =0.2	λ	0.2033	0.0274	0.1837	0.1998	0.2204	0.0276
	β_1^*	2.5567	0.4472	2.2492	2.508	2.8112	0.4508
	β_2^*	0.2875	0.1579	0.1812	0.2501	0.3509	0.1623
$\lambda_0 = 0.5$	λ	0.5077	0.0685	0.4592	0.4997	0.5507	0.0690
	β_1^*	2.5531	0.4461	2.2448	2.5057	2.8108	0.4493
	β_2^*	0.2865	0.1575	0.1803	0.2488	0.3490	0.1616
$\lambda_0 = 0.8$	λ	0.8045	0.1121	0.726	0.7978	0.8668	0.1122
	β_1^*	2.5201	0.4386	2.206	2.4564	2.7691	0.4391
	β_2^*	0.2807	0.1612	0.1743	0.2476	0.3355	0.1641
Powell's Estimator							
$\lambda_0 = 0.2$	λ	0.3858	0.3366	0.1604	0.2965	0.5422	0.3844
(N1=105, N2=0,	β_1^*	-1.379E5	2.137E6	1.5461	2.3560	6.2745	2.141E6
N3=52)	β_2^*	8.886E4	1.368.E6	0.1504	0.8232	10.731	1.370E6
$\lambda_0 = 0.5$	λ	0.7091	0.5303	0.3274	0.5579	1.0642	0.7351
(N1=124, N2=19,	β_1^*	5.3028	11.467	1.6474	2.5438	6.3440	11.805
N3=50)	β_2^*	8.4842	22.553	0.0971	0.3411	3.3047	24.009
$\lambda_0 = 0.8$	λ	0.7847	0.5219	0.3569	0.6528	1.1570	0.7838
(N1=173, N2=112,	β_1^*	3.2077	2.7245	1.3947	2.0478	3.9121	2.8150
N3=23)	β_2^*	0.7682	1.3651	0.0497	0.1476	0.7274	1.4602

True values: β_1^*=2.5 and β_2^*=0.25.

Following notations are used: STD, standard deviation; Ql, first quartile; Q3, third quartile; and MSE, mean squared error. For Powell's estimator, following notations are also used: N1, number of trials where S is minimized at $\lambda = 0.01$; N2, number of trials where S is minimized at $\lambda = 2.0$; and N3, number of trials where $S = 0$ becomes 0 at multiple values of θ.